NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Larotrectinib for treating NTRK fusion-positive advanced solid tumours [ID1299]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Larotrectinib for treating NTRK fusion-positive advanced solid tumours [ID1299]

Contents:
The **final scope and final stakeholder list** are available on the NICE website.

The following documents are made available to consultees and commentators:

1. **Company submission** from Bayer HealthCare

2. **Clarification questions and company responses**

3. **Patient group, professional group and NHS organisation submission** from:
   a. GIST Support UK
   b. Sarcoma UK- *endorsed by clinical expert Charlotte Benson*
   c. Royal College of Pathologists

4. **Expert personal perspectives** from:
   a. Alistair Reid – clinical expert, nominated by Royal College of Pathologists
   b. Charlotte Benson – clinical expert, nominated by Sarcoma UK
   c. Jayne Bressington- patient expert, nominated by GIST Support UK
   d. Peter Clark – CDF clinical lead

5. **Evidence Review Group report** prepared by Centre for Reviews and Dissemination and Centre for Health Economics – York

6. **Evidence Review Group report – factual accuracy check**

7. **Evidence Review Group report – addendum**

8. **Technical engagement response from company**
   a. Response form
   b. Additional evidence

9. **Technical engagement responses from consultees and commentators:**
   a. GIST Support UK

10. **Evidence Review Group critique of company response to technical engagement** prepared by Centre for Reviews and Dissemination and Centre for Health Economics – York
    a. Critique of technical engagement response from company
    b. Critique of additional evidence

11. Final Technical Report

Any information supplied to NICE which has been marked as confidential, has been redacted. All personal information has also been redacted.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Larotrectinib for treating NTRK fusion-positive advanced solid tumours [ID1299]

Document B

Company evidence submission

June 2019 (August 2019 update)

File name	Version	Contains confidential information	Date
ID1299_Larotrectinib_Document B_August 2019 update_redacted.docx	V1.0	Yes	August 2019

Contents

Contents 2
Tables and figures 3
B.1 Decision problem, description of the technology and clinical care pathway 13
B.2 Clinical effectiveness 38
B.3 Cost effectiveness 135
B.4 References 229
B.5 Appendices 241
Appendix C: Summary of product characteristics (SmPC) 242
Appendix D: Identification, selection and synthesis of clinical evidence 243
Appendix E: Subgroup analysis 312
Appendix F: Adverse reactions 318
Appendix G: Published cost-effectiveness studies 319
Appendix H: Health-related quality-of-life studies 337
Appendix I: Cost and healthcare resource identification, measurement and valuation 349
Appendix J: Clinical outcomes and disaggregated results from the model 361
Appendix K: Checklist of confidential information 363
Appendix L: Survival analysis 364
Appendix M: Detailed methodology for comparator inputs 468
Appendix M-3. Data sources by input parameter for individual tumour locations 573
Appendix N: Utility mapping study based on LOXO-TRK-15002 and LOXO-TRK-15003 576
Appendix O: Summary of available efficacy results from larotrectinib studies (data cut-off 19 February 2018) 604
Appendix P: Response assessment criteria 606
Appendix Q: GMI intra-patient comparison for patients treated with larotrectinib 608

Tables and figures

Table 1. The decision problem 14
Table 2: Technology being appraised 19
Table 3. NCCN guidelines incorporating NTRK screening and larotrectinib therapy guidance 33
Table 4. Larotrectinib – published clinical effectiveness evidence 39
Table 5. Comparative summary of trial methodology for larotrectinib studies 46
Table 6. Efficacy endpoints by study 51
Table 7. Relevant endpoints and measures in the pooled analysis / larotrectinib studies 52
Table 8. Baseline demographic and disease characteristics for pooled analysis and individual larotrectinib study populations (data cut-off 30th July 2018) (45) 56
Table 9. Definition of relevant larotrectinib data analysis sets included within this submission 63
Table 10. Summary of statistical analyses for the pooled analysis from LOXO-TRK-14001, SCOUT and NAVIGATE trials 64
Table 11. Downs and Blacks Checklist score for larotrectinib studies 68
Table 12. Summary of overall response rate (30 July 2018 data cut-off) 73
Table 13. Duration of response by IRC and investigator assessment for larotrectinib pooled analysis sets (30 July 2018 data cut-off) 77
Table 14. Time to first / best response in the pooled analysis (30 July 2018 data cut-off) 79
Table 15. PFS by IRC and investigator assessment for larotrectinib pooled analysis sets (30 July 2018 data cut-off) 82
Table 16. Summary of patients with Patient-Reported Outcome Data by study (ePAS2, 30 July 2018 data cut-off) 86
Table 17. Number of patients with MID improvement (ePAS2, 30 July 2018 data cut-off) 87
Table 18. Summary of baseline and best change from baseline in EORTC QLQ-C30 global health and functioning scores (Adults; ePAS2, 30 July 2018 data cut-off) 87
Table 19. Summary of baseline and best change from baseline in EORTC QLQ-C30 symptoms scores (Adults; ePAS2, 30 July 2018 data cut-off) 88
Table 20. Summary of baseline and best change from baseline in the EQ-5D-5L Visual Analogue Scale Health and Dimensions scores (Adults, ePAS2, 30 July 2018 data cut-off) 89
Table 21. Summary of baseline and best change from baseline in Paediatric quality of Life Inventory scores (Paediatrics age <2 years; ePAS2, 30 July 2018 data cut-off) 91
Table 22. Summary of baseline and best change from baseline in Paediatric quality of Life Inventory functioning and scale scores (Paediatrics age ≥2 years; ePAS2, 30 July 2018 data cut-off) 91
Table 23. GMI for patients with metastatic disease who had received ≥ 1 prior systemic therapy (ePAS2; IRC; July 30 2018 data cut-off) 93
Table 24. Results of sensitivity analyses for GMI 94
Table 25. Overview of Treatment Efficacy in Tumour Histologies where NTRK gene fusions occur 103
Table 26. Overview of Treatment Safety in Tumour Histologies where NTRK gene fusions occur 106
Table 27. TEAEs (all-cause and drug-related) occurring in ≥10% of patients with NTRK fusion-positive cancer in the pooled safety analysis of larotrectinib clinical trials (safety analysis set) 109

Table 28. Grade 3 or 4 TEAEs (all and drug-related) occurring in ≥2% of patients with NTRK fusion-positive cancer in the pooled analysis of larotrectinib clinical trials (safety analysis set)....111
Table 29. Ongoing studies....114
Table 30. Overview of Treatment Efficacy in Tumour Histologies where NTRK gene fusions occur....130
Table 31. End-of-life criteria....134
Table 32 Tumor site weightings in the economic model....144
Table 33. Assessment of larotrectinib vs. non-responder control....149
Table 34: Assessment of self-control comparison with previous line of therapy....150
Table 35. Parametric model coefficients and fit statistics for larotrectinib PFS survival models....161
Table 36. Parametric model coefficients and fit statistics for larotrectinib OS survival curves....162
Table 37. Base case model selection by tumour site....165
Table 38. Adverse events for larotrectinib and the pooled comparator treatment (grade 3 or higher; inclusion determined by weighted rates)....168
Table 39. EQ-5D-3L utility values applied in the base case for larotrectinib....173
Table 40. Summary of utility values for cost-effectiveness analysis....175
Table 41. Summary of utility decrement values for cost-effectiveness analysis....177
Table 42. Larotrectinib drug cost (expected list price)....179
Table 43. Treatment drug cost for intervention and comparators included in the model....180
Table 44. Drug administration costs....182
Table 45. Treatment cost (drug + administration) per cycle (week) by tumour location 183
Table 46. Data source used for HCRU....185
Table 47. Healthcare resource use components and associated unit costs....187
Table 48. Health state costs by tumour location....189
Table 49. NHS reference costs of adverse reactions....191
Table 50. Inputs summary and global inputs table....192
Table 51. Base-case model assumptions....194
Table 52. Conducted scenario analyses....196
Table 53. Basecase cost-effectiveness results....198
Table 54. Deterministic sensitivity analysis....200
Table 55. Probabilistic sensitivity analysis results: Expected list price....203
Table 56. Probability of larotrectinib cost-effectiveness....203
Table 57. Scenario analyses results....206
Table 58. Alternative modelling methods: using non-responding patients as a control 214
Table 59. Previous line of therapy comparison results....217
Table 60. Scenario 1 - survival adjustment for NTRK+ only applied to CRC....219
Table 61. Scenario 2 -survival adjustment for NTRK+ applied to all tumour sites where NTRK incidence (<25%)....219
Table 62. Comparison of base case model and trial outcomes....221
Table 63. Search strategy for evidence on the treatment of NTRK fusion-positive tumours (Medline & Embase)....244
Table 64. Eligibility criteria used in the search strategy....245
Table 65. Clinical evidence for larotrectinib in the treatment of solid NTRK fusion-positive tumours....248
Table 66. Publications excluded at full text stage of literature review on the treatment of NTRK fusion-positive solid tumours....258

Table 67. Search strategy for Natural history / prognosis of NTRK fusion-positive solid tumours (Medline & Embase) 268
Table 68. Eligibility criteria used in the prognosis search strategy 269
Table 69. Prognostic information on NTRK fusion-positive metastatic colorectal cancer 274
Table 70. Prognostic information on NTRK fusion-positive Congenital mesoblastic nephroma (CMN) 276
Table 71. Prognostic information on NTRK fusion positive Papillary Thyroid Carcinoma (PTC) 277
Table 72. Prognostic information on NTRK fusion-positive Lung cancer (NSCLC) 282
Table 73. Prognostic information on NTRK fusion-positive Mammary Analogue Secretory Carcinoma (MASC) 283
Table 74. Publications excluded at full text stage of literature review on prognosis of NTRK positive gene fusion solid tumours count 284
Table 75. Eligibility criteria used in the comparator SLR strategy 293
Table 76. Population / treatment inclusion criteria for comparator SLR for relevant tumours 294
Table 77. Summary of PRISMA flow diagrams of the included studies for tumours included in the comparator evidence SLR (Adapted from Xcenda comparator SLR) ... 300
Table 78. Overview of quality assessment of studies included in comparator SLR 301
Table 79. Patient disposition by study and pooled analysis sets (data cut-off 30 July 2018) 305
Table 80. Quality assessment of larotrectinib studies using Downs and Blacks Checklist 306
Table 81. ORR for larotrectinib according to patient age, and other baseline characteristics (IRC, ePAS2) 313
Table 82. ORR for larotrectinib according to tumour histology (IRC, ePAS2) 314
Table 83. ORR for larotrectinib according to NTRK gene fusion or major NTRK gene isoforms (IRC, ePAS2) 315
Table 84. Tumour sites with no objective response per IRC (ePAS2, 30 July 2018 data cut-off) 316
Table 85. GMI by subgroups (IRC, data cut-off July 2018) 317
Table 86. Search strategy for cost-effectiveness models on the treatment of NTRK fusion-positive tumours (Medline, Embase, EconLit, Congress abstracts) 320
Table 87. Search strategy for cost-effectiveness models for treatment of NTRK fusion-positive tumours (Cochrane Library) 321
Table 88: Inclusion and exclusion criteria used for search strategy for cost-effectiveness models in the treatment of NTRK fusion-positive cancer 322
Table 89. Eligibility criteria used in the cost-effectiveness SLR strategy 327
Table 90. Population and intervention/comparator criteria for Cost-effectiveness SLR for relevant tumours 328
Table 91. Summary of PRISMA flow diagrams of the included studies for tumours included in the cost-effectiveness SLR 332
Table 92. Eligibility criteria used in the SLR strategy 341
Table 93. Population criteria for HRQoL SLR for relevant tumours 343
Table 94. Summary of PRISMA flow diagrams of the included studies for tumours included in the HRQoL SLR 345
Table 95. Eligibility criteria used in the cost and resource utilisation SLR strategy 353
Table 96. Population and intervention/comparator criteria for Cost and resource utilisation SLR for relevant tumours 355

Table 97. Summary of PRISMA flow diagrams of the included studies for tumours included in the economic costs and resource use SLR 357
Table 98. Summary of Survival Outcomes 361
Table 99. Summary of QALY gain by health state 361
Table 100. Summary of costs by health state 362
Table 101. AIC/BIC comparison for models fitted to PFS (N=73) 369
Table 102. AIC/BIC comparison for models fitted to OS (N=73) 373
Table 103. Hazard ratio (HR) and confidence bounds for univariate OS-Cox models 374
Table 104. Time interval analysis of PFS: larotrectinib clinical trial programme 379
Table 105. AIC/BIC comparison for models fitted to PFS 398
Table 106. Estimated PFS by parametric form (years from start of treatment) 399
Table 107. Time interval analysis of OS: larotrectinib clinical trial programme 401
Table 108. AIC/BIC comparison for models fitted to OS 427
Table 109. Distribution of patient age in the larotrectinib clinical trial programme 427
Table 110. Estimated survival by parametric form (years from start of treatment) 428
Table 111. Time interval analysis of TTD: larotrectinib clinical trial programme 431
Table 112. AIC/BIC comparison for models fitted to TTD 450
Table 113. Estimated survival by parametric form (years from start of the model) 450
Table 114. Patient characteristics: Responder/ non-responder groups 452
Table 115. Extended Cox model results for PFS 457
Table 116. Super model results for PFS 457
Table 117. AIC/BIC comparison for models fitted to PFS with response status as a linear effect 457
Table 118. Estimated PFS by parametric form (years from start of treatment) 460
Table 119. Cox PH model results for OS 463
Table 120. AIC/BIC comparison for models fitted to OS with response status as a linear effect 463
Table 121. Estimated OS by parametric form (years from start of treatment) 467
Table 122. NSCLC-survival inputs 469
Table 123. AE rates in the scenario 472
Table 124. Treatment monitoring costs 473
Table 125. AE costs in the scenario 474
Table 126. Breakdown of end of life 475
Table 127. AE disutilities in the scenario 476
Table 128. NSCLC specific inputs 476
Table 129. Salivary gland cancer survival inputs 478
Table 130. AE rates in the scenario 481
Table 131. Treatment monitoring costs 481
Table 132. AE costs in the scenario 483
Table 133. End of life costs 484
Table 134. AE disutilties in the scenario 485
Table 135. Salivary gland cancer specific inputs 485
Table 136. Dosing schedule of drugs within the mixed chemotherapy arm of KEYNOTE-002 486
Table 137. Melanoma survival inputs 488
Table 138. AE rates 491
Table 139. Treatment monitoring costs 491
Table 140. AE cost 493
Table 141. End of life costs 493
Table 142. AE disutility 495
Table 143. Melanoma specific inputs 495

Table 144. Colorectal cancer survival inputs 498
Table 145. AE rates 501
Table 146. Treatment monitoring costs 502
Table 147. AE costs 503
Table 148. End of life costs 503
Table 149. Colorectal cancer AE disutilities 504
Table 150. Colorectal cancer specific inputs 505
Table 151. GIST survival inputs 507
Table 152. Treatment monitoring costs 510
Table 153. End of life costs 512
Table 154. STS adults (GIST) specific inputs 513
Table 155. Parametric model coefficients and fit statistics for the EORTC historical arm in nGIST for OS 515
Table 156. Treatment monitoring costs 517
Table 157. End of life cost 517
Table 158. STS adults (non-GIST) specific inputs 519
Table 159. Parametric model coefficients and fit statistics for the irinotecan + vincristine arm in STS paediatrics for OS and PFS 522
Table 160. AE rates 525
Table 161. Treatment monitoring costs 526
Table 162. AE costs 526
Table 163. End of life cost 527
Table 164. AE disutilities 527
Table 165. STS paediatrics specific inputs 528
Table 166. Parametric model coefficients and fit statistics for TPC in breast cancer for OS and PFS 530
Table 167. AE rates 533
Table 168. Breakdown of treatment monitoring costs 534
Table 169. AE cost 534
Table 170. Breakdown of end of life costs 535
Table 171. AE disutilities 536
Table 172. Breast cancer specific inputs 537
Table 173. Parametric model coefficients and fit statistics for the gemcitabine + cisplatin arm in cholangiocarcinoma for OS and PFS 539
Table 174. AE rates 542
Table 175. Treatment monitoring cost 543
Table 176. AE costs 544
Table 177. End of life cost 544
Table 178. AE disutility 545
Table 179. Cholangiocarcinoma specific inputs 546
Table 180. AE rates 550
Table 181. Treatment monitoring costs 551
Table 182. AE cost 552
Table 183. End of life costs 552
Table 184. AE disutilities 552
Table 185. Pancreatic cancer specific inputs 553
Table 186. Parametric model coefficients and fit statistics for lomustine in gliomas for OS and PFS 556
Table 187. AE rates in the scenario 559
Table 188. Treatment monitoring costs 559
Table 189. AE cost in the scenario 560

Table 190. End of life cost 560
Table 191. AE disutilities in the scenario 561
Table 192. Glioma (CNS) specific inputs 562
Table 193. Parametric model coefficients and fit statistics for placebo in thyroid papillary/follicular for OS and PFS 564
Table 194. Treatment monitoring costs 567
Table 195. End of life cost 568
Table 196. Thyroid follicular and papillary cancer -specific inputs 569
Table 197. Larotrectinib-specific inputs 570
Table 198. Larotrectinib survival distribution parameters used in the PSA 572
Table 199. Data sources by input parameter for individual tumour locations 573
Table 200. Summary of HRQoL Data in the Larotrectinib Clinical Trial Programme 578
Table 201. Summary of Top-Performing Models from Khan et al. (2014) 581
Table 202. Review of Patient Characteristics Across Trials: Full trial population versus patients included in the HRQoL analysis 587
Table 203. EQ-5D-3L index score per cycle, stratified for progression status 595
Table 204. Summary Results 598
Table 205. Goodness-of-fit measures for OLS and mixed models (AIC) using EQ-5D-3L utility values 600
Table 206. Recommended Utility Values for Use in the CEM 603
Table 207. Response Evaluation Criteria in Solid Tumours (RECIST 1.1) Guidelines for Tumour Response 606
Table 208. Overall Response Criteria 606
Table 209. RANO Response Criteria Incorporating MRI and Clinical Factors 607
Table 210. Methodology for post hoc GMI analysis 608

List of figures

Figure 1. Efficacy Results With Larotrectinib in the Integrated Analysis (Investigator Assessment) 22
Figure 2. Efficacy Results With Larotrectinib in the Integrated Analysis by Patient Age (Investigator Assessment) 23
Figure 3. Distribution of NTRK gene fusions across tumour histologies 24
Figure 4. TRK signalling pathways 26
Figure 5. NTRK gene fusion 27
Figure 6. Mechanism of action – larotrectinib (Vitrakvi®) 32
Figure 7. What is a basket trial? 42
Figure 8. Infantile fibrosarcoma case study – reduction in tumour mass 75
Figure 9. Waterfall Plot of Best Change in Tumour Size Based on IRC 75
Figure 10. Waterfall Plot of Best Change in Tumour Size Based on Investigator Assessments (SAS3, July 30 2018 data cut-off) 76
Figure 11. Duration of Response based on IRC Assessment in the ePAS2 (30 July 2018 data cut-off) 78
Figure 12. Swimmer Plot of Time to Response and Overall Treatment Duration (ePAS2, July 30 2018 data cut-off) 80
Figure 13. Swimmer Plot of Time to Response and Overall Treatment Duration (SAS3, July 30 2018 data cut-off) 81
Figure 14. Kaplan-Meier Plot of PFS (ePAS2, IRC assessment; 30 July 2018 data cut-off) 83

Figure 15. Kaplan-Meier Plot of PFS (SAS3, investigator assessment; 30 July 2018 data cut-off) 83
Figure 16. Kaplan-Meier Plot of OS (ePAS2, IRC assessment; 30 July 2018 data cut-off) 84
Figure 17. Kaplan-Meier Plot of OS (SAS3, investigator assessment; 30 July 2018 data cut-off) 85
Figure 18. Waterfall plot of best absolute change from baseline in the EQ-5D-5L Visual Analogue Scale Health score (Adults, ePAS2, 30 July 2018 data cut-off) 90
Figure 19. Waterfall plot of Growth Modulation Index for eligible patients from ePAS2 dataset (IRC; July 30 2018 data cut-off) 93
Figure 20. Case 1: Adult LMNA-NTRK1 Fusion Sarcoma 95
Figure 21. Case 2: Adult ETV6-NTRK3 GIST 96
Figure 22. Case 3: Adult ETV6-NTRK3 Mammary Analogue Secretory Carcinoma of the Salivary Gland (MASC) 97
Figure 23. Case 4: Paediatric TPM3-NTRK1 Papillary Thyroid Carcinoma (PTC) 98
Figure 24. Case 5: Paediatric ETV6-NTRK3 Secretory breast carcinoma (SBC) 99
Figure 25. Larotrectinib arm of partitioned survival model 140
Figure 26. Economic model schematic 142
Figure 27: Parametric model fittings for larotrectinib for PFS and OS across time horizons 159
Figure 28. Larotrectinib PFS curves based on the larotrectinib clinical trial programme patients: KM versus extrapolated curves 161
Figure 29. Larotrectinib OS curves based on the larotrectinib clinical trial programme patients: KM versus extrapolated curves 162
Figure 30. Pooled comparator PFS and OS survival models 166
Figure 31. Disposition of patients and assessments included in utility analysis 171
Figure 32. DSA Tornado diagram 201
Figure 33. PSA convergence, QALYs, costs and ICER 202
Figure 34. Larotrectinib cost-effectiveness acceptability curve (CEAC) 203
Figure 35. Cost-effectiveness plane 204
Figure 36. GMI assessment 215
Figure 37. PRISMA flow diagram of the included studies - NTRK+ solid tumours treatment 247
Figure 38. PRISMA flow diagram of the included studies – NTRK+ solid tumours prognosis 270
Figure 39. PRISMA flow diagram of the included cost-effectiveness studies - NTRK+ solid tumours treatment 323
Figure 40. PFS parametric fittings for N=73 patients 365
Figure 41. Assessment of spline model fitting of PFS (n=73) 367
Figure 42. Assessment of penalized spline function fitting for PFS 368
Figure 43. OS parametric fittings for N=73 patients 370
Figure 44. Assessment of spline model fitting of OS 371
Figure 45. Assessment of penalized spline function fitting for OS 372
Figure 46. Kaplan Meier curve for PFS based on the larotrectinib clinical trial programme 378
Figure 47. Cumulative number of events and censoring: PFS 380
Figure 48. Log cumulative hazard plot of KM curve for PFS 381
Figure 49. Non-parametric smoothing using the LOWESS method: PFS 383
Figure 50. Parametric model fittings: PFS at 40 months 384
Figure 51. Parametric model fittings: PFS at 40 months - exponential 385
Figure 52. Parametric model fittings: PFS at 40 months – Weibull 386

Figure 53. Parametric model fittings: PFS at 40 months – log normal 387
Figure 54. Parametric model fittings: PFS at 40 months – log-logistic 388
Figure 55. Parametric model fittings: PFS at 40 months – generalized gamma 389
Figure 56. Parametric model fittings: PFS at 40 months - Gompertz 390
Figure 57. Parametric model fittings: PFS at 480 months (40 years) 391
Figure 58. Parametric model fittings: PFS at 480 months (40 years) - exponential 392
Figure 59. Parametric model fittings: PFS at 480 months (40 years) – Weibull 393
Figure 60. Parametric model fittings: PFS at 480 months (40 years) – log normal 394
Figure 61. Parametric model fittings: PFS at 480 months (40 years) – log logistic 395
Figure 62. Parametric model fittings: PFS at 480 months (40 years) – generalised gamma 396
Figure 63. Parametric model fittings: PFS at 480 months (40 years) – Gompertz 397
Figure 64. Kaplan Meier curve for OS based on the larotrectinib clinical trial programme 400
Figure 65. Cumulative number of events and censoring: OS 402
Figure 66. Log cumulative hazard plots for OS 403
Figure 67. Non-parametric smoothing using the LOWESS method: OS 405
Figure 68. Parametric model fittings: OS at 40 months 406
Figure 69. Parametric model fittings: OS at 40 months – exponential 407
Figure 70. Parametric model fittings: OS at 40 months – Weibull 408
Figure 71. Parametric model fittings: OS at 40 months – log normal 409
Figure 72. Parametric model fittings: OS at 40 months – log-logistic 410
Figure 73. Parametric model fittings: OS at 40 months – generalised gamma 411
Figure 74. Parametric model fittings: OS at 40 months – Gompertz 412
Figure 75. Parametric model fittings: OS at 480 months (40 years) 413
Figure 76. Parametric model fittings: OS at 480 months (40 years) – exponential 414
Figure 77. Parametric model fittings: OS at 480 months (40 years) – Weibull 415
Figure 78. Parametric model fittings: OS at 480 months (40 years) – log normal 416
Figure 79. Parametric model fittings: OS at 480 months (40 years) – log logistic 417
Figure 80. Parametric model fittings: OS at 480 months (40 years) – generalised gamma 418
Figure 81. Parametric model fittings: OS at 480 months (40 years) – Gompertz 419
Figure 82. Parametric model fittings: OS at 960 months (80 years) 420
Figure 83. Parametric model fittings: OS at 960 months (80 years) – exponential 421
Figure 84. Parametric model fittings: OS at 960 months (80 years) – Weibull 422
Figure 85. Parametric model fittings: OS at 960 months (80 years) – log normal 423
Figure 86. Parametric model fittings: OS at 960 months (80 years) – log logistic 424
Figure 87. Parametric model fittings: OS at 960 months (80 years) – generalised gamma 425
Figure 88. Parametric model fittings: OS at 960 months (80 years) – Gompertz 426
Figure 89. Kaplan Meier curve for TTD based on the larotrectinib clinical trial programme 430
Figure 90. Cumulative number of events and censoring: TTD 432
Figure 91. Log cumulative hazard plots for TTD 433
Figure 92. Non-parametric smoothing using the LOWESS method: TTD 435
Figure 93. Parametric model fittings: TTD at 40 months 436
Figure 94. Parametric model fittings: TTD at 40 months – exponential 437
Figure 95. Parametric model fittings: TTD at 40 months – Weibull 438
Figure 96. Parametric model fittings: TTD at 40 months – log normal 439
Figure 97. Parametric model fittings: TTD at 40 months – log logistic 440
Figure 98. Parametric model fittings: TTD at 40 months – generalised gamma 441

Figure 99. Parametric model fittings: TTD at 40 months – Gompertz 442
Figure 100. Parametric model fittings: TTD at 480 months (40 years) 443
Figure 101. Parametric model fittings: TTD at 480 months (40 years) – exponential .. 444
Figure 102. Parametric model fittings: TTD at 480 months (40 years) – Weibull 445
Figure 103. Parametric model fittings: TTD at 480 months (40 years) – log normal.... 446
Figure 104. Parametric model fittings: TTD at 480 months (40 years) – log logistic.... 447
Figure 105. Parametric model fittings: TTD at 480 months (40 years) – generalised gamma 448
Figure 106. Parametric model fittings: TTD at 480 months (40 years) – Gompertz 449
Figure 107. Log rank test for significant difference in PFS between responders and non-responders..... 455
Figure 108. Schoenfeld test for proportional hazards assumption for response status: PFS 456
Figure 109. Parametric model fittings: PFS stratified by response status (40 months) 458
Figure 110. Parametric model fittings: PFS stratified by response status (40 years) .. 459
Figure 111. Log rank test for significant difference in OS between responders and non-responders..... 461
Figure 112. Schoenfeld test for proportional hazards assumption for response status: OS 462
Figure 113. Parametric model fittings: OS stratified by response status (40 months) . 464
Figure 114. Parametric model fittings: OS stratified by response status (40 years) 465
Figure 115. Parametric model fittings: OS stratified by response status (80 years) 466
Figure 116: Progression free survival model of NSCLC patients based on the placebo arm in NICE TA374: KM versus exponential curve 470
Figure 117. Overall survival model of NSCLC patients based on the placebo arm in NICE TA374: KM versus exponential curve..... 471
Figure 118. Progression free survival model of salivary gland cancer patients based on cisplatin + vinorelbine arm in TA490: exponential curve 479
Figure 119. Overall survival model of salivary gland cancer patients based on cisplatin + vinorelbine arm in TA490: KM versus extrapolated models 480
Figure 120. Progression free survival model of melanoma patients based on the mixed chemotherapy arm in TA357: KM versus piecewise model..... 489
Figure 121. Overall survival model of melanoma patients based on the mixed chemotherapy arm in TA357: KM versus extrapolated models..... 490
Figure 122. Progression free survival model of colorectal cancer patients based on placebo arm in TA405: KM versus extrapolated model..... 499
Figure 123. Overall survival model of colorectal cancer patients based on placebo arm in TA405: KM versus extrapolated model. 500
Figure 124. Progression free survival model of GIST patients based on placebo arm in TA488: KM versus extrapolated model. 508
Figure 125. Overall survival model of GIST patients based on placebo arm in TA488: KM versus extrapolated model. 509
Figure 126. Overall survival model of nGIST patients based on TA185: KM versus extrapolated model. 516
Figure 127. PFS model of STS paediatrics based on the irinotecan + vincristine arm in Mascarenhas et al: KM versus Gompertz model. 523
Figure 128. OS model of STS paediatrics based on the irinotecan + vincristine arm in Mascarenhas et al: KM versus extrapolated models..... 524
Figure 129. Progresion free survival model of breast cancer patients based on TPC arm in TA423: KM model. 531

Figure 130. Overall survival model of breast cancer patients based on TPC arm in TA423: KM model verus piecewise model. 532
Figure 131. PFS model of cholangiocarcinoma patients based on the gemcitabine + cisplatin arm in Valle et al: KM versus extrapolated models. 540
Figure 132. OS model of cholangiocarcinoma patients based on the gemcitabine + cisplatin arm in Valle et al: KM versus extrapolated models. 541
Figure 133. PFS model of pancreas cancer patients based on the 5-FU + leucovorin arm in TA440: KM with exponential tail 548
Figure 134. OS model of pancreas cancer patients based on the 5-FU + leucovorin arm in TA440: KM with exponential tail 549
Figure 135. PFS model of glioma patients based on lomustine arm in Batchelor et al: KM versus extrapolated models 557
Figure 136. OS model of glioma patients based on lomustine arm in Batchelor et al: KM versus extrapolated models 558
Figure 137. PFS model of thyroid papillary/follicular cancer patients based on placebo arm in TA535: KM versus piecewise models. 565
Figure 138. OS model of thyroid papillary/follicular cancer based on placebo arm in TA535: KM versus piecewise models. 566
Figure 139. Disposition of patients and assessments included in utility analysis 585
Figure 140. Histograms of Assessments During Pre-Progression by Trial 592
Figure 141. Histograms of Assessments Post-Progression by Trial 593

B.1 Decision problem, description of the technology and clinical care pathway

B.1.1 Decision problem

The submission covers the technology's full marketing authorisation for this indication.

Table 1. The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	People with NTRK fusion-positive advanced solid tumours who: • have either progressed on or not responded to prior therapies • are unfit for chemotherapy or for whom no curative therapy exists	[redacted]	In line with the anticipated marketing authorisation.
Intervention	Larotrectinib	Larotrectinib	

Comparator(s)	Established management without larotrectinib	The comparator selected is in line with the final scope issued by NICE. There are no treatment options available for patients that specifically target NTRK gene fusion cancers. The approach taken to identifying the comparator is to consider standard of care after patients have exhausted all satisfactory treatment options, [redacted] [redacted]. The comparators identified are specific to tumour sites, meaning there are a number of relevant comparators that need to be considered in this appraisal. We have weighted these by patient enrolment per tumour location in the clinical trials. In the absence of any data after the final line of approved active treatment, we use a proxy such as the last line of active treatment.	
Outcomes	The outcome measures to be considered include: • overall survival • progression-free survival • response rate • duration of response • adverse effects of treatment • health-related quality of life.	The outcome measures considered include: • overall survival • progression-free survival • response rate • duration of response • adverse effects of treatment • health-related quality of life.	

Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year.	Cost-effectiveness is expressed in terms of incremental cost per quality-adjusted life year. The company submission adopts a cost-utility approach using partitioned survival analysis and adheres as closely as possible to the reference case and previously accepted submission approaches.	
	The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared.	The time horizon for estimating clinical and cost effectiveness is lifetime.	
	Costs will be considered from an NHS and Personal Social Services perspective.	Costs are considered from an NHS and Personal Social Services perspective.	
	The use of larotrectinib is conditional on the presence of NTRK fusion. The economic modelling should include the costs associated with diagnostic testing for NTRK fusion in people with advanced solid tumours who would not otherwise have been tested. A sensitivity analysis should be provided without the cost of the diagnostic test. See section 5.9 of the Guide to the Methods of Technology Appraisals.	Testing costs are not included within the model as patients will be tested routinely according to NHS plans.	While testing for NTRK gene fusions is currently only carried out for some cancer patients in the UK, genomics is identified as transformative and an area of innovation in the NHS Long Term Plan, with an aim for the NHS to be the first national health care system to offer whole genome sequencing (WGS) as part of routine care. As WGS delivers a comprehensive view of the whole genome, then one test can

			provide information about multiple targets, not just the one under investigation. As such, it is not appropriate in cost-effectiveness modelling of a single innovation, to assign the cost of the test to that treatment, as there will be wider healthcare benefits of WGS. Furthermore the methods guide states that *"… If a diagnostic test to establish the presence or absence of this biomarker is carried out solely to support the treatment decision for the specific technology, the associated costs of the diagnostic test should be incorporated into the assessments of clinical and cost effectiveness*." Since other technologies for the treatment of people with NTRK fusions will be available any diagnostic testing will not be solely to support larotrectinib.
Subgroups to be considered	If evidence allows, subgroup analyses by: • tumour site • previous therapy will be considered.	No subgroups are considered in the cost-effectiveness analysis.	Larotrectinib acts as a histology independent precision medicine. As such, there are no subgroups that are considered in the cost-effectiveness analysis. Whilst supportive analyses was carried out to assess for consistency across selected subgroups, the study designs and patient numbers do not allow for any robust conclusions to be drawn from these analyses.

Special considerations including issues related to equity or equality		Larotrectinib is an innovative technology and a 'step change' in the management of NTRK fusion-positive cancer. It is a precision medicine, designed to selectively target NTRK fusion cancer, providing a specific treatment for NTRK fusion-positive solid tumours where previously no treatment was available. Larotrectinib can be described as a histology independent therapy, providing a treatment for adults and children within one indication. **Importantly, if approved by EMA, larotrectinib will be the first histology independent drug approval in the EU**. As a rare disease, data come from single arm basket studies, considered best suited to assess the efficacy of targeting genomic alterations that occur at low frequencies across a wide variety of tumour types (1). Indeed, the MHRA agreed that a single arm study was an appropriate design to support an MAA, given the extreme rarity of NTRK fusion cancers. Indication. There is no precedence or guidance for evaluating the cost effectiveness of histology independent treatments where activity and clinical evidence is not confined to a particular tumour location. A publication by Love-Koh et al 2018(2) raised a number of relevant challenges for the evidence analysis of precision medicines such as an increasing use of new trial designs that involve smaller populations, complex clinical pathways, high numbers of comparators and the difficulty in obtaining head-to-head estimates of comparative effectiveness. This led to the conclusion that HTA bodies will need to adapt their methods and processes to facilitate evaluation. The company submission adheres as closely as possible to the NICE reference case. On-going workstreams to understand the suitability of HTA methods to precision medicines suggest that there will inevitably be deviations from the reference case. The trial design and high number of comparators are inherent to the rarity of the gene fusion and the innovative nature of this product and need to be taken into consideration in such a way that patients with rare gene fusions are not inequitably disadvantaged. Eligible patients should not be denied access whilst HTA methods evolve.

B.1.2 Description of the technology being appraised

In appendix C include the summary of product characteristics or information for use, and the European public assessment report, scientific discussion or drafts. **Please note – the summary of product characteristics is draft, pending finalisation of the marketing authorisation application process.**

Table 2: Technology being appraised

UK approved name and brand name	Larotrectinib (Vitrakvi)
Mechanism of action	Larotrectinib is an orally bioavailable, adenosine triphosphate (ATP)-competitive and highly selective tropomyosin receptor kinase (TRK) inhibitor, rationally designed to avoid activity with off-target kinase. The target for larotrectinib is the TRK family of proteins inclusive of TRKA, TRKB, and TRKC that are encoded by NTRK1, NTRK2 and NTRK3 genes, respectively (see section 1.3). In-frame gene fusion events resulting from chromosomal rearrangements of the human genes NTRK1, NTRK2, and NTRK3 lead to the formation of oncogenic TRK fusion proteins. These resultant novel chimeric oncogenic proteins are aberrantly expressed, driving constitutive kinase activity subsequently activating downstream cell signalling pathways involved in cell proliferation and survival and leading to TRK fusion cancer. In a broad panel of purified enzyme assays, larotrectinib inhibited TRKA, TRKB, and TRKC with half maximal inhibitory concentration (IC50) values between 5-11 nM. The only other kinase activity occurred at 100-fold higher concentrations (3). Results from the two Phase 1/2-studies (4, 5), which also included patients without NTRK gene fusions showed high objective response rates in patients with tumours harbouring the target, i.e. a TRK fusion protein (as evidenced by NTRK gene fusions), while almost no responses were seen in patients without the target, thus supporting the proposed mechanism of action. A relatively favourable safety profile may be interpreted as clinical support of a low degree of off-target effects (see section B.2.10).
Marketing authorisation/CE mark status	• [redacted] • [redacted]
Indications and any restriction(s) as described in	[redacted]

the summary of product characteristics (SmPC)	[redacted] [pending definitive EC decision].
Method of administration and dosage	Larotrectinib is available as hard capsules (25mg, 100mg) to be taken orally, or as an oral solution (20mg/mL). • Adults: 100 mg larotrectinib, twice daily, until disease progression or unacceptable toxicity occurs. • Paediatric population: Dosing in paediatric patients is based on body surface area (BSA). 100 mg/m^2 larotrectinib, twice daily with a maximum of 100 mg per dose until disease progression or unacceptable toxicity occurs.
Additional tests or investigations	The presence of an NTRK gene fusion in a tumour specimen should be confirmed by a validated test prior to initiation of treatment with larotrectinib. Please refer to section B.1.3 on diagnosis.
List price and average cost of a course of treatment	The proposed NHS list price is £[redacted] for a 30 day supply for an adult. The dose and therefore cost for a paediatric patient is based on body surface area, with a proposed price per mg of £[redacted].
Patient access scheme (if applicable)	Bayer has applied for a Patient Access Scheme to PASLU, representing a simple discount.

B.1.3 Health condition and position of the technology in the treatment pathway

It is anticipated that larotrectinib will be licensed for [redacted]

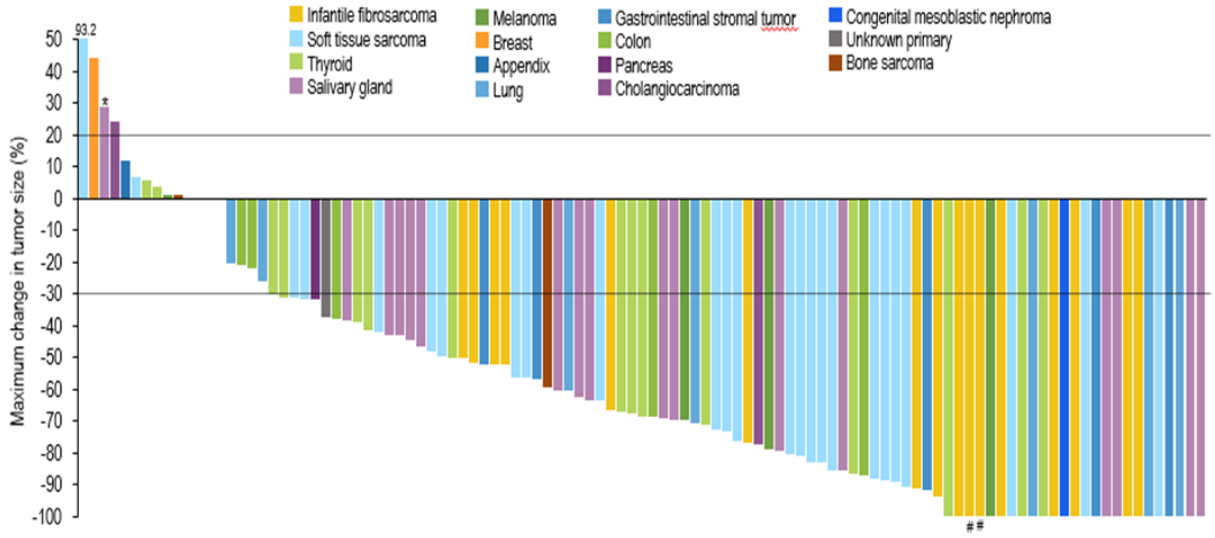
NTRK fusion cancer

Neurotrophic Tyrosine Kinase (NTRK) gene fusion as a primary oncogenic driver and underlying cause of cancer is known to occur across a diverse range of solid tumour sites, affecting both adult and paediatric patients (6-10). A systematic review, identified limited published data on the prognosis of patients with the NTRK gene fusion; only six publications in three tumour sites included a comparison with patients without the NTRK gene fusion. The presence of an NTRK gene fusion has been shown to be associated with a worse prognosis or more aggressive tumour in patients with metastatic colorectal cancer (mCRC), and papillary thyroid carcinoma (PTC) (11, 12). Patients with cellular CMN featuring an NTRK gene appeared to have a better prognosis than cellular CMN without an NTRK fusion (13).

Aligned with the anticipated indication for larotrectinib, this overview focuses on '**NTRK fusion-positive cancer**' as the disease entity, as opposed to individual discussion of each of the multiple tumour sites known to harbour NTRK gene fusions and treatable by larotrectinib. This is because the selection of larotrectinib as a treatment will be based solely on the presence of an NTRK gene fusion (the oncogenic driver) rather than the location of the tumour. In this way, larotrectinib represents a paradigm shift and step-change in the way cancer is treated, enabling cancer treatment to be delivered according to causation (i.e. the presence of NTRK gene fusion [NTRK+; NTRK fusion-positive cancer]) as opposed to tumour site e.g. lung, prostate, thyroid, as has been done traditionally. Larotrectinib is thus termed a 'tumour-agnostic' or

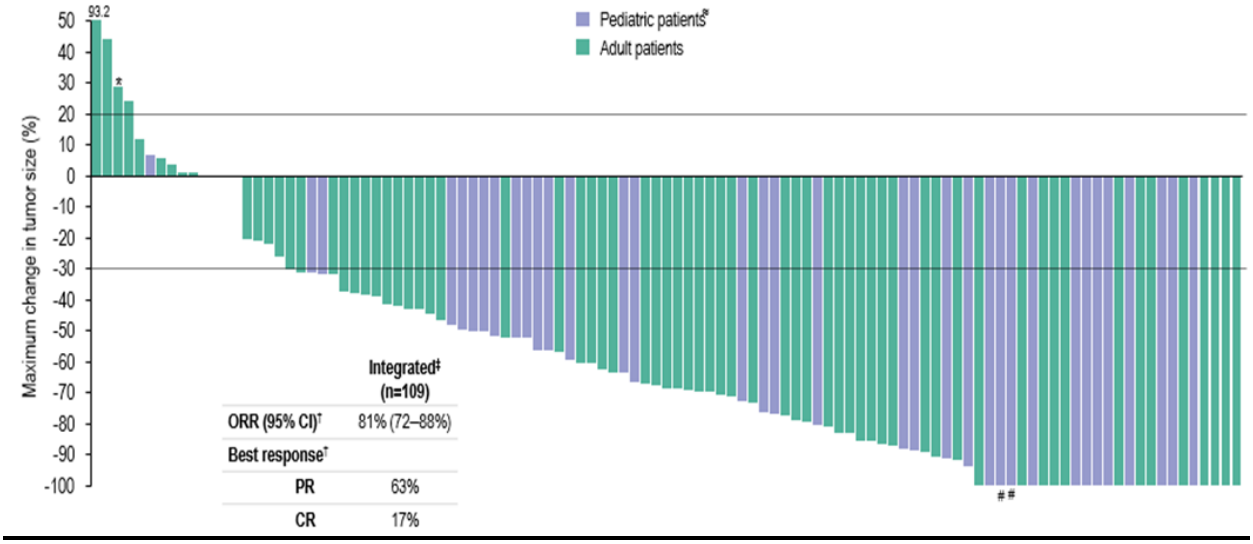
'histology-independent' therapy. This is highlighted in the figures below which demonstrates the histology independent and age independent efficacy of larotrectinib, giving a treatment option for both adults and children with NTRK fusion-positive solid tumours where previously no treatment was available.

Figure 1. Efficacy Results With Larotrectinib in the Integrated Analysis (Investigator Assessment)(14)



- surgical CR

Figure 2. Efficacy Results With Larotrectinib in the Integrated Analysis by Patient Age (Investigator Assessment)(14)



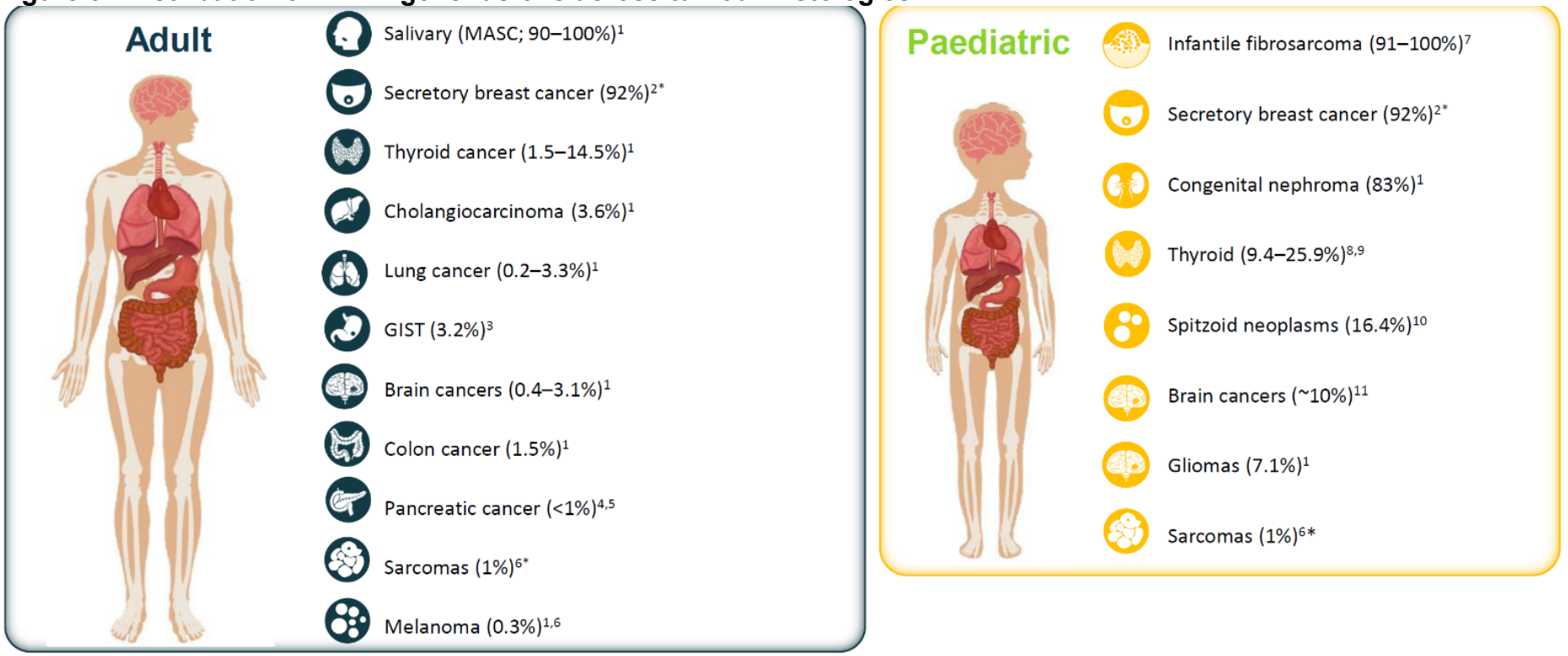
- surgical CR

Epidemiology

NTRK fusion-positive cancer is considered a rare disease, with less than 1% of solid tumours having NTRK gene fusions (10, 15, 16). In England, it has been estimated that the potential eligible patient population is <[redacted] patients. [redacted] is the number of patients estimated to be receiving last line of cancer therapy for various tumour sites harbouring NTRK gene fusion, however not all of these patients would harbour the NTRK gene fusion or be appropriate for a further line of therapy.

The frequency of NTRK gene fusions varies considerably according to tumour histology, occurring rarely (<0.1% to 3%) in common histologies, such as non-small cell lung cancer (NSCLC) and colorectal cancer (CRC), and more often (>90%) in several uncommon tumours, such as secretory breast carcinoma and infantile fibrosarcoma (IFS) (see Figure 3).

Figure 3. Distribution of NTRK gene fusions across tumour histologies



*Frequency in adult vs. paediatric patients not specified. GIST=gastrointestinal stromal tumour; MASC=mammary analogue secretory carcinoma; NTRK=neurotrophic tyrosine receptor kinase.1. Vaishnavi A, *et al. Cancer Discov.* 2015;5:25-34; 2. TognonC, *et al. Cancer Cell.*2002;2:367-376; 3. Brenca M, *et al. J Pathol.*2016;238:543-549; 4. Pishvaian MJ, *et al.Clin Cancer Res.* 2018; DOI: 10.1158/1078-0432.CCR-18-0531; 5. Cocco E, *et al. Nat Rev Clin Oncol.* 2018 15(12):731-747; 6. Stransky N, *et al. Nat Commun.* 2014 10;5:4846; 7. Bourgeois JM, *et al. Am J Surg Pathol.* 2000;24:937-946; 8. Ricarte-Filho JC, *et al. J Clin Invest.* 2013;123:4935-4944; 9. Prasad ML, *et al. Cancer.* 2016;122(7)1097-1107; 10. Wiesner T, *et al. Nat Commun.* 2014;5:3116; 11. Wu G, *et al. Nat Genet.* 2014;46(5):444-450.

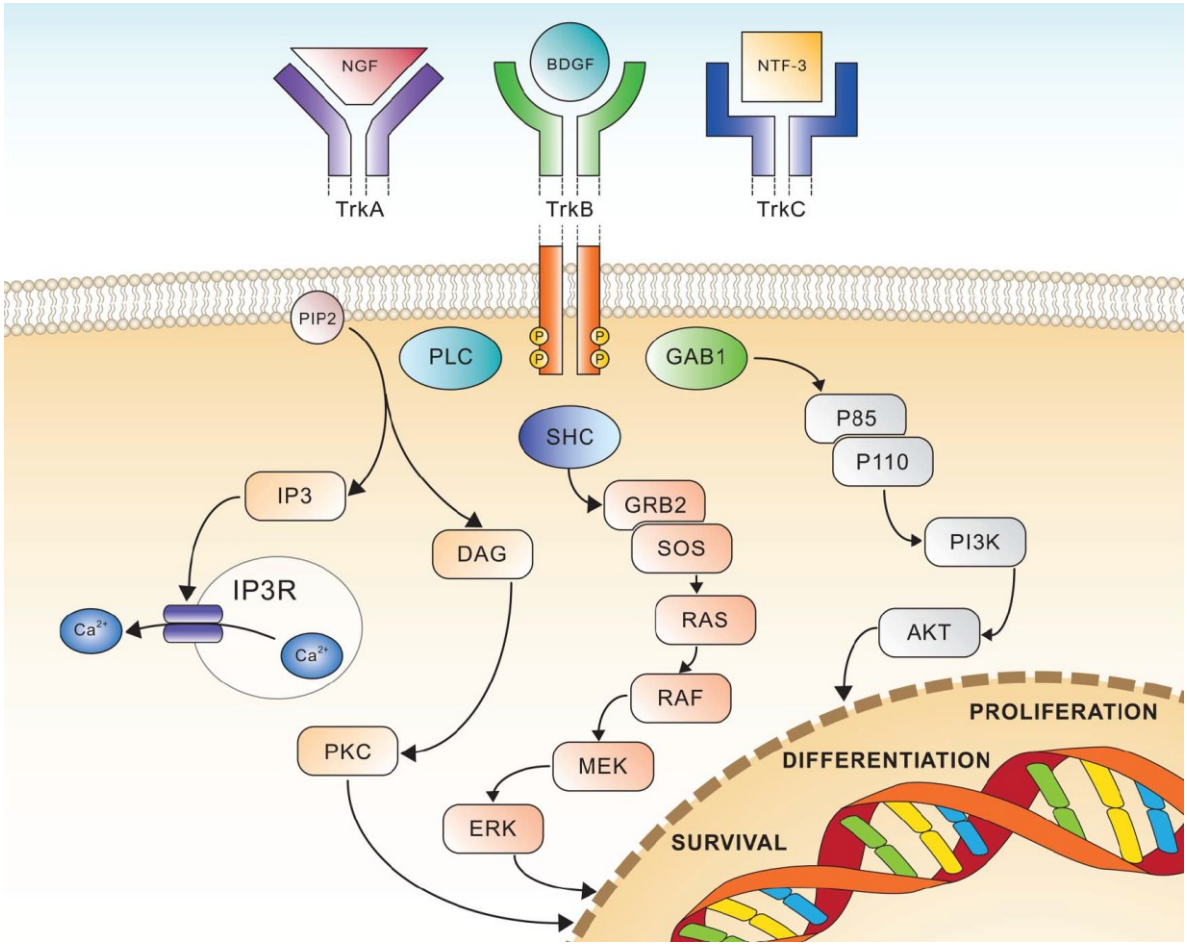
Of note, as a rare disease, data comes from single arm basket studies which enrolled patients who have the same molecular feature across anatomically and histologically diverse solid tumours. In contrast to a traditional, organ-site-specific trial, the central organizing principle of a basket study is the genomic alteration. A basket trial tests a particular therapy among patients with the same genomic alteration across multiple cancer types. Research into the most appropriate methods for these cases has indicated that basket trials as opposed to traditional tumour site specific trials, are considered suited to assess the efficacy of targeting genomic alterations that occur at low frequencies across a wide variety of tumour sitesl.(1)

Aetiology

- **NTRK signaling pathway**

Under normal physiologic conditions, the NTRK gene family (NTRK1, NTRK2, and NTRK3) encodes the tropomyosin receptor kinase (TRK) proteins TRKA, TRKB, and TRKC, regulating the proliferation, growth, and survival of neurons, through activation of neurotrophins (17-19) (see Figure 4). This process is responsible for normal development and function of the central and peripheral nervous system (e.g. pain, thermoregulation, proprioception, appetite, memory).

Figure 4. TRK signalling pathways (17)

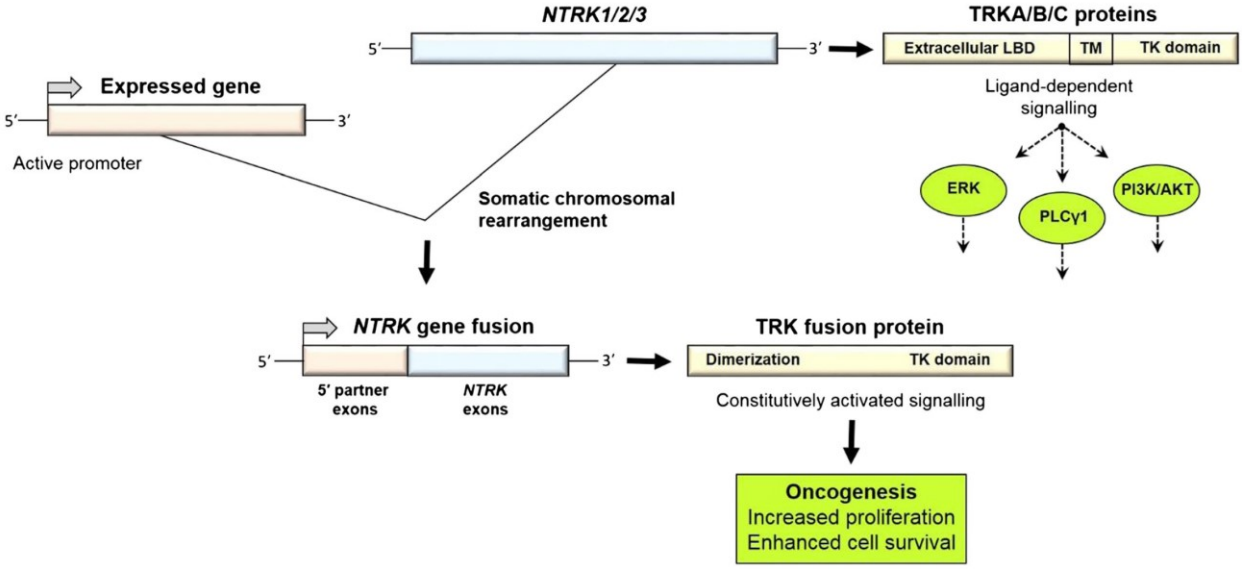


Adapted from Amatu A, Sartore-Bianchi A, Siena S (2016) (17).
AKT =v-AKT murine thymoma viral oncogene homologue; BDGF = brain-derived growth factor; DAG = diacyl glycerol; ERK = extracellular signal-regulated kinase; GAB1 = GRB2-associated binding protein 1; GRB2 = growth factor receptor-bound protein 2; IP3 = inositol trisphosphate; MEK = mitogen-activated protein kinase; NGF = nerve growth factor; NTF-3 = neurotrophin 3; PI3K = phosphatidylinositol-4,5-bisphosphate 3-kinase; PIP2 = phosphatidylinositol 4,5-bisphosphate; PKC = protein kinase C; PLC = phospholipase C; RAF = rapidly accelerated fibrosarcoma kinase; RAS = rat sarcoma kinase; SHC = Src homology 2 domain containing.

- **NTRK gene fusions**

NTRK fusion-positive tumours arise from a gene rearrangement involving fusion of a portion of the NTRK1, NTRK2 or NTRK3 gene with another unrelated gene (17). Gene fusions are a well-established class of primary oncogenic drivers. In all reported NTRK oncogenic gene fusions, the 3' region of the NTRK gene is broken apart and is fused together with a 5' region of an unrelated gene (fusion partner), causing the TRK fusion protein to become activated / expressed even in the absence of its ligand. This promotes cancer formation by driving unchecked cell proliferation and tumour growth, through the TRK pathway (9, 17, 20).

Figure 5. NTRK gene fusion (21)



LBD=ligand binding domain; TM=transmembrane; TK=tyrosine kinase;

Over 80 different NTRK fusion partners have been identified and shown to contribute to the development of NTRK fusion cancer across various histologic tumour types (6, 9, 17, 21, 22). The fusion partners vary based on histologic cancer type, with more common cancers typically having a higher number of different fusion partners, whereas rarer histologies commonly have one known fusion partner e.g. mammary analogue secretory carcinoma of the salivary gland / ETV6-NTRK3.

Personalised 'precision' medicine

Evidence has shown that use of targeted therapy paired with a specific oncogenic driver leads to better outcomes for patients than using a "one-size-fits-all" treatment approach with standard of care therapies (23-26). Use of targeted therapies has been shown to provide maximum benefit and have the potential to improve patient quality of life (QoL) (27). It is also expected to reduce the overall cost for the healthcare system, as patients ultimately avoid treatment unlikely to benefit them or potentially cause harm (24-28). Indeed, the NHS England report *Improving Outcomes Through Personalised Medicine* states 'Personalised medicine will help to maximise the value we can secure from the £15billion that the NHS currently spends on drugs each year.'

Importantly, if approved by EMA, larotrectinib will be the first histology independent drug approval in the EU

Management of NTRK fusion-positive cancer

- **Diagnosis**

Multiple testing methods are available to identify patients with tumours harbouring NTRK gene fusions. Next generation sequencing (NGS), allows for efficient testing with the ability to find NTRK gene fusions and other genomic targets simultaneously. Other detection methods include fluorescence in situ hybridisation (FISH), immunohistochemistry (IHC), reverse transcription polymerase chain reaction (RT-PCR), and whole-genome sequencing (WGS).

UK clinical practice

Immunohistochemistry can give an indication of the tumour's NTRK status but may require confirmation through subsequent FISH or genetic sequencing. However, the rollout of whole genome sequencing (WGS) for adult and paediatric cancers as part of the national NHS Genomic Medicine Service means that the requirement for undertaking multiple confirmatory diagnostic tests to determine the presence of NTRK gene fusions should become redundant. Routine access to WGS, particularly the rare mutations like NTRK gene fusion cancers, will enable more prompt diagnoses and precise clinical decision-making, with the potential for better outcomes for patients and also inform future cancer research in line with the ambitions set out in the UK's life sciences industrial strategy(29).

In the UK, the recently unveiled 'NHS Long Term Plan'(30) commits to dramatically improving cancer survival by 2028, partly by increasing the proportion of cancers diagnosed early, from a half to three quarters, facilitated by, among other aspects, accelerating access to diagnosis and treatment and maximising the number of cancers identified through screening.

Section 3.63 of the plan says *"We will extend the use of molecular diagnostics and, over the next ten years, the NHS will routinely offer genomic testing to all people with*

cancer for whom it would be of clinical benefit, and expand participation in research. The NHS will begin from 2020/21 to offer more extensive genomic testing to patients who are newly diagnosed with cancers so that by 2023 over 100,000 people a year can access these tests."

Furthermore, genomics is identified as transformative and an area of innovation in the NHS Long Term Plan, with an aim for the NHS to be the first national health care system to offer whole genome sequencing as part of routine care.

It is clear that the intention to implement WGS for all cancers is not due to any single advance in diagnosis or the possible introduction of any single product. Harnessing genomics into routine care has been recognised as fundamental to delivery of NHS cancer care, with the aim to improve diagnosis, survival and patient experience, as well as supporting future research across the whole disease pathway and maintaining the UK's leading position in the area of genomics.

The Genomic Medicines Service was launched in October 2018, building upon the success of the 100,000 Genomes Project and acting as a key step in the NHS move towards its vision of "Improving Outcomes Through Personalised Medicine"(31).

The 2018/19 final draft National Genomics Test Directory FAQ(32) states: *"The NHS Genomic Medicine Service aims to provide consistent and equitable access to cutting-edge genomic testing to England's 55 million population through consolidating existing services and improving access to the best of current NHS practice, while providing the foundation to deliver future technologies and approaches as they arise."*

The routine provision of fair and equitable access to relevant genetic testing for all cancer patients has been reiterated by Professor Sue Hill OBE, Chief Scientific Officer for England and the Senior Responsible Officer for Genomics in NHS England earlier this year:

"This transformation and commitment puts the NHS in the remarkable world-leading position of being the first country to have a national NHS Genomic Medicine Service (NHS GMS). Launched last October, and rolling out over the next 18 months, the NHS GMS will provide fair and equitable access to the full range of genomic testing to the

country's entire 55 million population. It is unique in providing a comprehensive Genomics offer from single gene to WGS – all embedded in routine care and working with clinicians across the NHS determine the significance of the result and the actionability and access to the right treatments."

Within The NHS Long Term Plan, the partnership between Genomics England and the NHS has an ambition that, during 2019, seriously ill children who are likely to have a rare genetic disorder, children with cancer, and adults suffering from certain rare conditions or specific cancers, will begin to be offered whole genome sequencing. Also, beginning from 2020/21, there is an aim to extend genomic testing to all people with cancer for whom it would be of clinical benefit, and patients who are newly diagnosed with cancer.

Currently, genomic testing for NTRK gene fusions is listed in the National genomic test directory for cancer(33) for Secretory Carcinoma (Salivary Gland), Infantile fibrosarcoma and Histiocytosis (where there is diagnostic uncertainty between benign and malignant process). All paediatric tumours are eligible to have WGS. NTRK gene fusions in Congenital mesoblastic nephroma (CMN) can also be specifically tested for.

The NICE processes guide stipulates that in instances where a diagnostic test to establish the presence or absence of a biomarker is carried out solely to support the treatment decision for the specific technology, the associated costs of the companion diagnostic test should be incorporated into the assessments of clinical and cost effectiveness. It is clear from the NHS Long Term Plan and the NHS "Improving Outcomes Through Personalised Medicine" that the intention to implement WGS for all cancers is not due to any single advance in diagnosis or the possible introduction of any single product, rather a wider ambition to improve health outcomes, expand the ability of centres in England to participate in research and establish the NHS as the first country to have a national NHS Genomic Medicine Service.

Furthermore, since other technologies for the treatment of people with NTRK fusions will be available any diagnostic testing will not be solely to support larotrectinib.

Given the aspirations of the UK to lead in the area of genomics, early access for patients to larotrectinib as a histology-independent cancer medicine, offers the Genomic Medicine Service an opportunity to move towards fulfilling its ambitions.

- **Treatment**

Currently, there are no approved treatment options in the UK specifically for patients with NTRK fusion-positive solid tumours and, to date, treatment recommendations regarding NTRK fusion-positive cancer have not been included within any UK guidelines.

Patients are currently treated per treatment guideline recommendations for the specific tumour site, irrespective of NTRK status.

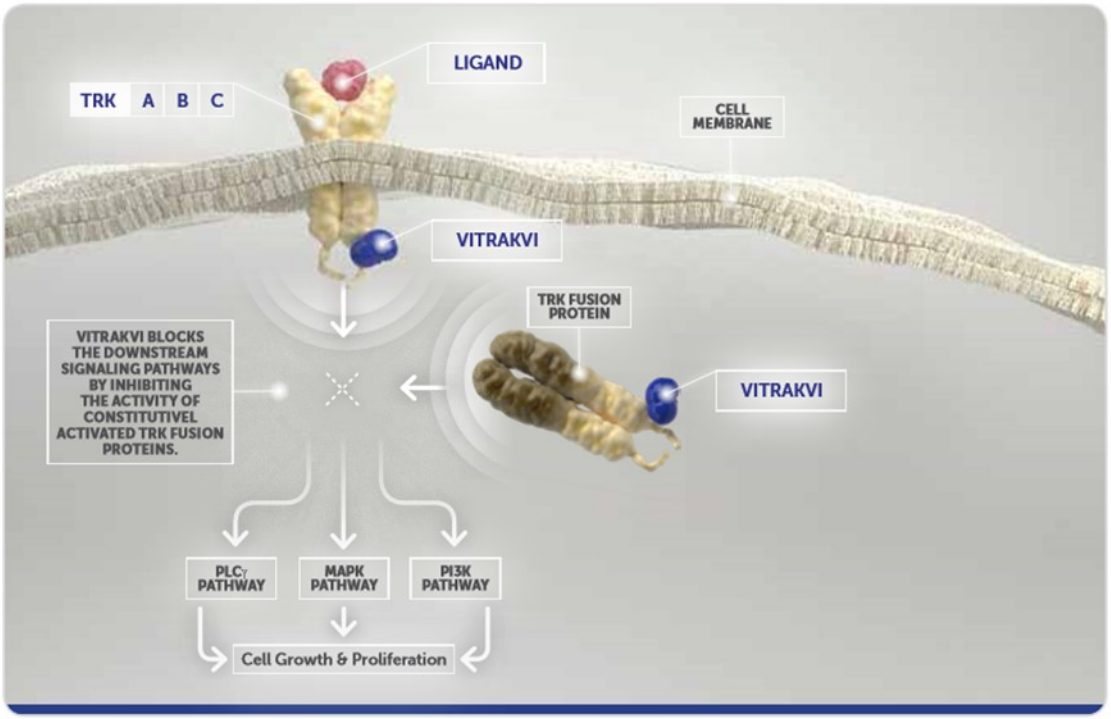
Treatment recommendations vary by tumour site. More common tumour sites such as NSCLC, CRC, melanoma and pancreatic have guideline recommendations for multiple lines of therapy (such as chemotherapy, targeted therapy, and/or immunotherapy); less frequent tumour sites/types such as appendix, salivary gland, and secretory breast carcinoma have limited or no treatment guidelines or recommendations due to scarcity of evidence supporting systemic therapy. These rarer tumours are mainly treated with chemotherapy and/or surgery, or patients are enrolled in clinical trials.

In line with the anticipated marketing authorisation, the patients eligible for larotrectinib will be those who have [redacted] This represents a small yet diverse group, ranging from infants to adults with multiple tumour sites / histologies but with a commonality of a high unmet medical need.

- **Larotrectinib**

Larotrectinib is an innovative technology that specifically targets the protein product of the NTRK fusion genes (i.e. TRK fusion proteins), irrespective of the location or histology of the tumour, turning off signalling pathways that usually allow NTRK fusion-positive cancers to grow.

Figure 6. Mechanism of action – larotrectinib (Vitrakvi®)



The histology-independent, targeted mechanism of action of larotrectinib has been clearly demonstrated in terms of the efficacy observed in diverse tumour types with NTRK gene fusions (4, 5, 14, 15). Patients with NTRK gene fusion cancers enrolled in clinical studies for larotrectinib had locally advanced / metastatic cancer that either could not be sufficiently controlled with available therapies or were likely to result in significant morbidity such as limb amputation to achieve control. Many patients who had received prior therapy had received multiple systemic therapies. Treatment of NTRK fusion-positive tumours with larotrectinib exhibited rapid, substantial antitumour activity with durable disease control that appears to be independent of NTRK isoform, tumour type and patient age (see Appendix E). There was no effect in patients without an NTRK fusion, irrespective of tumour type. This is not surprising given the mechanism of action of larotrectinib as a potent and selective inhibitor of TRKA, TRKB, and TRKC.

Larotrectinib is effective across a broad range of tumours including rare tumours and rare subsets of more common tumours, and in paediatric and adult patients ranging in age from [redacted] years. The safety profile is characterised by recognisable toxicities, which are predictable and can be monitored. These data demonstrate the ability to treat a patient based on the type of mutation (gene fusion) their tumour

contains, regardless of where the cancer originated. This is considered a therapeutic advance when compared with traditional chemotherapy which can be of limited benefit in many tumours, non-specific with respect to molecularly defined targets, generally associated with significant toxicities, and often unsuitable for certain patient populations (e.g. very young paediatric patients or adult patients who may be elderly or frail).

After FDA approval of larotrectinib, larotrectinib has been rapidly incorporated into National Comprehensive Cancer Network (NCCN) guidelines for many solid tumours in the US (see Table 3).

Table 3. NCCN guidelines incorporating NTRK screening and larotrectinib therapy guidance

Guideline title	NTRK fusion testing recommendation	Therapy Guidance [a]	Date / ref
NCCN Guidelines for Colon Cancer and Rectal Cancer	Testing should include the neurotrophic receptor tyrosine kinase (NTRK) gene fusion.	Larotrectinib added as a treatment option for patients with metastatic colorectal cancer that is NTRK gene fusion positive.	Version 1.2019
NCCN Guidelines for Cutaneous Melanoma	-	Systemic Therapy for Metastatic or Unresectable Disease Second-line or subsequent therapy: Useful in certain circumstances: 'Larotrectinib for NTRK gene fusion positive tumors'	Version 2.2019
Salivary Gland (NCCN Guidelines for Head and Neck Cancers)	-	NTRK therapy (e.g. larotrectinib) has been included as an option for recurrent NTRK gene fusion-positive salivary gland tumours with distant metastases, PS 0-3 (on page SALI-4).	V.1.2019
NCCN Guidelines for Non-Small Cell Lung Cancer	Testing for advanced or metastatic disease should include the NTRK gene fusion; fusion; if positive, see NSCL-26	NSCL-26 (new page added) Larotrectinib was added as a treatment option for first-line or subsequent therapy of NTRK gene fusion positive metastatic NSCLC	V.3.2019
NCCN Guidelines for Occult Primary	Per physician discretion, TRK protein testing can be considered as part of a broad IHC testing (a positive test should then be confirmed with NGS)	No recommendation	V.2.2019

NCCN Guidelines for Pancreatic Adenocarcinoma	-	'Useful in certain circumstances for second-line therapy for locally advanced/metastatic disease and therapy for recurrent disease, if good performance status: Larotrectinib (if NTRK gene fusion positive)'	Version 2.2019
NCCN Guidelines for Soft Tissue Sarcoma	No recommendation	Guidelines updated to include larotrectinib (for NTRK gene-fusion sarcomas) as a single agent for systemic therapy for soft tissue sarcoma subtypes with non-specific histologies.	V.1.2019
NCCN Guidelines for Thyroid Carcinoma	For advanced, progressive, or threatening disease, genomic testing to identify actionable mutations is recommended [b]	Guidelines updated to include larotrectinib as an option for NTRK gene fusion positive structurally persistent /recurrent locoregional or distant metastatic disease – for anaplastic (preferred regimen), follicular, Hürthle cell and papillary thyroid carcinoma	Version 3.2018

IHC=immunohistochemistry; NGS=next generation sequencing; NTRK=neurotrophic tyrosine kinase receptor; TRK=tropomyosin receptor kinase.
[a] NCCN guidelines that have added larotrectinib treatment state recommendations have been made using 'category 2a' evidence (i.e. Category 2A: Based upon lower-level evidence, there is uniform NCCN consensus that the intervention is appropriate.)
[b] This recommendation pertaining to testing and treatment is for papillary, follicular, Hürthle cell, and anaplastic carcinoma.

ESMO has also recently launched the ESMO Scale for Clinical Actionability of molecular Targets (ESCAT) (34). This scale provides a framework to assign DNA alterations into tiers that reflect their clinical utility for selecting patients for treatment with targeted therapies. This was developed by leading cancer specialists in Europe and North America with the aim of 'optimising patient care by making it easier to identify patients with cancer who are likely to respond to precision medicines and help make treatment more cost effective' (ESMO Press release 21st August 2018). Based on the strength of clinical evidence supporting them (Tier I-V), the new grading system classes alterations in tumour DNA according to their relevance as markers for selecting patients for targeted treatment. Using the ESCAT scale, larotrectinib is designated 'tier I-C', designated where clinical trials in multiple tumour types, or basket

clinical trials, have demonstrated a clinically meaningful benefit for the target–drug pair with similar magnitude of benefit across the different tumour types. In this scenario, the clinical value of a target–drug match can be accepted across cancers that harbour the target abnormality (34).

The availability of larotrectinib, a highly selective oral precision medicine, represents a significant therapeutic innovation in the field of precision medicine and targeted therapies. Its introduction is in alignment with the aspirations of NHS Long Term Plan in respect of personalised medicines / genomic-directed therapy.

B.1.4 Equality considerations

As highlighted in section B1.3, larotrectinib reflects a new paradigm where cancer treatment targets the oncogenic driver rather than the tumour location. The histology-independent nature of larotrectinib and the rarity of the NTRK gene fusion cancers targeted by larotrectinib present unorthodox challenges to the traditional technology assessment process, for example:

- **Trial design.** As a rare disease, data come from single arm basket studies that enrolled patients who have the same molecular feature across anatomically and histologically diverse solid tumours. In contrast to a traditional, organ-site-specific trial, the central organizing principle of a basket study is the genomic alteration. A basket trial tests a particular therapy among patients with the same genomic alteration across multiple cancer types. Research into the most appropriate methods for these cases has indicated that basket trials as opposed to traditional tumour site-specific trials are considered suited to assess the efficacy of targeting genomic alterations that occur at low frequencies across a wide variety of tumour sites (1). Indeed, the MHRA agreed that a single arm study was an appropriate design to support an MAA, given the extreme rarity of NTRK fusion cancers.

- **High numbers of comparators.** There are no existing comparator treatments for patients with NTRK fusion cancer. As an RCT is not appropriate or feasible in this rare disease with multiple tumour types and complex treatment pathways, the standard of care selected for this appraisal, reflects a mixed basket of last-line standard of care approaches [redacted] across histologically diverse tumours.

In addition to the uncertainties commonly associated with oncology appraisals such as immature overall survival data, the aforementioned complexities add to the uncertainty for a histology independent treatment. However, these uncertainties are inherent to the rarity of the gene fusion and the innovative nature of this product and therefore need to be taken into consideration so that patients with rare gene fusions are not inequitably disadvantaged. A recent publication by Love-Koh et al 2018 (2), exploring the appropriateness of HTA methods for evaluating precision medicines

suggested that HTA bodies will need to adapt their methods and processes to facilitate evaluation of these new technologies. Indeed, HTA assessment, like clinical practice, has traditionally considered cost-effectiveness of cancer therapies based on tumour location. There is no precedence or guidance for assessing the cost effectiveness of histology independent treatments where evidence is obtained through basket studies and there are multiple 'standard of care' comparators. In particular, there is lack of guidance on methods for controlling for single arm basket-studies, aggregation of comparator data to facilitate a comparative assessment, or how confounding factors (treatment effect modifiers) may be controlled for.

In this appraisal the Committee is asked to give balanced consideration to downward as well as upward uncertainty that is associated with evaluating this histology independent innovation. Further, that a recommendation to enter the CDF will go towards addressing much of the uncertainty without denying patients an effective treatment in a timely manner.

Given the current level of uncertainty, Bayer proposes that whilst data mature, larotrectinib is made available in a timely manner through the Cancer Drugs Fund.

B.2 Clinical effectiveness

B.2.1 Identification and selection of relevant studies

Three clinical studies were identified for the indication being appraised: a phase 1 adult clinical trial (LOXO-TRK-14001, NCT02122913), a phase 1/2 paediatric clinical trial (LOXO-TRK-15003, NCT02637687, SCOUT), and a phase 2 adolescent and adult clinical trial (LOXO-TRK-15002, NCT02576431, NAVIGATE). SCOUT and NAVIGATE trials are still actively enrolling patients. See appendix D for full details of the process and methods used to identify and select the clinical evidence relevant to larotrectinib in the treatment of NTRK fusion-positive advanced solid tumours.

B.2.2 List of relevant clinical effectiveness evidence

See Table 4.

Table 4. Larotrectinib – published clinical effectiveness evidence

Study	Pooled Analyses of NCT02576431, NCT02122913, and NCT02637687	NAVIGATE (Phase II) NCT02576431 LOXO-TRK-15002	(Phase I) NCT02122913 LOXO-TRK-14001	SCOUT (Phase I / II) NCT02637687 LOXO-TRK-15003
Latest publications	**Lassen (2018) (14)** 55 + 67 (109 evaluable) patient cohort (as at 30th July 2018) (conference presentation) 55 + 35 patient cohort (as at 19th Feb 2018): - abstract from ESMO submitted prior to conference	See pooled analyses	**Hong (2019) (4)** data cut-off Feb 2018	**Laetsch (2018) (5)** Results of phase I dose escalation cohort
Other publications	**Drilon (2018) (15)** 55 patient cohort (as at July 2017) Also publications by tumour type: **Brose 2018 (35)** [thyroid]; **Farago (2018) (36)** [NSCLC]; **Nathenson 2018 (37)** [GI tumours]; **Wirth 2018 (38)** [Thyroid and salivary gland tumours]	See pooled analyses		**DuBois (2018) (39)** [Sarcomas]
Study design	See individual studies	Phase II, open-label, basket study	Phase I, open-label, multicentre, 3+3 dose-escalation with expansion phase in patients with NTRK gene fusions only	Phase 1, open-label, dose escalation study Phase 2, single arm open-label study in IFS, other extracranial solid tumours, and primary CNS tumours
Population	Additional patients with TRK fusion cancer enrolled after reporting of the primary analysis set Drilon 2018: N=55 with NTRK fusion-positive cancer 17 unique fusion-types	Adults and adolescents (≥12yrs) with advanced or metastatic NTRK fusion-positive solid tumour.	Adults (≥18yrs). Locally advanced or metastatic, solid tumours refractory to standard therapies Expansion phase: patients with NTRK fusion-positive tumours only. NTRK fusion-positive: n=8; No documented NTRK fusion: n=62 23 unique cancer diagnoses	Paediatric (1mo. - <21yrs) with locally advanced or metastatic solid tumour or primary CNS tumours
Intervention / Comparator	Larotrectinib (100mg or 150mg orally b.d.)	Larotrectinib 100mg b.d. in continuous 28-day cycles	Larotrectinib administered on continuous 28-day schedule	Larotrectinib in continuous 28-day cycles Cohorts 1 & 2: Calculated on basis of age / body weight to

			Dose escalation: 50mg o.d. – 200mg b.d. Expansion phase: 100mg b.d.	provide doses equivalent to adult dose of 100mg / 150mg b.d. Cohort 3 and phase 2: oral larotectinib dose 100mg/m^2 b.d. (not to exceed 100mg b.d.)
Table continued… **Study**	**Pooled Analyses of NCT02576431, NCT02122913, and NCT02637687**	**NAVIGATE (Phase II) NCT02576431 LOXO-TRK-15002**	**(Phase I) NCT02122913 LOXO-TRK-14001**	**SCOUT (Phase I / II) NCT02637687 LOXO-TRK-15003**
Trial supports application for marketing authorisation?	YES. FDA: Primary analysis set (PAS) n=55 (enrolled patients with sufficient duration of follow-up) EU: extended PAS (ePAS) n=73 (enrolled patients with sufficient duration of follow-up)	YES, the efficacy and safety evidence for larotrectinib were based on results from these studies, primarily as pooled analyses.		
Trial used in economic model?	Yes. As there are now more evaluable patients, data (as at July 2018) for 102 patients is used	Data from a pooled analysis of NTRK fusion-positive patients from these trials have been used in the economic model.		
Rationale for use / non-use in model	Largest / most comprehensive dataset providing clinical and safety evidence for larotrectinib in NTRK fusion-positive solid tumours.			
Reported outcomes specified in the decision problem	Primary: **ORR by IRC assessment.** Secondary: **ORR (investigator assessment)**, **DoR**, **PFS (6m/12m rate), OS (12m rate)** and **safety**.	Primary: ORR (CR+PR) according to RECIST v1.1 or RANO criteria Secondary: Best overall response, DoR, PFS, OS, **exploratory quality of life,** safety	Primary: safety, including dose-limiting toxicity. Secondary endpoints: ORR (CR+PR) and DoR.	Primary: (phase 1): safety, including dose-limiting toxicity. (phase II): ORR. Secondary (phase I): DoR, best OR, **health-related quality of life (HR-QoL),** safety. (phase II): DoR, safety
All other reported outcomes	Secondary: time to response / best response, time on treatment, disease control rate (DCR)	Secondary: CBR (proportion of patients with confirmed CR, PR or SD lasting ≥16 weeks),	Pharmacokinetics.	

AE=adverse event; b.d.=twice daily; CBR=clinical benefit rate; CNS=central nervous system; CR=complete response; DCR=disease control rate; DoR =duration of response; ePAS=extended Primary analysis set; FDA=Food and Drug Administration; GI=gastrointestinal; GIST=gastrointestinal stromal tumour; HR-QoL=health-related quality of life; IFS=infantile fibrosarcoma; IRC=independent review committee; MASC= mammary analogue secretory carcinoma; mo.=months; o.d.=once daily; ORR=overall response rate: OR=objective response; PAS=primary analysis set; PR=partial response; PFS=progression-free survival; RANO=Response Assessment in Neuro-oncology; RECIST=Response Evaluation Criteria in Solid Tumours; SD=stable disease; TRK=tyrosine receptor kinase; yrs=years;

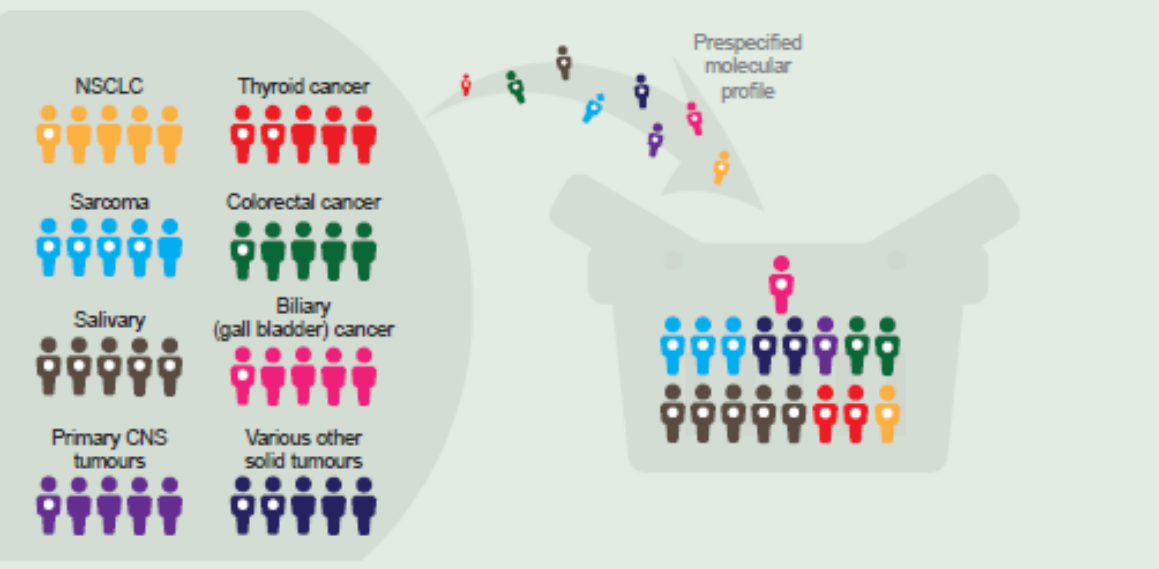
As discussed earlier in this submission, an RCT would not be appropriate or feasible for evaluating larotrectinib as a histology independent treatment.

While larotrectinib was in phase 1, the sponsor collaborated with global health authorities to devise a feasible drug development approach for larotrectinib that recognised the rarity of NTRK fusion cancer and that NTRK gene fusions were found across many different tumour types. As there are no available targeted therapies for patients with solid tumours with NTRK gene fusions and, thus, no common comparator that could be used in a comparative trial across all tumour types affected, a single-arm study approach and basket trial design (see Figure 7) was considered most appropriate. Indeed, the MHRA agreed that a single arm study was an appropriate design to support an MAA, given the extreme rarity of NTRK fusion cancers.

The use of a single-arm basket trial study design, provides clinical evidence for an indication based on an oncogenic driver irrespective of the primary disease histology, allowing extrapolation of the observed treatment effect to diverse tumour histologies (40).

The characteristics of the three clinical trials identified for larotrectinib are described within the clinical section, alongside the pooled analysis methodology. The pooled analysis results are used as the source of clinical effectiveness data for larotrectinib in this submission, including within the economic model. Early in the development programme - based on the rare nature of NTRK gene fusions, the heterogeneity of the cancer types, and advice from global regulators (15) - the decision was made to pool efficacy data across all 3 studies from patients with a solid tumour harbouring an NTRK gene fusion. This was possible due to the consistency of treatment response, safety, and tolerability across tumours and age groups for larotrectinib, and the common eligibility criteria and study procedures. The pooled analysis approach provides a more robust estimate of the responses in patients with NTRK fusion cancer and was agreed with regulatory agencies. The pooled analysis was used for both the US Food and Drug Administration (FDA) and the EMA regulatory submissions.

Figure 7. What is a basket trial?



Enrolment in a basket trial is based on a molecular profile, not tumour type; as such, basket trials are tumour-agnostic. Regardless of the location or histology of the tumour, if the patient has the pre-specified molecular profile (for example, an NTRK gene fusion), the patient is eligible for the trial.

B.2.3 Summary of methodology of the relevant clinical effectiveness evidence

This submission focuses on results from the pooled analysis of the three larotrectinib studies briefly introduced in Table 4 and described below. The consistency of treatment response, safety, and tolerability across tumours and age groups across common eligibility criteria and study procedures permitted the pooling of interim data in support of global regulatory submissions. A comparative summary of the methodology of the Adult Phase 1 (LOXO-TRK-14001), SCOUT (LOXO-TRK-15003) and NAVIGATE (LOXO-TRK-15002) trials and the pooled analysis is presented in Table 5.

Pooled analysis (2)

The primary endpoint for the pooled efficacy analysis was overall response rate (ORR) by Independent Review Committee (IRC) assessments, based on RECIST (version 1.1) [Response Evaluation Criteria in Solid Tumours] for non-CNS solid tumours.

Secondary objectives included duration of response (DoR), progression-free survival (PFS) and safety (15). Details of the rationale, study design and methodology were published along with early results from the pooled analysis (data cut-off July 2017) (15). The larotrectinib regulatory submission to the FDA, was based on this analysis involving 55 patients. Subsequently, an updated analysis has been presented at ESMO (Lassen 2018 (14)– ESMO abstract, data cut-off February 2018 and ESMO presentation slides, data cut-off July 2018) and used within the EMA regulatory submission. Unpublished aspects of the pooled analysis are drawn from Statistical analysis plans (SAPs) (41-43) and the manufacturer licence application submission to the European Medicines Agency (EMA) (44-46).

Unless otherwise specified, the results and analyses of all efficacy and safety outcomes in this submission are presented for events occurring up to the most recent pooled analysis of 30 July 2018. The pooled analysis of efficacy is presented as two datasets – ePAS2 and SAS3 [see 'Analysis sets', section B2.4 for further information]. ePAS2 or 'Extended Primary Analysis set 2' reflects an updated version of the original primary analysis set of non-CNS NTRK fusion solid tumours (PAS) in the FDA submission, including data from additional patients recruited since that data-cut of July 2017. SAS3 or 'Supplementary analysis set 3' consists of 9 paediatric and adult patients with primary CNS tumours but who otherwise met PAS eligibility criteria and were enrolled before the data cut-off. The pre-specified integrated primary analysis excluded patients with primary CNS tumours before enrolment of any CNS patient. Surgery and radiation treatments can lead to varying amount of oedema / inflammation and scarring, which can impact the radiological assessment in patients with primary CNS tumours. SAS3 utilised disease assessments performed by the Investigator as opposed to central assessment.

LOXO-TRK-14001: A Phase 1 Study of the Oral TRK Inhibitor LOXO-101 in Adult Patients with Solid Tumours (NCT02122913) (4, 15)

LOXO-TRK-14001 is a multicentre, phase I, open-label, dose escalation (5 planned dose cohorts with 3 to 6 patients per cohort) and dose expansion study (2 planned cohorts) in adult patients with advanced solid tumours. The primary objective of the dose-escalation portion of the study was to characterise safety, in terms of dose-

limiting toxicities (DLT) and maximum tolerated dose (MTD), with the aim of identifying an appropriate dose of larotrectinib for further study. The dose-escalation phase of the study is now closed and results fully published with information on efficacy to a data cut-off of February 19th 2018 (Hong 2019 (4)). Following dose escalation, a dose expansion phase was initiated only for patients with documented NTRK gene fusion cancer. Based on its tolerability and on the durability of response in patients with NTRK fusion cancer, 100 mg b.d. was set as the recommended phase II dose. Eight patients from this study contribute to the pooled analysis.

LOXO-TRK-15003: A Phase 1/2 Study of the Oral TRK Inhibitor LOXO-101 in Paediatric Patients with Advanced Solid or Primary Central Nervous System Tumours (NCT02637687; SCOUT) (5, 15)

SCOUT is an ongoing, international, multicentre, open-label phase I / II study in paediatric patients aged 1 month to 21 years with advanced solid or primary CNS tumours. The primary objective of the phase I portion of the study was to assess the safety of larotrectinib in paediatric patients, with the aim of identifying an appropriate paediatric dose for further study. SCOUT is thus divided into a dose escalation phase, dose expansion (both phase I), and a Phase II portion where enrolment is restricted to patients with documented NTRK gene fusion cancer (3 cohorts: IFS, other extracranial solid tumours, and primary CNS tumours). Results of the phase I dose escalation cohort (now complete) have been fully published (Laetsch 2018 (3)). The recommended phase 2 dose was defined as 100mg/m² twice daily (maximum 100 mg per dose) for infants, children, and adolescents, regardless of age. The dose expansion phase and phase 2 portion are ongoing. Thirty-two patients from this study contribute to the pooled analysis.

LOXO-TRK-15002: A Phase II Basket Study of the Oral TRK Inhibitor LOXO-101 in Subjects with NTRK Fusion-Positive Tumours (NCT02576431; NAVIGATE) (15)

NAVIGATE is an ongoing, international, multicentre, phase II, open-label “basket” study in patients 12 years of age or older. The study has 8 cohorts of patients with recurrent, advanced solid tumours with a documented NTRK gene fusion, including non-small cell lung carcinoma (NSCLC), thyroid cancer, sarcoma, colorectal cancer,

salivary gland cancer, biliary cancer, primary CNS tumours, and a cohort that enrolled patients of all other histologic types or patients without measurable disease. NAVIGATE has yet to be published, however evaluable patients are included in the pooled analyses, published at various analysis time-points: Drilon 2018 (15) – data cut-off July 2017; Lassen 2018 (14) – ESMO abstract, data cut-off February 19th 2018 and ESMO presentation slides, data cut-off July 2018 (see Table 4). Sixty-two patients from this study contribute to the pooled analysis.

Available efficacy results by individual study (to 19 February 2018 data cut-off) are presented for completeness in Appendix O. These data also provide 'proof of concept' of larotrectinib, in that they demonstrate lack of activity in solid tumours not harbouring an NTRK gene-fusion.

Table 5. Comparative summary of trial methodology for larotrectinib studies (4, 5, 15, 44, 45)

Trial number (acronym)	Pooled Analysis of patients from LOXO-TRK-14001, SCOUT and NAVIGATE trials	NAVIGATE LOXO-TRK-15002 (NCT02576431)	Adult Phase I LOXO-TRK-14001 (NCT02122913)	SCOUT LOXO-TRK-15003 (NCT02637687)
Trial design	Integrated safety and efficacy analysis of adult and paediatric patients with prospectively identified NTRK fusion cancers enrolled and treated in 1 of 3 clinical studies for larotrectinib (NCT02122913, SCOUT and NAVIGATE)	Phase II, international, multicentre, open-label basket study in patients with recurrent, advanced solid tumours with a documented NTRK gene fusion	Multicentre, open-label, phase I, dose-escalation and dose expansion study in adult patients with a locally advanced or metastatic, solid tumour refractory to standard therapies	International, multicentre, open-label, phase I/II study in paediatric patients (aged 1 month to 21 years) with advanced solid or primary CNS tumours
Location	Patients included in the analysis to date are from: Asia (South Korea), Australia, Europe (Denmark, Germany, France, Ireland, Italy, Spain) and US Across 38 study sites	Asia (Singapore, South Korea), Europe (Denmark, France, Ireland, Portugal, Spain, UK) 35 study sites	US 8 study sites	Australia, North America (Canada, US), Europe (Denmark, France, Germany, Ireland, Italy, Spain, Sweden, Switzerland, UK) 26 study sites
Duration of study	March 2015 –July 30th, 2018 (latest analysis data cut-off)	October 15, 2015 – ongoing	May 1, 2014 – ongoing	December 22, 2015 – ongoing
Method of randomisation	Not applicable - single arm studies			
Method of blinding	Single-arm studies - open label for study patients and investigators			
Eligibility criteria	Inclusion: -**ePAS2 only:** Locally advanced or metastatic non-CNS primary solid tumour or **-SAS3 only:** primary CNS tumour with a documented NTRK gene fusion assessable according to RECIST, version 1.1 (non-CNS)	Inclusion: -Age ≥12 years -Locally advanced or metastatic solid tumour with documented NTRK gene fusion that could be assessed according to RECIST, version 1.1	Inclusion: -Age ≥18 years -Locally advanced or metastatic solid tumour (with documented NTRK gene fusion for expansion phase of study)	Inclusion: -Age 1 month–21 years; -Locally advanced or metastatic solid tumour or primary CNS tumour or patients with locally advanced IFS who required disfiguring surgery or limb amputation to achieve surgical CR

Trial number (acronym)	Pooled Analysis of patients from LOXO-TRK-14001, SCOUT and NAVIGATE trials	NAVIGATE LOXO-TRK-15002 (NCT02576431)	Adult Phase I LOXO-TRK-14001 (NCT02122913)	SCOUT LOXO-TRK-15003 (NCT02637687)
	or RANO (primary CNS tumours) criteria -Previously treated with standard therapy (if available or possible) -ECOG PS 0-3 -adequate major organ function - received 1 or more doses of larotrectinib Exclusion: -Current treatment with a strong CYP3A4 inhibitor or inducer -receipt of an investigational or anticancer therapy within 2 weeks, or major surgery within 4 weeks, prior to enrolment -previous treatment with kinase inhibitors (NB one patient enrolled before amendment) -clinically significant cardiovascular disease or history of prolonged QT interval corrected for heart rate (QTc) - Symptomatic or unstable brain metastases -any conditions affecting oral absorption	-Previously treated with standard therapy (if available or possible) -ECOG PS 0–3 -adequate organ function -life expectancy of ≥3 months -Patients with primary CNS tumours or metastasis who were neurologically stable Exclusion: -Current treatment with a strong CYP3A4 inhibitor or inducer -receipt of an investigational or anticancer therapy within 2 weeks, or major surgery within 4 weeks, prior to enrolment -previous treatment with kinase inhibitors (NB one patient enrolled before amendment) -clinically significant cardiovascular disease or history of prolonged QT interval corrected for heart rate (QTc) - Symptomatic or unstable brain metastases -any conditions affecting oral absorption	-Previously treated with standard therapy (if available or possible) -ECOG PS 0–2 -adequate organ function -life expectancy of ≥3 months -Patients with primary CNS tumours or metastasis who were neurologically stable, and did not require steroid management of CNS symptoms within the 2 weeks prior to study entry, could enroll Exclusion: -Current treatment with a strong CYP3A4 inhibitor or inducer -receipt of an investigational or anticancer therapy within 2 weeks, or major surgery within 4 weeks, prior to enrolment - clinically significant cardiovascular disease or history of prolonged QT interval corrected for heart rate (QTc) -any conditions affecting oral absorption	[Patients must have measurable disease (per RECIST v1.1, RANO criteria, or International Neuroblastoma Response Criteria)] (with documented NTRK gene fusion for expansion phase / phase II) -Previously treated with standard therapy (if available or possible) - Karnofsky (≥16 years) or Lansky (<16 years) PS of ≥50 Exclusion: -Clinically significant cardiovascular disease or corrected QT interval >480 ms -an active uncontrolled systemic infection -any conditions affecting oral absorption Current treatment with a strong CYP3A4 inhibitor or inducer -receipt of an investigational or anticancer therapy within 2 weeks, or major surgery within 4 weeks, prior to enrolment
Trial drugs and method of administration	According to the dose expansion or phase II portion of	Oral larotrectinib 100 mg b.d. in 28-day cycles.	**Dose escalation:** Oral larotrectinib, once- or twice-daily, on a continuous 28-	**Dose escalation:** Oral larotrectinib (capsule or liquid formulation)

Trial number (acronym)	Pooled Analysis of patients from LOXO-TRK-14001, SCOUT and NAVIGATE trials	NAVIGATE LOXO-TRK-15002 (NCT02576431)	Adult Phase I LOXO-TRK-14001 (NCT02122913)	SCOUT LOXO-TRK-15003 (NCT02637687)
	the study protocol the patient was treated under. **Adults** [redacted]/68 adults received 100mg b.d dosing of larotrectinib [redacted]68 received 150mg b.d. **Paediatrics** At least [redacted]/34 paediatric patients received 100mg/m^2 b.d. dosing of larotrectinib	Larotrectinib was administered as capsules unless patients could not swallow capsules, in which case a liquid formulation was available.	day schedule, in increasing dose levels according to a standard 3+3 dose escalation scheme. Dose levels: 50 mg q.d. / 100 mg q.d. / 200 mg q.d. / 100 mg b.d. / 150 mg b.d. / 200 mg b.d. Dose escalation proceeded through planned dose levels, according to dose-limiting toxicity (DLT) in cycle 1, until the maximum tolerated dose (MTD) was reached. **Expansion Phase** Oral larotrectinib 100 mg b.d.	Cohort 1: Doses ranging from 17%–96% of the BSA-adjusted recommended adult phase 2 dose of 100 mg b.d. Cohort 2: Doses ranging from 30%–208% of the BSA-adjusted adult dose of 150 mg b.d. Cohort 3: 100 mg/m² b.d. (maximum of 100 mg per dose) Dosing was continuous for 28-day cycles. **Phase II:** Oral (capsule or liquid formulation) larotrectinib 100 mg/m2 b.d., not to exceed 100 mg b.d.
	Larotrectinib was administered until disease progression, the occurrence of unacceptable toxicity, or the withdrawal of patient consent. Dose interruptions of up to 4 weeks to allow for recovery were specified for clinically significant adverse events. Upon recovery, patients could either continue at the assigned dose of larotrectinib or have the dose reduced. Patients who had drug-related toxicity requiring a recovery period longer than 4 weeks were withdrawn from study drug administration, unless there was compelling evidence of response and no alternative treatment.			
Permitted and disallowed concomitant medication	As per individual protocols, which were broadly similar	**Permitted** -Palliative radiotherapy to specific sites of disease -Standard supportive medications (e.g., haematopoietic growth factors, transfusions, anti-emetics, anti-diarrhoeals, and glucocorticoids in short courses); -patients could continue standard of care medications that they had been receiving for the previous 28 days at stable doses e.g. gonadotropin-releasing hormone or luteinising hormone-releasing hormone agonists for patients with prostate cancer		

Trial number (acronym)	Pooled Analysis of patients from LOXO-TRK-14001, SCOUT and NAVIGATE trials	NAVIGATE LOXO-TRK-15002 (NCT02576431)	Adult Phase I LOXO-TRK-14001 (NCT02122913)	SCOUT LOXO-TRK-15003 (NCT02637687)
		-glucocorticoids could be administered to primary CNS tumour patients to reduce peritumoural oedema and improve neurological deficits **Disallowed** Other anti-tumour approved or investigational agents that were being used with the intent to effect tumour shrinkage (e.g. chemotherapy); known strong inhibitors or inducers of CYP3A4; any other investigational agents		
Outcomes	see Table 4, Table 6, and Table 7 for outcomes, scoring methods and timings of outcome assessments			
Pre-specified subgroup analyses	ORR was investigated by sex, age group, race, tumour type, baseline disease characteristics, and NTRK gene fusion type.	No pre-specified subgroup analyses		

Key: b.d.=twice-daily; CBR=clinical benefit rate; CNS=central nervous system; CR=complete response; CTCAE=Common Toxicity Criteria Adverse Events; DLT=Dose-limiting toxicity; DOR=duration of response; ECOG=Eastern Cooperative Oncology Group; ePAS2=extended primary analysis set 2; MTD=maximum tolerated dose; ORR=overall response rate; PR=partial response; PS=performance status; q.d.=once-daily; RANO= Response Assessment in Neuro-Oncology; SAS3=supplementary analysis set 3; SD=stable disease.

Missed tablets and Compliance

Late doses (i.e., 4 or more hours after scheduled time) should be noted in the diary and taken as soon as remembered. Doses that are late by more than 6 hours (LOXO-TRK-14001 and SCOUT) or 8 hours (NAVIGATE) should be skipped and recorded in the dosing diary as missed.

A diary was kept to record dosing compliance, which was assessed at each clinic visit by means of a capsule count in the returned bottle, or liquid level verification in the solution bottle(s).

Efficacy outcome measures used in the economic model or specified in the scope

The primary efficacy outcome in the pooled analysis was the overall response rate (ORR) according to independent review, using RECIST version 1.1 (see Appendix P for RECIST criteria).

Table 4 lists all outcomes from the 3 larotrectinib studies and the pooled analysis and Table 7 provides details of measures and timings for those endpoints included in the model. Similar efficacy endpoints were analysed across the studies, however there was variation as to which endpoints were primary or secondary in each study (see Table 6).

With the exception of the SAS3 dataset (primary CNS tumours), Independent review committee (IRC) assessments served as the principal data source for response, time to response, time to best response, duration of response, disease control rate, and PFS (44).

Table 6. Efficacy endpoints by study (4, 5, 15, 44)

	Pooled Analysis of patients from LOXO-TRK-14001, SCOUT and NAVIGATE trials	NAVIGATE LOXO-TRK-15002 (NCT02576431)	Adult Phase I LOXO-TRK-14001 (NCT02122913)	SCOUT LOXO-TRK-15003 (NCT02637687)
Overall response rate independent review	Primary			
Overall response rate investigator assessment	Secondary	Primary	Secondary	Phase I: Secondary Phase II: Primary
Duration of response	Secondary	Secondary	Secondary	Secondary
Best overall response	Secondary	Secondary		
Disease Control Rate (or Clinical benefit rate)	Secondary	Secondary		
Time to response / best response	Secondary			
Time on treatment	Secondary			
Progression-free survival	Secondary	Secondary		
Overall survival	Secondary	Secondary		
Quality of life	As available	Exploratory		Secondary

Table 7. Relevant endpoints and measures in the pooled analysis / larotrectinib studies (4, 5, 15, 44)

Endpoint	Definition & timing of assessment / measure
Primary Efficacy Endpoint (pooled analysis)	
Overall Response Rate (ORR) by independent review committee (IRC)	Best overall response of confirmed CR or confirmed PR Response assessment was made as appropriate to tumour type using Response Evaluation Criteria in Solid Tumours (RECIST; version 1.1) for non-CNS tumours (47) or RANO for CNS tumours (48)(See Appendix P for RECIST and RANO criteria). Disease status was assessed according to relevant criteria at baseline and then on day 1 of every other cycle (for the duration of treatment LOXO-TRK-14001) or, for NAVIGATE and SCOUT, between cycles 1-12 and every 12 weeks thereafter until the end of treatment / disease progression). In Study LOXO-TRK-15002, patients enrolled into cohort 7 (primary CNS tumour) underwent radiographic disease assessment after every cycle for cycle 1 through cycle 4, followed by every other cycle from cycle 5 through cycle 13. Tumours were assessed by computed tomography (CT), magnetic resonance imaging (MRI), and, in the case of cutaneous lesions, clinically with electronic callipers. All tumour responses were confirmed at least 4 weeks after the initial response.
Secondary Endpoints (pooled analysis)	
Overall Response Rate (ORR) by investigator assessment	CR+PR See above for method – performed by investigators instead of IRC.
Disease Control Rate (DCR) [also referred to as Clinical Benefit Rate (CBR)]	Proportion of patients with confirmed best response of CR, PR or SD lasting ≥16weeks.
Duration of Response (DoR)	Defined as the number of months from the start of CR or PR (whichever response is recorded first) and subsequently confirmed to the first date that recurrent or progressive disease is documented or death.
Best overall response	The best response designation as of the data cut-off date for each patient recorded between the date of the first dose of larotrectinib and the date of documented disease progression per RECIST v1.1, the date of subsequent therapy or cancer-related surgery, or the data cut-off date, whichever occurred first. Patients who underwent surgical resection on therapy with no viable tumour cells and negative margins on post-surgical pathology report were considered a CR by surgery/pathology.
Time to response / best response	Time from therapy initiation to the date of confirmed response / best response.
Time on treatment	TOT (months) = (Last Dose/visit Date – First Dose Date + 1) / 30.4375
Progression-free survival (PFS)	Including PFS rate at 6 and 12 months after initiation of larotrectinib. Number of months from initiation of larotrectinib to the earlier of disease progression or death due to any cause.
Overall survival (OS)	Including survival rate at 12 months after initiation of larotrectinib. Number of months from the initiation of larotrectinib to the date of death due to any cause.
Safety	Adverse events (AEs) were monitored throughout the study and for 28 days after treatment and graded according to the National Cancer

	Institute Common Terminology Criteria for Adverse Events (CTCAE), version 4.0. Laboratory monitoring for toxicity and symptom directed neurological examinations for close monitoring of neurological toxicities were performed weekly during cycle 1 and every 4 weeks thereafter.
Exploratory endpoints	
Health Related Quality of Life (HRQoL): ***(Secondary endpoint in SCOUT)***	Baseline and subsequent quality of life instruments were administered during the same visits during which radiographic assessments were performed but were given prior to the patient learning the results of his or her restaging, as the radiographic results could influence the patient's response to the questionnaires. **NAVIGATE:** 1) Age ≥18 years: EORTC QLQ-C30 and EQ-5D. 2) Age 12 – 17 years: Paediatrics Quality of Life-Core Module (PedsQL-Core). **SCOUT** (phase II): 1) Infants 1-24mo.: PedsQL-Infant scale 2) ≥25 mo.: PedsQL-Core 3) Wong-Baker Faces Scale (FACES) to assess pain in patients 3 years and older. Minimally important differences (MID), (i.e. clinically meaningful) defined in literature include: - a difference of 10 for EORTC QLQ-C30 and EQ-5D-5L - a difference of 4.5 for Pediatric Quality of Life Inventory [PedsQL] total score). Within the submission the following definitions apply for EORTC QLQ-C30 and EQ-5D-5L (PedsQL) respectively: ▪ MID improvement: change from baseline ≥ 10 (4.5) ▪ MID no change/slight improvement: 0 ≤ change from baseline < 10 (4.5) ▪ MID slight deterioration: (−4.5) −10 < change from baseline < 0 ▪ MID deterioration: change from baseline ≤ (−4.5) −10 ▪ Not evaluable: if the change from baseline was not evaluable because either baseline value was missing or the postbaseline questionnaire was not available
European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire-Core 30 (QLQ-C30) (version 3.0)	EORTC QLQ-C30 is a well-validated instrument that assesses HRQoL in cancer patients. It includes 30 items, with scales evaluating physical (5 items), emotional (4 items), role (2 items), cognitive (2 items), and social (2 items) functioning, as well as global health status (3 items). Higher mean scores represent better functioning. There are also 3 symptom scales measuring nausea and vomiting (2 items), fatigue (3 items), and pain (2 items), and 6 single items assessing financial impact and various physical symptoms. A change of at least 10 points on the EORTC QLQ-C30 total score is considered clinically meaningful (49, 50) – referred to as minimally important difference (MID).
European Quality of Life 5-Dimension 5-Levels Health Questionnaire (EQ-5D-5L)	EuroQol/EQ-5D is a validated instrument consisting of the EQ-5D descriptive system and the EQ Visual Analogue Scale (EQ VAS). The EQ-5D 5 level version (EQ-5D-5L) is comprised of 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has 5 response levels: 1 (no problems) to 5 (unable). A change of at least 0.10 to 0.12 points on the EQ-5D index is considered clinically meaningful. For the EQ-5D VAS, higher scores represent better health status. A change of at least 7 points on EQ VAS is considered as clinically meaningful (51).
Paediatrics Quality of Life-Core Module (PedsQL-Core 4.0)	PedsQL Infant Scale is completed by the parent or caregiver. The PedsQL 4.0 Core Module inventory uses child self-reporting as a generic core measure integrated into disease-specific modules to provide one assessment. The Generic Core Scales for children/adolescents consist

	of 23 items and 4 dimensions (physical, emotional, social, and school functioning). A 5-point Likert scale from 0 (Never) to 4 (Almost always) is reported for each item.
Wong-Baker Faces Scale (FACES) – pain scale	FACES: used in patients ≥ 3 years old to assess pain. The scale shows a series of 6 faces ranging from a happy face at 0, or "no hurt," to a crying face at 10, which represents "hurts like the worst pain imaginable." Based on the faces and written descriptions, the subject chooses the face that best describes their level of pain (52).
Growth Modulation index (GMI) (*post hoc* analysis)	Uses each patient as his / her own control to compare the effects of larotrectinib versus the effect of the previous line of treatment the patient had received. A summary of the methodology / statistical approach used for this post hoc analysis is provided in Appendix Q. GMI was defined by Von Hoff (64, 65), and calculated as follows: $GMI_{Laro}=PFS_{Laro}/TTP_{-1}$ where PFS_{Laro} is the time from the date of the first dose of larotrectinib and the earliest date of documented disease progression or death from any cause (based on IRC assessed data); and TTP_{-1} is the time from the start of therapy to the date of disease progression on that therapy for the most recent prior systemic anti-cancer therapy. A GMI > 1.33 was defined by Von Hoff as the sign of drug activity. The everyday observation that underscores this approach is that TTP tends to become shorter with successive chemotherapy lines. Since successive TTPs tend to become shorter, a GMI >1.0 (or,more conservatively >1.33 to eliminate chance fluctuations) should be considered as a sign of activity.

AE=adverse events; CBR=clinical benefit rate; CNS=central nrevous system; CR=complete response; CT= computed tomography; CTCAE=Common Terminology Criteria for Adverse Events; DCR=disease control rate; DoR=duration of response; EORTC- QLQ-C30=European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; EQ-5D-5L=European Quality of Life 5-Dimensions 5-Levels Health Questionnaire; HRqol=Health-related quality of life; GMI=growth modulation index; IRC=independent review committee; MID=minimally important difference; MRI=Magnetic resonance imaging; ORR=overall response rate; OS=overall survival; PedsQL-Core=Paediatrics Quality of Life-Core Module; PFS=progression-free survival; PR=partial response; RANO= Response Assessment in Neuro-Oncology; RECIST=Response Evaluation Criteria in Solid Tumours; RECIST=Response Evaluation Criteria in Solid Tumours; SD=stable disease; TOT=time on treatment;

See section B.2.7 and Appendix E for details of pre-planned subgroup analyses.

Patient Baseline characteristics

Patient characteristics for the pooled analysis and individual clinical trials evaluating larotrectinib are summarised in Table 8.

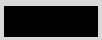
The pooled analysis datasets include 93 patients with solid non-CNS tumours and 9 patients with primary CNS tumours, all with an NTRK gene fusion.

The median age of patients with non-CNS solid tumours in the pooled analysis was [redacted] years (range [redacted] years) and for patients with primary CNS tumours was [redacted] years [redacted] ([redacted] years). Patients with primary CNS were [redacted] whereas patients with solid non-CNS tumours were [redacted], mostly between the ages of [redacted] and [redacted]. Most patients were 'White' - [redacted] patients in the non-CNS group and [redacted] in the primary CNS tumour group. There were slightly more men in the analysis. Most patients had an Eastern Cooperative Oncology Group (ECOG) performance status (PS) of [redacted]. All primary CNS tumour patients and almost all non-CNS patients had received prior cancer therapy (surgery, radiotherapy or systemic therapy) [redacted]. Larotrectinib was the initial systemic therapy for [redacted] of non-CNS tumour patients where no standard of care systemic treatment existed. [redacted] different tumour histologies are represented in the pooled analysis dataset, the most common tumour types being soft tissue sarcoma (n=[redacted]), salivary gland tumour (n=[redacted]), infantile fibrosarcoma (n=[redacted]), thyroid cancer (n=[redacted]), primary CNS cancer (n=[redacted]), and lung and melanoma cancer (n=[redacted] for each). The non-CNS tumours were mainly either [redacted] or [redacted] fusions, whereas most of the primary CNS tumours studied harboured an [redacted] fusion. [redacted] unique upstream fusion partners were identified.

The distributions of ages for patients enrolled in each study reflect the different selection criteria across the studies. Patients had to be ≥18 years of age to enroll in LOXO-TRK-14001, ≥12 years of age to enroll in NAVIGATE and <21 years of age to enroll in the SCOUT study.

Table 8. Baseline demographic and disease characteristics for pooled analysis and individual larotrectinib study populations (data cut-off 30th July 2018) (45)

Trial number (acronym) / Baseline Characteristics	Pooled Analysis of patients from LOXO-TRK-14001, SCOUT and NAVIGATE trials[a] N=102		NAVIGATE LOXO-TRK-15002 NCT02576431 N=83[a]	Adult Phase 1 LOXO-TRK-14001 NCT02122913			SCOUT LOXO-TRK-15003 NCT02637687		
	ePAS2 n=93	SAS3 n=9		NTRK N=10	Non-NTRK N=62	Total N=72 [a]	NTRK N=46	Non-NTRK N=8	N=54 [a]
Age, n (%)									
Median age, years (range)									
Mean									
< 2 yr									
2-<6 yr									
6-<12 yr									
12-<16 yr									
16-<18 yr									
18-<45 yr									
45-<65 yr									
65<75 yr									
≥ 75 yr									
Sex, n (%)									
Male									
Female									
Race, n (%)									
White									
Black or African American									
Asian									

Trial number (acronym) Baseline Characteristics	Pooled Analysis of patients from LOXO-TRK-14001, SCOUT and NAVIGATE trials[a] N=102		NAVIGATE LOXO-TRK-15002 NCT02576431 N=83[a]	Adult Phase 1 LOXO-TRK-14001 NCT02122913			SCOUT LOXO-TRK-15003 NCT02637687		
	ePAS2 n=93	SAS3 n=9		NTRK N=10	Non-NTRK N=62	Total N=72 [a]	NTRK N=46	Non-NTRK N=8	N=54 [a]
American Indian or Alaska Native									
Native Hawaiian or Pacific Islander									
Multiple / Other									
Declined to state/Not reported									
ECOG PS, n (%)									**PS Karnofsky / Lansky) n (%)**
0									
1									
2									
Not reported/unknown									
Primary tumour type, n (%)									
NSCLC									
IFS									
STS									
Colon									
Salivary gland									
Breast									
Pancreas									
Thymus									
Thyroid									
Bone sarcoma									
Cholangiocarcinoma									
Gastric									

Trial number (acronym) Baseline Characteristics	Pooled Analysis of patients from LOXO-TRK-14001, SCOUT and NAVIGATE trials[a] N=102		NAVIGATE LOXO-TRK-15002 NCT02576431 N=83[a]	Adult Phase 1 LOXO-TRK-14001 NCT02122913			SCOUT LOXO-TRK-15003 NCT02637687		
	ePAS2 n=93	SAS3 n=9		NTRK N=10	Non-NTRK N=62	Total N=72 [a]	NTRK N=46	Non-NTRK N=8	N=54 [a]
GIST									
Hepatic									
Melanoma									
Anal									
Appendix									
Cancer of unknown primary									
Endometrial									
Larynx									
Neuroblastoma									
Oral									
Ovarian									
Primary CNS									
Renal									
Congenital mesoblastic nephroma									
Ewing sarcoma									
Other									
Stage at initial diagnosis, n (%)									
I									
II									
III									
IV									
Not reported/Unknown									

Trial number (acronym) Baseline Characteristics	Pooled Analysis of patients from LOXO-TRK-14001, SCOUT and NAVIGATE trials[a] N=102		NAVIGATE LOXO-TRK-15002 NCT02576431 N=83[a]	Adult Phase 1 LOXO-TRK-14001 NCT02122913			SCOUT LOXO-TRK-15003 NCT02637687		
	ePAS2 n=93	SAS3 n=9		NTRK N=10	Non-NTRK N=62	Total N=72 [a]	NTRK N=46	Non-NTRK N=8	N=54 [a]
Disease extent at enrollment n (%)									
Locally advanced									
Metastatic									
Other / not reported									
Prior cancer therapy - Yes, n (%)									
Surgery									
Radiotherapy									
Systemic therapy									
0 prior systemic									
1-2									
≥3									
Mean no. prior systemic									
Median no. prior systemic									
***NTRK* gene fusion status, n (%)**									
None / not known									
NTRK1									
NTRK2									
NTRK3									
Inferred NTRK3									
***NTRK* gene fusion partner, n (%)**									
Fusion Partner not reported									

Trial number (acronym) Baseline Characteristics	Pooled Analysis of patients from LOXO-TRK-14001, SCOUT and NAVIGATE trials[a] N=102		NAVIGATE LOXO-TRK-15002 NCT02576431	Adult Phase 1 LOXO-TRK-14001 NCT02122913			SCOUT LOXO-TRK-15003 NCT02637687		
	ePAS2 n=93	SAS3 n=9	N=83[a]	NTRK N=10	Non-NTRK N=62	Total N=72[a]	NTRK N=46	Non-NTRK N=8	N=54[a]
ETV6-NTRK3									
Inferred ETV6-NTRK3									
TPM3-NTRK1									
LMNA-NTRK1									
IRF2BP2-NTRK1									
CTRC-NTRK1									
MYO5A-NTRK3									
PPL-NTRK1									
SPECC1L-NTRK3									
SQSTM1-NTRK1									
SQSTM1-NTRK3									
TPM4-NTRK3									
TPR-NTRK1									
TRIM63-NTRK1									
AFAP1-NTRK1									
BCR-NTRK2									
GNAQ-NTRK2									
GON4L-NTRK1									
NFASC-NTRK1									
PLEKHA6-NTRK1									
TRAF2-NTRK2									
EML4-NTRK3									
PDE4DIP-NTRK1									
STRN-NTRK2									

Trial number (acronym) Baseline Characteristics	Pooled Analysis of patients from LOXO-TRK-14001, SCOUT and NAVIGATE trials[a] N=102		NAVIGATE LOXO-TRK-15002 NCT02576431 N=83[a]	Adult Phase 1 LOXO-TRK-14001 NCT02122913			SCOUT LOXO-TRK-15003 NCT02637687		
	ePAS2 n=93	SAS3 n=9		NTRK N=10	Non-NTRK N=62	Total N=72 [a]	NTRK N=46	Non-NTRK N=8	N=54 [a]
KANK-NTRK2									
KANK2-NTRK2									
SPECC1L-NTRK2									
AGTPBP1-NTRK2									
EPS15-NTRK1									
DIAPH1-NTRK1									
RBPMS-NTRK2									

CNS=central nervous system; ECOG=Eastern Cooperative Oncology Group; GIST=gastrointestinal stromal tumour; IFS=infantile fibrosarcoma; n=number; no.=number; NSCLC=non-small cell lung cancer; NTRK=neurotrophin receptor tyrosine kinase; PS=performance status; STS=soft tissue sarcoma; yr=year;
[a] ePAS2 and SAS3: efficacy evaluable patients; demographics for individual studies based on safety analysis sets;
[b] Cholangiocarcinoma, also known as bile duct cancer. Recorded under 'Other' in NAVIGATE trial.
Note: due to rounding, percentages may not total to exactly 100%

B.2.4 Statistical analysis and definition of study groups in the relevant clinical effectiveness evidence

Analysis sets

Analyses were performed according to the intention-to-treat principle. To provide a better estimate of treatment efficacy in patients with NTRK gene fusions, data from the 3 clinical studies (LOXO-TRK-14001, NAVIGATE and SCOUT) were pooled. The rationale to pool efficacy data from patients with NTRK gene fusions across all three clinical studies was made during the clinical development programme, based on the rarity of NTRK fusion cancer, the heterogeneity of the tumour histologies, and in conjunction with global regulatory advice (15). The pooled population was representative of the anticipated target population for larotrectinib in regard to the range of patient ages ([redacted] years, median [redacted] years) and a broad range of tumour types represented ((ePAS2: [redacted] years, median [redacted] years; SAS3[redacted] years, median [redacted] years).

The primary population for analysis for the clinical effectiveness of larotrectinib in this submission comprises the ePAS2 (extended primary analysis set 2; n=93) plus SAS3 (supplementary analysis set 3; n=9) datasets. This includes all evaluable patients with documented NTRK gene fusion, advanced solid (ePAS2) or primary CNS tumours (SAS3) with measurable disease as at the data cut-off 30 July 2018 (n=102) (see Table 9). Participants were required to have received at least one dose of larotrectinib. The population for safety analysis consists of all patients with an NTRK solid or primary CNS tumour who have received at least one dose of larotrectinib in any of the three larotrectinib studies, regardless of whether they were evaluable for efficacy. These datasets have been used for the EMA marketing authorisation for larotrectinib. At the time of the data cut-off for the ePAS2 and SAS3 datasets, all 3 trials were ongoing, with patients still being treated and new patients being enrolled.

The original primary analysis set (PAS) of the first 55 evaluable patients, using the cut-off date of 17 July 2017, was used to support the regulatory application to the FDA. As patient numbers increase in the three individual clinical studies, the pooled analysis dataset is extended and analysed at different cut-off timepoints e.g. 19th Feb 2018

(ePAS1) and 30th July 2018 (ePAS2). The approach for pooling of the data was similar to that for the original integrated analysis for the FDA, hence the original statistical analysis plan (SAP) for the integrated analysis (see protocol appendix, Drilon 2018 (15)) was not updated. The ePAS2 dataset includes an additional 38 patients who were recruited after the 55th patient until 30 July 2018 and, apart from this, fulfilled all criteria for the original PAS. The SAS3 dataset includes patients who met criteria for the PAS / ePAS2, except for having a primary CNS tumour. SAS3 includes 9 paediatric and adult patients with primary CNS tumours and utilised investigator assessments of disease as opposed to central assessment.

Table 9. Definition of relevant larotrectinib data analysis sets included within this submission (44, 45)

Analysis set	Definition	Number of valid patients in treatment group			
		From NAVIGATE LOXO-TRK-15002 (NCT02576431)	From Adult Phase I LOXO-TRK-14001 (NCT02122913)	From SCOUT LOXO-TRK-15003 (NCT02637687)	Pooled analysis Total
Population for efficacy analysis (NICE submission)					
ePAS2	Patients (≥1 dose larotrectinib) with NTRK fusion cancer using RECIST, version 1.1 at baseline (excluding primary CNS tumours and patients without measurable disease) (as at 30 July 2018)	[redacted]	[redacted]	[redacted]	N=93
SAS3	Patients (≥1 dose larotrectinib) with primary CNS tumours, who otherwise met PAS / ePAS2 eligibility criteria and were enrolled before 30 July 2018	[redacted]	[redacted]	[redacted]	N=9
Original PAS (basis of FDA approval)	Patients (≥1 dose larotrectinib) with NTRK fusion cancer (excluding primary CNS tumours and patients without measurable disease using RECIST, version 1.1 at baseline) (as at 17 July 2017)	N=35	N=8	N=12	N=55
Safety analysis set	**All treated patients (≥1 dose larotrectinib) with NTRK fusions, regardless if evaluable for efficacy or not (as at 30 July 2018)**	[redacted]	[redacted]	[redacted]	**N=137**

ePAS=extended primary analysis set; N=number; PAS=primary analysis set;

Overview of statistical analyses

Table 10. Summary of statistical analyses for the pooled analysis from LOXO-TRK-14001, SCOUT and NAVIGATE trials (15)

Trial number (acronym)	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
Pooled Analysis of patients from LOXO-TRK-14001, SCOUT and NAVIGATE trials	Under the planned primary analysis of effectiveness, a true ORR of ≥ 50% is hypothesised when larotrectinib is administered to patients with NTRK fusion cancers.	**Statistical analyses were performed using SAS® version 9.2 or later.** **NB. SAS3 analysis set includes 9 paediatric and adult patients with primary CNS tumours and utilised investigator assessments of disease as opposed to central assessment.** **Primary statistical method:** - Performed according to the ITT principle - using data as at cut-off 30 July 2018 The primary endpoint was ORR determined by an independent review committee (IRC). Response rates were summarised descriptively by number and percentage. The agreement rate between IRC and Investigator assessments of response was tabulated. The best overall response was summarised descriptively by number and percentage. Point estimates were accompanied by a two-sided 95% exact binomial CI using the Clopper-Pearson method. **Supportive analyses of primary endpoint:** -ORR was calculated based on local investigator assessment of response and agreement rate between IRC and investigator response assessment calculated. - Change in tumour burden was calculated for each patient as the percentage change from baseline in the sum of diameters of target tumour lesions at each time point. The best tumour-burden change was summarised descriptively, by calculating the median and interquartile range across patients, and presenting as a waterfall plot. Spider plots displayed change in tumour burden over time	A sample size of 55 patients was estimated to provide 80% power to achieve a lower boundary of the 2-sided 95% exact binomial CI about the estimated ORR exceeding 30%. Ruling out a lower limit of 30% for ORR was considered clinically meaningful and consistent with the estimated response rates seen with approved targeted therapies in genetically-defined patient populations who have progressed on prior therapies. Under the primary analysis, the lower limit of the 95% CI would exceed 30% when the estimated ORR was 46% or greater (Clopper-Pearson method).	**Handling of missing data:** No imputations were performed on missing data for data summarised over time by visit. All analyses were based on observed data only. The effective sample sizes at each assessment visit were based on the total number of patients with non-missing data for the parameter of interest at that visit. **Censoring:** DOR and PFS were right-censored for patients who: -had amputation, surgical resection of tumour or subsequent anticancer therapy in the absence of documented disease progression, or -died or had documented disease progression after missing two or more consecutively scheduled disease assessment visits, or -were alive and without documented disease progression on or before the data cut-off date, or (PFS only) had no post-baseline disease assessments unless death occurred prior to the first planned assessment.

		for each patient. Swimmer plots showed occurrence of clinical outcomes of interest over time (e.g. TTR, DOR, disease progression, treatment discontinuation, death). **Secondary endpoints:** -The Kaplan–Meier method was used to evaluate duration of response (DoR), progression-free survival (PFS) and overall survival (OS), with the two-sided 95% CI about the median calculated using Greenwood's formula. - Analysis of disease control rate was based on the methods described for ORR. IRC assessments served as the principal data source for TTR, TTBR, DOR, DCR and PFS. Supplemental analyses based on local investigator assessments were provided. Time to response (TTR) & time to best response (TTBR) (calculated for responders only) were summarised descriptively by calculating the median, interquartile range, and minimum and maximum values. The number and percentage of patients by the milestone time points were tabulated. Kaplan-Meier curves were used to graphically present the time to response/time to best response distribution over time.		

CI=confidence interval; DCR=disease control rate; DoR=duration of response; HR=hazard ratio; IRC=independent review committee; ITT=intention-to treat; ORR=overall response rate; OS=overall survival; PFS=progression-free survival; TT(B)R=time to (best) response;

See Appendix D for patient disposition in the pooled analysis and individual larotrectinib studies as at data cut-off 30 July 2018.

***Post hoc* analysis**

A *post hoc* analysis using each patient as his / her own control was performed to compare the effects of larotrectinib versus the effect of the previous line of treatment the patient had received. This analysis aims to test the hypothesis that if larotrectinib has an anti-tumour effect, it will change the natural history of the disease. Given the natural history, one would expect time to progression (TTPn) to be shorter on the nth treatment compared to the TTP on n-1th treatment (TTPn-1) (53-56). Therefore, if TTP on larotrectinib is greater than TTP on the therapy prior to larotrectinib, then it is likely that larotrectinib is having an effect on natural history of that patient's tumour(57). Results of this post hoc analysis are presented in Table 23. Clinical effectiveness results and a summary of the methodology / statistical approach used is provided in Appendix Q.

B.2.5 Quality assessment of the relevant clinical effectiveness evidence

As highlighted in section B2.2, evidence of larotrectinib's clinical effectiveness is derived from three non-randomised studies (NRS), data from which inform the pooled analysis.

No specific recommendations are made within the NICE User Guide as to a preferred critical appraisal tool for quality assessment of NRS and there appears to be no consensus regarding the most appropriate critical appraisal tool for NRS (58). As a validated tool for assessing the risk of bias in non-randomised studies, the Downs and Black checklist (59, 60) was selected.

The situation for larotrectinib is unorthodox when compared with many NICE Technology appraisals. All studies providing evidence for larotrectinib in the submission are ongoing and many aspects yet to be published. Critical appraisal has therefore been performed with reference to multiple sources e.g. publications of interim analyses, the Summary of Clinical Efficacy as submitted to the EMA, study protocols (which have been published as a supplementary appendix to Drilon 2018 (15)). The clinical programme for larotrectinib (including study designs) was developed

in collaboration with global health authorities. Publication of study methodology, analyses and results from all studies within the programme is approached with the same rigour and transparency.

The larotrectinib studies were evaluated for: quality of reporting (10 items); external validity (3 items); bias (7 items); and confounding (6 items) using the sub scales of Downs and Black scoring system.

The studies scored 14-15 points (summary Table 11; detailed Table 80). The study question was specifically stated and well defined, with appropriate outcome measures in all included studies. The intervention was clearly defined, and adverse events reported. Patient characteristics and study findings were well described, including patients lost to follow-up. In terms of external validity across the studies, it was considered that the patient pool was representative of that found for the disease in the general population. Patients received treatment at their usual hospital, however there was no information on patients that were asked to participate in the study but declined or were screening failures. Although statistical tests used were appropriate, treatment compliance was reliable and outcome measures accurate, there was significant risk of bias due to the studies being non-randomised, and the analyses being interim. Also, confounding factors and measures for internal validity were not discussed sufficiently.

Table 11. Downs and Blacks Checklist score for larotrectinib studies (59)

	Reporting										Score	Total Score
Study name	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**		
NAVIGATE (NCT02576431, LOXO-TRK-15002)	1	1	1	1	0	1	1	1	1	0	8	
LOXO-TRK-14001 (NCT02122913)	1	1	1	1	0	1	1	1	1	0	8	
SCOUT (NCT02637687, LOXO-TRK-15003)	1	1	1	1	0	1	1	1	1	0	8	
	External Validity			**Internal validity - bias**							**Score**	
Study name	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**		
NAVIGATE	1	0	1	0	0	0	0	1	1	1	5	
LOXO-TRK-14001	1	0	1	0	0	0	0	1	1	1	5	
SCOUT	1	0	1	0	1	0	0	1	1	1	6	
	Internal validity – confounding (selection bias)						**Power**				**Score**	
Study name	**21**	**22**	**23**	**24**	**25**	**26**	**27**					
NAVIGATE	0	1	0	0	0	0	0				1	**14**
LOXO-TRK-14001	0	1	0	0	0	0	0				1	**14**
SCOUT	0	1	0	0	0	0	0				1	**15**

From: Downs SH, Black N. (1998) The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health. 52(6): 377-384.

The larotrectinib trials were conducted across North America, Europe and Asia, taking place in 15 countries including the UK. Although no patients from the UK are included in the pooled analysis, the design of the studies enabled a study population generally reflective of patients who would be seen within clinical practice and considered for treatment with larotrectinib in England.

NTRK gene fusions can be found across many different types of solid tumours. In some tumours, the incidence of an NTRK gene fusion is very low e.g. <1% to 3% in lung adenocarcinomas or colon cancer (10, 61), whereas in others, an NTRK gene fusion can be a defining characteristic of the tumour and found in most cases e.g. MASC (62), IFS (63-65). In addition, patients with NTRK fusion cancer represent a diverse group, ranging from infants to adults. The manufacturer collaborated with global health authorities to devise a feasible drug development approach for larotrectinib that recognised the rarity of NTRK fusion cancers and their tumour heterogeneity. Regulatory authourities encouraged broad but harmonised eligibility criteria across protocols that accommodated diverse tumour types in both adult and paediatric patients with NTRK fusion cancers. This strategy was confirmed by

examining the baseline characteristics of the population with NTRK gene fusions included in the pooled analysis (see Table 8), where the age of patients ranges from [redacted] months to [redacted] years ((ePAS2: [redacted] years; SAS3: [redacted] years), and the range of tumour types treated corresponds closely with literature on NTRK gene fusion detection (6-8). Also, the large number of different NTRK fusion partners (n=[redacted]) in patients in the larotrectinib analysis is characteristic of published findings.

Current treatments for different solid tumour cancers include surgery, radiotherapy, and systemic therapies such as chemotherapy, hormone therapy, or immunotherapy. Larotrectinib studies recruited patients with metastatic (pooled analysis, ePAS2 [redacted]%) or locally advanced disease (pooled analysis, ePAS2 [redacted]%), with most patients having received prior therapy for their cancer and been treated with therapies such as those listed above.

With exception of the initial dose-finding phases in the SCOUT and LOXO-TRK-14001 studies where presence of an NTRK gene fusion was not a pre-requisite, the population specified in the scope of this technology appraisal - also [redacted] - corresponds with the expansion phase / phase II larotrectinib study inclusion criteria i.e adult and paediatric patients with NTRK fusion-positive solid tumours who have a locally advanced or metastatic disease or where surgical resection is likely to result in severe morbidity, and have no satisfactory treatment options. Within the studies, most patients (n=[redacted]%) received a starting dose of larotrectinib in line with the recommended licenced dose, however [redacted]% patients initially started with a different dose due to part of the SCOUT and LOXO-TRK-14001 studies incorporating initial dose-finding elements. Treatment was continued in the study in accordance with the label posology.

Patients were recruited and seen within the same setting of the hospitals and cancer centres that they would usually attend, regardless of trial participation, and outcomes and disease assessment during the trial was carried out in line with that of normal practice.

See section B.2.13.2 for more detailed quality assessment.

Please see Appendix D for the full detailed quality assessment.

B.2.6 Clinical effectiveness results of the relevant trials

Notes:

1) Unless otherwise specified, the results and analyses of all efficacy and safety outcomes presented in this submission are based on the most recent data cut-off of 30 July 2018.

2) The reporting of efficacy results will focus on the pooled analysis of patients with NTRK fusion-positive tumours from the three larotrectinib studies: LOXO-TRK-14001, NAVIGATE (LOXO-TRK-15002) and SCOUT (LOXO-TRK-15003). A summary of available efficacy results from the individual studies for the 19th February 2018 cut-off is reported in Appendix O.

Summary of efficacy in the pooled analysis

The pooled analysis met its primary efficacy endpoint at the data cut-off of 17 July 2017 and was published in 2018 (15). Analysis in the updated and expanded dataset (ePAS2; n=93) - the main focus of this submission - further confirmed the robust effectiveness of larotrectinib in NTRK fusion-positive solid non-CNS tumours.

According to independent review, [redacted]% (95% CI: [redacted]) of patients with NTRK fusion-positive non-CNS solid tumours treated with larotrectinib exhibited an objective antitumour response, regardless of tumour type. Best response included [redacted]% (n=[redacted]) with a CR, [redacted]% (n=[redacted]) with a PR, [redacted]% (n=[redacted]) with SD, [redacted]% (n=[redacted]) with PD, and [redacted]% (n=[redacted]) who could not be evaluated due to early withdrawal for clinical deterioration.

Response to larotrectinib was rapid and durable, with median time to response at [redacted] months (range: [redacted]). After a median follow-up of [redacted] months, the median duration of response (DOR) [redacted]. As of the data cutoff date, among the [redacted] patients who had a response, [redacted]% (n=[redacted]) were still in response.

Median change in tumour size was a decrease of [redacted]%. Tumour shrinkage provides a further significant, and potentially life-changing benefit of treatment, particularly in children where larotrectinib treatment enabled an increased rate of limb sparing surgery, avoiding amputations or other disfiguring surgery.

In patients with NTRK fusion-positive primary CNS tumours (SAS3 dataset), to date, [redacted] ([redacted]%) has been observed (time to response [redacted]_months), where, after a follow-up of two months the patient was alive and well. [redacted] patients ([redacted]%) have achieved stable disease, and in [redacted] of these this has lasted [redacted].

Survival data, although immature and analysis ongoing, supports durability of larotrectinib effect. In ePAS2 (30 July 2018), the median PFS is [redacted] months (95% CI: [redacted]) (Kaplan-Meier). Six-month and one-year PFS rates are [redacted]% (95% CI: [redacted]) and [redacted]% (95% CI: [redacted]) respectively. Median PFS in the SAS3 patient group is [redacted] months (95% CI: [redacted]) with 6-month PFS rate at [redacted]% (95% CI: [redacted]). The median duration of overall survival is

[redacted] after a median follow-up of [redacted] months (25th: [redacted], 75th [redacted]) for ePAS2 ([redacted] patients alive; [redacted]%) and a median follow-up of [redacted] (25th: [redacted], 75th [redacted]) for SAS3 ([redacted] patients alive; [redacted]%). One-year OS rate for ePAS2 is [redacted]% (95%: [redacted]) and was [redacted] in SAS3. **Given the current level of uncertainty*, Bayer proposes that whilst data mature, larotrectinib is made available in a timely manner through the Cancer Drugs Fund.***

Patients with NTRK fusion-positive cancer also experienced a rapid and sustained improvement in health-related quality of life (HRQoL) compared to baseline, which occurred within [redacted] months of the start of treatment. For the majority of patients, scores [redacted] that translates into a meaningful improvement to the patient.

Primary Efficacy Outcome

Overall Response Rate (ORR) by independent review committee (IRC) assessment

As at 30 July 2018, the independently assessed ORR in patients with non-CNS NTRK fusion-positive tumours [ePAS2 pooled analysis set] was [redacted]% ([redacted]/93). This included [redacted] patients ([redacted]%) with complete responses (CR), [redacted] ([redacted]%) of which was a surgical CR (sCR). [redacted] percent of patients (n=[redacted]) experienced a partial response (PR), while stable disease (SD) was achieved in [redacted] patients ([redacted]%).

Patients with primary CNS tumours were analysed separately from the ePAS2 patients (SAS3 dataset; n=9) and in these patients, disease was not independently assessed. The investigator-assessed ORR was [redacted]%. [redacted] patients achieved stable disease ([redacted]%).

Table 12. Summary of overall response rate (30 July 2018 data cut-off) (45)

	ePAS2 (IRC) n=93	ePAS2 [a] (INV) n=93	**SAS3 (INV) N=9**
ORR (CR+sCR+PR) [95% CI] [b]	[redacted]	[redacted]	[redacted]
CR, confirmed	[redacted]	[redacted]	[redacted]
sCR	[redacted]	[redacted]	[redacted]
PR, confirmed	[redacted]	[redacted]	[redacted]
Stable disease ≥ 16 weeks	[redacted]	[redacted]	[redacted]
Stable disease <16 weeks	[redacted]	[redacted]	[redacted]
Progressive disease	[redacted]	[redacted]	[redacted]
Not evaluable	[redacted]	[redacted]	[redacted]
Disease Control rate (DCR) [95% CI] [b]	[redacted]	[redacted]	[redacted]

CR=complete response; DCR=disease control rate; INV=investigator assessment; IRC=independent review committee; PR=partial response; sCR=surgical complete response;
[a] Agreement by IRC and investigator n=[redacted] ([redacted]%); CR/sCR/PR by investigator, non-responder by IRC n=[redacted] ([redacted]%); CR/sCR/PR by IRC, non-responder by investigator n=[redacted] ([redacted]%);
[b] 95% confidence interval was calculated using Clopper-Pearson method

Note. For ePAS2 disease response was assessed using RECIST (version 1.1). For SAS3 disease response was assessed using RANO (see Appendix P).

Responses were observed in a wide range of tumour types, regardless of patient age, or NTRK gene fusion (see Appendix E for results of subgroup analysis of ORR).

Secondary Efficacy Outcomes

ORR by investigator assessment

Table 12 also presents ORR as determined by investigator assessment. There was [redacted]% agreement rate between the IRC assessment and the Investigator assessment of tumour response.

Disease Control Rate (DCR)

Disease control rate (defined as the proportion of patients with best overall response of confirmed CR, surgical CR, PR, or stable disease lasting 16 weeks or more following the initiation of larotrectinib) was [redacted]% (95% CI: [redacted]) for the ePAS2 (per IRC review) and [redacted]% (95% CI: [redacted]) by investigator assessment. DCR in the primary CNS tumour group (SAS3) was [redacted]% (95% CI: [redacted]) (see Table 12).

The median change in tumour size was a decrease of [redacted]%. The impact of larotrectinib treatment on tumour size is shown in the following waterfall plots (see Figure 9 and Figure 10) within the images within the five cases presented at the end of this section (Figure 20 to Figure 24). Evidence suggests that even in patients without an objective response, benefit may still be derived from receiving larotrectinib treatment, due to tumour shrinkage and disease stabilisation. This is further illustrated when examining cases of tumour types with no objective responses per IRC (see Appendix E).

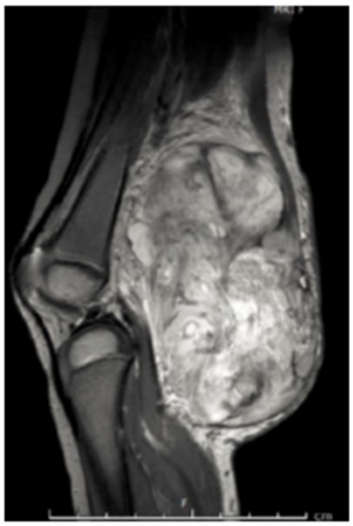
The impressive reductions of tumour size achieved by larotrectinib provides a further significant, and potentially life-changing benefit of treatment, particularly in children. Some paediatric patients with NTRK fusion-positive cancer require surgery that may result in the functional loss of a limb (e.g. patients with infantile fibrosarcoma). In the SCOUT trial, [redacted] patients were listed as having no other curative options besides amputation or disfiguring surgery. Larotrectinib treatment enabled an increased rate of limb sparing surgery. In all [redacted] patients, amputation was avoided). After larotrectinib, [redacted] patients required no surgical procedures, [redacted] patients underwent resection, but amputation was avoided, [redacted] patients had biopsy but no other surgery, [redacted] had colostomy takedown and port removal, and [redacted] had central venous catheter (CVC) placement. As at July 30, 2018 data cut-off, [redacted] of the [redacted] patients remained on therapy with larotrectinib, [redacted] patients discontinued therapy after their resection, [redacted] patient discontinued therapy due to family preference, and only [redacted] patients discontinued larotrectinib due to disease progression. Larotrectinib treatment therefore enables paediatric patients to avoid disfiguring surgery, such as amputation, which can have devastating, lifelong consequences (39, 66, 67).

Figure 8. Infantile fibrosarcoma case study – reduction in tumour mass

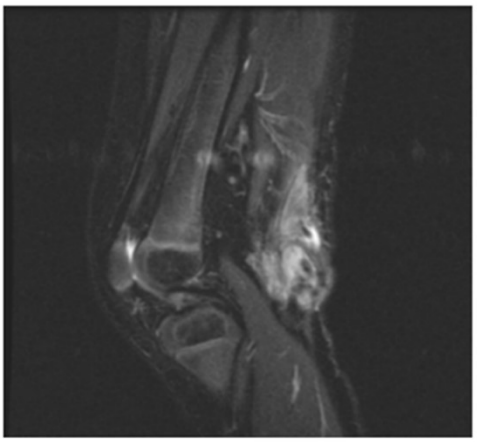
Patient B, Age 2, diagnosed with infantile fibrosarcoma

Underwent 2 cycles of vincristine, actinomycin-D and cyclophosphamide before regression and potential need for leg amputation.

After 4 cycles of larotrectinib, patient was referred for surgery. Patient had a complete response with clear margins, with no functional deficit post-surgery.



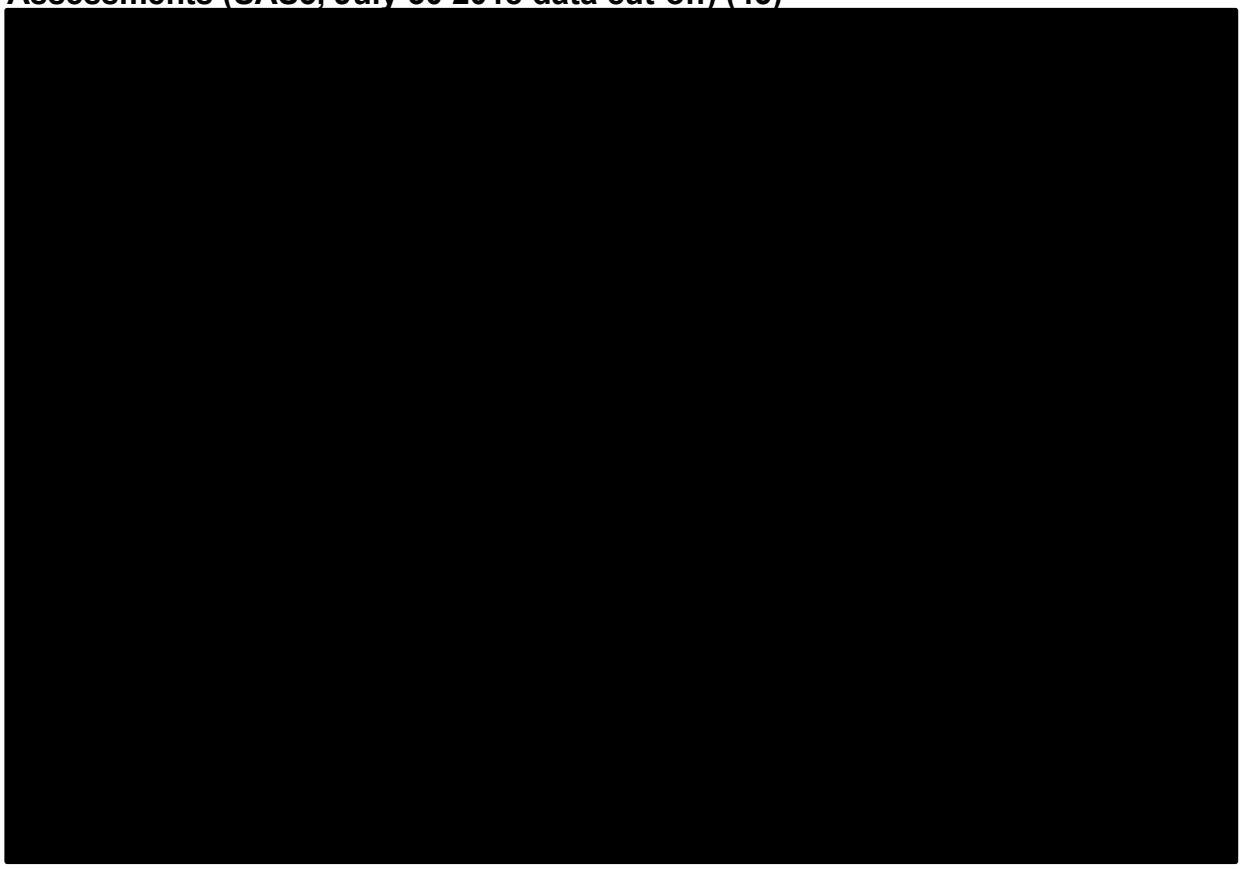
Baseline



Cycle 3

Figure 9. Waterfall Plot of Best Change in Tumour Size Based on IRC Assessments (ePAS2, July 30 2018 data cut-off) (45)

Figure 10. Waterfall Plot of Best Change in Tumour Size Based on Investigator Assessments (SAS3, July 30 2018 data cut-off) (45)



Duration of response

ePAS2: After a median follow-up of [redacted] months (IQR: [redacted]), the median duration of response [redacted]. At the time of the 30 July 2018 data cut-off, [redacted] patients in the ePAS2 had achieved a response by IRC assessment. Of these, [redacted]% of patients were still in response (see Table 13). The probability of retaining a response at the 6-month milestone was [redacted]% (95% CI: [redacted]) and at the 12-month milestone the probability was [redacted]% (95% CI: [redacted]) by Kaplan-Meier estimate (see Figure 11).

Duration of response measures based on investigator assessment were slightly [redacted] (due to [redacted] additional patients in the analysis) than those by IRC assessment. After a median follow-up of [redacted] months (IQR: [redacted]), the median duration of response [redacted].

SAS3: The [redacted] with a response in the primary CNS group was alive and well after a follow-up of [redacted] months at the 30 July 2018 data cut-off.

Table 13. Duration of response by IRC and investigator assessment for larotrectinib pooled analysis sets (30 July 2018 data cut-off) (45)

	ePAS2 (IRC) n=93	ePAS2 (INV) n=93	**SAS3 (INV) N=9**
Responding patients	[redacted]	[redacted]	[redacted]
Censored (still in response), n (%)	[redacted]	[redacted]	[redacted]
Disease progression	[redacted]	[redacted]	[redacted]
Duration of response			
≤ 6 months	[redacted]	[redacted]	[redacted]
>6 to 12 months	[redacted]	[redacted]	[redacted]
>12 to 18 months	[redacted]	[redacted]	[redacted]
>18 to 24 months	[redacted]	[redacted]	[redacted]
>24 months	[redacted]	[redacted]	[redacted]
Median Duration of response (months) [95% CI] [c] Minimum, Maximum	[redacted]	[redacted]	[redacted]

CI=confidence interval; ePAS=extended primary analysis set; NE=not estimable; SAS=supplementary analysis set; sCR=surgical complete response;
[a] Alive without documented disease progression n=[redacted] ([redacted]%); Surgical resection of tumour without sCR n=[redacted] ([redacted]%);
[b] Alive without documented disease progression n=[redacted] ([redacted]%); Surgical resection of tumour without sCR n=[redacted] ([redacted]%);
[c] 95% confidence interval was calculated using Greenwood's formula.

Figure 11. Duration of Response based on IRC Assessment in the ePAS2 (30 July 2018 data cut-off) (45)

Duration of response is defined as the time from the start date of CR or PR (whichever response is recorded first) to the earlier of documented disease progression or death due to any cause. Disease assessments were performed by investigators using RECIST, version 1.1. Vertical tick marks represent censored patients.

Time to response

Median time to response for the [redacted] responder patients in the ePAS2 with a confirmed response according to IRC assessment was [redacted] months (range: [redacted]) which corresponds to the protocol-mandated initial tumour assessment of response at 8 weeks ((15)). Median time to response for the [redacted] in SAS3 was [redacted] months (45). Time to best response was similar, with most patients achieving their best response within [redacted] months of initiating larotrectinib treatment. Such a quick time to response enables a rapid onset of patient benefit and an early understanding of whether the therapy is effective.

Table 14. Time to first / best response in the pooled analysis (30 July 2018 data cut-off) (45)

	ePAS2 (IRC) n=93		ePAS2 (INV) n=93		**SAS3 (INV) N=9**	
	First response	**Best response**	First response	Best response	**First response**	**Best response**
Responding patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Median time (months)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
25th, 75th percentiles	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Range	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Time of response						
≤ 2 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
>2 to 4 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
>4 to 6 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
>6 to 9 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
>9 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

The analysis of outcomes, based around disease response, presented so far highlight the ability of larotrectinib to achieve rapid, effective and durable responses in NTRK fusion-positive solid tumours. The following swimmer plots present a summary of each patient's journey while receiving larotrectinib, clearly showing the high response rate, quick time to response and the subsequent durability of most responses. These results are noteworthy given that most patients in the pooled analysis had metastatic disease, had received prior surgery and / or radiotherapy for their cancer and a mean of [redacted] prior systemic therapies.

Figure 12. Swimmer Plot of Time to Response and Overall Treatment Duration (ePAS2, July 30 2018 data cut-off) (45)

Figure 13. Swimmer Plot of Time to Response and Overall Treatment Duration (SAS3, July 30 2018 data cut-off) (45)

Time on treatment (45)

Median time on treatment in the ePAS2 was [redacted] months (range: [redacted]7 months) and [redacted] months (range: [redacted] months) for the SAS3. In the ePAS2, [redacted]% patients had received larotrectinib for 12 months or more and [redacted]% had received larotrectinib for 18 months or more, with follow-up ongoing at the time of the analysis. To date, [redacted] patients in the SAS3 have received treatment for ≥ 12 months.

Progression-free survival

By Kaplan-Meier methodology, the median PFS is [redacted] months (95% CI: [redacted]) for patients in ePAS2 (median follow-up [redacted] months [[redacted]] (see Figure 14). Six-month and one-year PFS rates are [redacted]% (95% CI: [redacted]) and [redacted]% (95% CI: [redacted]) respectively (45). Median PFS in the SAS3 patient group is [redacted] months (95% CI: 2[redacted]) with 6-month PFS rate at [redacted]% (95% CI: [redacted]).

Table 15. PFS by IRC and investigator assessment for larotrectinib pooled analysis sets (30 July 2018 data cut-off) (45)

	ePAS2 (IRC) n=93	ePAS2 (INV) n=93	**SAS3 (INV) N=9**
Progression status (as of patient's last disease assessment on or before 30 July 2018)			
Censored (still in response), n (%)	[redacted]	[redacted]	[redacted]
Disease progression	[redacted]	[redacted]	[redacted]
Duration of progression-free survival			
≤ 6 months	[redacted]	[redacted]	[redacted]
>6 to 12 months	[redacted]	[redacted]	[redacted]
>12 to 18 months	[redacted]	[redacted]	[redacted]
>18 to 24 months	[redacted]	[redacted]	[redacted]
>24 months	[redacted]	[redacted]	[redacted]
Median duration of PFS (months) [95% CI] [d] Minimum, Maximum	[redacted]	[redacted]	[redacted]

CI=confidence interval; ePAS=extended primary analysis set; INV=investigator assessment; IRC=independent review committee; NE=not estimable; PFS=progression-free survival; SAS=supplementary analysis set; sCR=surgical complete response;

[a] Alive without documented disease progression n=[redacted] ([redacted]%); Surgical resection of tumour without sCR n=[redacted] ([redacted]%); No evaluable post-baseline disease assessments n=[redacted] ([redacted]%)

[b] Alive without documented disease progression n=[redacted] ([redacted]%); Surgical resection of tumour without sCR n=[redacted] ([redacted]%); No evaluable post-baseline disease assessments n=[redacted] ([redacted]%)

[c] Alive without documented disease progression n=[redacted] ([redacted]%); No evaluable post-baseline disease assessments n=[redacted] ([redacted]%)

[d] 95% confidence interval was calculated using Greenwood's formula.

Figure 14. Kaplan-Meier Plot of PFS (ePAS2, IRC assessment; 30 July 2018 data cut-off) (45)

Figure 15. Kaplan-Meier Plot of PFS (SAS3, investigator assessment; 30 July 2018 data cut-off) (45)

Overall survival (OS)

As at 30 July 2018, [redacted] (of 93; [redacted]%) patients in ePAS2 and [redacted] ([redacted]%) patients in SAS3 are alive. The median duration of overall survival is [redacted] for either dataset after a median follow-up of [redacted] months (25th: [redacted], 75th [redacted]) for ePAS2 (median OS: [redacted] (Min [redacted], Max [redacted])) and a median follow-up of [redacted] months (25th: [redacted], 75th [redacted]) for SAS3 (median OS: [redacted] (Min [redacted], Max [redacted])). At 1 year, the probability of survival was [redacted]% (95%: [redacted]) in the ePAS2 and was [redacted] in SAS3. Kaplan-Meier plots of OS are shown overleaf.

Figure 16. Kaplan-Meier Plot of OS (ePAS2, IRC assessment; 30 July 2018 data cut-off) (45)

[redacted]

Figure 17. Kaplan-Meier Plot of OS (SAS3, investigator assessment; 30 July 2018 data cut-off) (45)

Exploratory endpoints

Health-related quality of life

In summary, HRQoL analyses show that using larotrectinib to treat NTRK fusion-positive cancers resulted in early and sustained clinically meaningful improvement in quality of life—both in adult and paediatric patients – correlating with the clinical efficacy and disease response. However, because of the small sample size, most results were not statistically significant, and these results should be interpreted cautiously.

Table 16 presents the patient-reported outcome completion rates for the ePAS2. HRQoL was added as a part of a protocol amendment in 2015 after trial initiation, so PRO data is not available for all patients in the trials. Although some patients had missing assessments, most were due to administrative reasons; therefore, the data can be

considered as 'missing at random'. For paediatric questionnaires, some patients grew out of the infant category, and no data were available for those cycles. There were several instances where questionnaires were returned late or without any date. [redacted] of the [redacted] patients were from NAVIGATE which did not have the same scheduled visits as patients from the SCOUT study, hence the fluctuation of the number of patients with assessments at each cycle in various analyses.

Table 16. Summary of patients with Patient-Reported Outcome Data by study (ePAS2, 30 July 2018 data cut-off)(68)

NB. No data were collected in LOXO-TRK-14001.

Patients	NAVIGATE (n = 58)	SCOUT (n = 27)	Total (n = 93)
Adult patient under treatment at baseline	[redacted]	[redacted]	[redacted]
EORTC QLQ-C30/EQ-5D analysis population	[redacted]	[redacted]	[redacted]
Baseline	[redacted]	[redacted]	[redacted]
Baseline and at least 1 postbaseline	[redacted]	[redacted]	[redacted]
Pediatric patient under treatment at baseline (≥ 2 years)	[redacted]	[redacted]	[redacted]
PedsQL analysis population	[redacted]	[redacted]	[redacted]
Baseline	[redacted]	[redacted]	[redacted]
Baseline and at least 1 postbaseline	[redacted]	[redacted]	[redacted]
Pediatric patient under treatment at baseline (< 2 years)	[redacted]	[redacted]	[redacted]
PedsQL analysis population	[redacted]	[redacted]	[redacted]
Baseline	[redacted]	[redacted]	[redacted]
Baseline and at least 1 postbaseline	[redacted]	[redacted]	[redacted]

EORTC QLQ-C30=European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; EQ-5D=European Quality of Life 5-Five Dimensions 5-Levels Health Questionnaire; PedsQL=Pediatric Quality of Life Inventory;

The number of patients with MID-improvement (i.e. clinically meaningful improvement) for the EORTC QLQ-C30 (Global Health Score), EQ-5D 5L (VAS), and PedsQL (Total Score) are presented in Table 17. Improvements were rapid (by cycle [redacted] or [redacted]), seen across most tumour types, and sustained a minimum of [redacted] cycles.

Table 17. Number of patients with MID improvement (ePAS2, 30 July 2018 data cut-off)(68)

	Number of patients with baseline and 2 post-baseline measurements	At least one best post-baseline score > baseline score n (%)	MID Improvement [a] n (%)	Patients evaluable for sustained improvement	Sustained improvement n (%)
EORTC QLQ-C30 (Global Health Score)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EQ-5D 5L VAS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PedsQL (Total Score)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] MID: >10 points for EORTC QLQC-30 and EQ-5D 5L; > 4.5 points for PedsQL.
EORTC QLQ-C30=European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; EQ-5D=European Quality of Life 5-Dimensions 5-Levels Health Questionnaire; MID=minimally important difference; PedsQL-Core=Pediatrics Quality of Life–Core Module.

EORTC QLQ-C30

Adults

Within the EORTC QLQ-C30, scores for [redacted] and [redacted] had the largest improvements, whereas [redacted] had the lowest improvement during treatment. All functions and global health showed, on average, an improvement from baseline during treatment across patients (Table 18). Similarly, all symptoms showed, on average, an improvement from baseline during treatment (Table 19).

Table 18. Summary of baseline and best change from baseline in EORTC QLQ-C30 global health and functioning scores (Adults; ePAS2, 30 July 2018 data cut-off)(68)

	Baseline (N = [redacted])		Best Change from baseline (N = [redacted])	
Function[a]	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)
Global health	[redacted]	[redacted]	[redacted]	[redacted]
Physical	[redacted]	[redacted]	[redacted]	[redacted]
Role	[redacted]	[redacted]	[redacted]	[redacted]
Cognitive	[redacted]	[redacted]	[redacted]	[redacted]
Emotional	[redacted]	[redacted]	[redacted]	[redacted]
Social	[redacted]	[redacted]	[redacted]	[redacted]

IQR=interquartile range; QLQ-C30=Quality of Life Questionnaire–Core Module; SD=standard deviation.
[a] The scores are on a scale from 0 to 100, and higher scores indicate better function. A positive change from baseline indicates an improvement in the function.

Table 19. Summary of baseline and best change from baseline in EORTC QLQ-C30 symptoms scores (Adults; ePAS2, 30 July 2018 data cut-off)(68)

	Baseline (N = [redacted])		Best Change from baseline (N = [redacted])	
Symptom[a]	**Mean (SD)**	**Median (IQR)**	**Mean (SD)**	**Median (IQR)**
Fatigue	[redacted]	[redacted]	[redacted]	[redacted]
Pain	[redacted]	[redacted]	[redacted]	[redacted]
Nausea and vomiting	[redacted]	[redacted]	[redacted]	[redacted]
Dyspnoea	[redacted]	[redacted]	[redacted]	[redacted]
Loss of appetite	[redacted]	[redacted]	[redacted]	[redacted]
Insomnia	[redacted]	[redacted]	[redacted]	[redacted]
Constipation	[redacted]	[redacted]	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]	[redacted]	[redacted]
Financial impact of disease	[redacted]	[redacted]	[redacted]	[redacted]

IQR=interquartile range; QLQ-C30=Cancer Quality of Life Questionnaire–Core Module; SD=standard deviation.
[a] The scores are on a scale from 0 to 100, and higher scores indicate more symptoms. A negative change from baseline indicates an improvement of the symptom.

The estimated 25th percentile of time to improvement from baseline at the next visit was equal to [redacted] months (95% CI, [redacted]). [redacted] of the [redacted] patients at risk of event (i.e., patients with baseline and at least two post-baseline measures) had a sustained improvement in global health. For [redacted] of [redacted] patients, the sustained improvement lasted until the end of measurements, duration between [redacted] and [redacted] months.

EQ-5D-5L

Adults

Overall health status as assessed by the EQ-5D-5L results were consistent with the global health status as assessed by EORTC QLQ-C30.

All dimensions and VAS health status showed improvement from baseline during treatment based on best change from baseline for each patient (Table 20). VAS [redacted]t had the largest improvements, whereas [redacted] had the lowest improvements. Improvements generally correlated with disease response (Figure 18).

Table 20. Summary of baseline and best change from baseline in the EQ-5D-5L Visual Analogue Scale Health and Dimensions scores (Adults, ePAS2, 30 July 2018 data cut-off)(68)

Function	Baseline (N = [redacted])		Best Change from Baseline (N = [redacted])	
	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)
VAS health[a]	[redacted]			
Mobility[b]				
Self-care[b]				
Usual activities[b]				
Pain/discomfort[b]				
Anxiety/depression[b]				

IQR=interquartile range; SD=standard deviation; VAS=visual analogue scale.
[a] Higher VAS scores indicate better health, and lower dimension scores indicate better functioning. A positive change from baseline in VAS score indicates an improvement in health.
[b] A negative change from baseline in the function dimension scores represents an improvement in the function.

The estimated 25th percentile of time to improvement was equal to [redacted] months (95% CI, [redacted]). For [redacted] of [redacted] patients with sustained improvement, this lasted until the end of measurements, with sustained improvement duration between [redacted] and [redacted] months.

Figure 18. Waterfall plot of best absolute change from baseline in the EQ-5D-5L Visual Analogue Scale Health score (Adults, ePAS2, 30 July 2018 data cut-off)(68)

CR=complete response; IRC=independent review committee; PD=progressive disease.
Note:
.

See section B.3.4 and Appendix N for more detail.

PedsQL

Analysis of data from paediatric patients are consistent with the adult findings.

In the **paediatric population < 2 years of age**, the largest average of best change from baseline was for , and the lowest was for (Table 21). Because of the small sample size, no further analyses were performed on the group of paediatric patients younger than 2 years old.

Table 21. Summary of baseline and best change from baseline in Paediatric quality of Life Inventory scores (Paediatrics age <2 years; ePAS2, 30 July 2018 data cut-off)(68)

Function	Baseline (N = [redacted])		Best Change From Baseline (N = [redacted])	
	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)
Total score	[redacted]	[redacted]	[redacted]	[redacted]
Physical score	[redacted]	[redacted]	[redacted]	[redacted]
Physical functioning	[redacted]	[redacted]	[redacted]	[redacted]
Physical symptoms	[redacted]	[redacted]	[redacted]	[redacted]
Psychosocial score	[redacted]	[redacted]	[redacted]	[redacted]
Emotional functioning	[redacted]	[redacted]	[redacted]	[redacted]
Social functioning	[redacted]	[redacted]	[redacted]	[redacted]
Cognitive functioning	[redacted]	[redacted]	[redacted]	[redacted]

IQR=interquartile range; SD=standard deviation.
Note: The function scores are reverse transformed on scale from 0 to 100, and higher scores indicate a better function. A positive change from baseline indicates an improvement in the function.

In **paediatric patients aged ≥ 2 years**, all functions and scale scores from the PedsQL showed improvement during treatment (Table 22), with [redacted] showing the largest best improvements.

Table 22. Summary of baseline and best change from baseline in Paediatric quality of Life Inventory functioning and scale scores (Paediatrics age ≥2 years; ePAS2, 30 July 2018 data cut-off)(68)

Function [a]	Baseline (N = [redacted])		Best Change from Baseline (N = [redacted])	
	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)
Total score	[redacted]	[redacted]	[redacted]	[redacted]
Physical score [b]	[redacted]	[redacted]	[redacted]	[redacted]
Physical functioning	[redacted]	[redacted]	[redacted]	[redacted]
Psychosocial score [b]	[redacted]	[redacted]	[redacted]	[redacted]
Emotional functioning	[redacted]	[redacted]	[redacted]	[redacted]
Social functioning	[redacted]	[redacted]	[redacted]	[redacted]
School functioning [c]	[redacted]	[redacted]	[redacted]	[redacted]

IQR=interquartile range; SD=standard deviation.
Note: The function scores are reverse transformed on scale from 0 to 100, and higher scores indicate a better function. A positive change from baseline indicates an improvement in the function.
[a] Higher score means better functioning. [b] [redacted] patients had a missing score. [c] [redacted] patients had missing scores.

The estimated 25th percentile of time to improvement was equal to [redacted] months (95% CI, [redacted]). For [redacted] of [redacted] patients with sustained improvement, this lasted until the end of measurements, with sustained improvement duration between [redacted] and [redacted] months.

FACES

The [redacted] patients with FACES scores at baseline had a mean and standard deviation of [redacted] and [redacted], respectively, and the interquartile range equal to 0; this indicates that at least [redacted] of the [redacted] children ticked no pain. The average change from baseline was within [redacted] for all cycles, except for [redacted] cycles and end of treatment visit. Although the FACES instrument was analysed there was little variability in the scores.

***Post hoc* analysis - Intra-patient comparison(57)**

As outlined in section 2.4 a *post hoc* analysis using each patient as his / her own control was performed to compare the effects of larotrectinib versus the effect of the previous line of treatment the patient had received. A summary of the methodology / statistical approach used for this *post hoc* analysis is provided in Appendix Q.

The analysis assessed the growth modulation index (GMI) for patients from the ePAS2 dataset (n=93) who had received at least one prior systemic therapy in the metastatic setting (n=[redacted]). GMI was defined by Von Hoff (55, 56), and calculated as follows:

$$GMI_{Laro}=PFS_{Laro}/TTP_{-1}$$

where PFS_{Laro} is the time from the date of the first dose of larotrectinib and the earliest date of documented disease progression or death from any cause (based on IRC assessed data); and TTP_{-1} is the time from the start of therapy to the date of disease progression on that therapy for the most recent prior systemic anti-cancer therapy. A GMI > 1.33 was defined by Von Hoff as the sign of drug activity.

Using IRC assessed response data, [redacted]% of patients in the analysis had GMI ≥ 1.33. Results were consistent across key subgroups and sensitivity analyses.

Table 23. GMI for patients with metastatic disease who had received ≥ 1 prior systemic therapy (ePAS2; IRC; July 30 2018 data cut-off)(57)

		GMI	
		INV	IRC
All patients (n=[redacted])	Mean (SD)	[redacted]	[redacted]
	Median (Min, Max)	[redacted]	[redacted]
GMI category	< 1 [a], n (%)	[redacted]	[redacted]
	≥ 1, n (%)	[redacted]	[redacted]
	1 to <1.33, n (%)	[redacted]	[redacted]
	≥ 1.33, n (%)	[redacted]	[redacted]

GMI=growth modulation index; INV=investigator assessed; IRC=independent review committee assessed; PFS_{Laro}=Progression-free survival on larotrectinib; TTP_{-1}=time-to-progression on previous line of therapy;
[a] GMI <1 category: [redacted] patients censored to PFS (INV); [redacted] patients censored for PFS (IRC).

Figure 19. Waterfall plot of Growth Modulation Index for eligible patients from ePAS2 dataset (IRC; July 30 2018 data cut-off)

Sensitivity analyses of GMI endpoint

See Table 24.

Table 24. Results of sensitivity analyses for GMI(57)

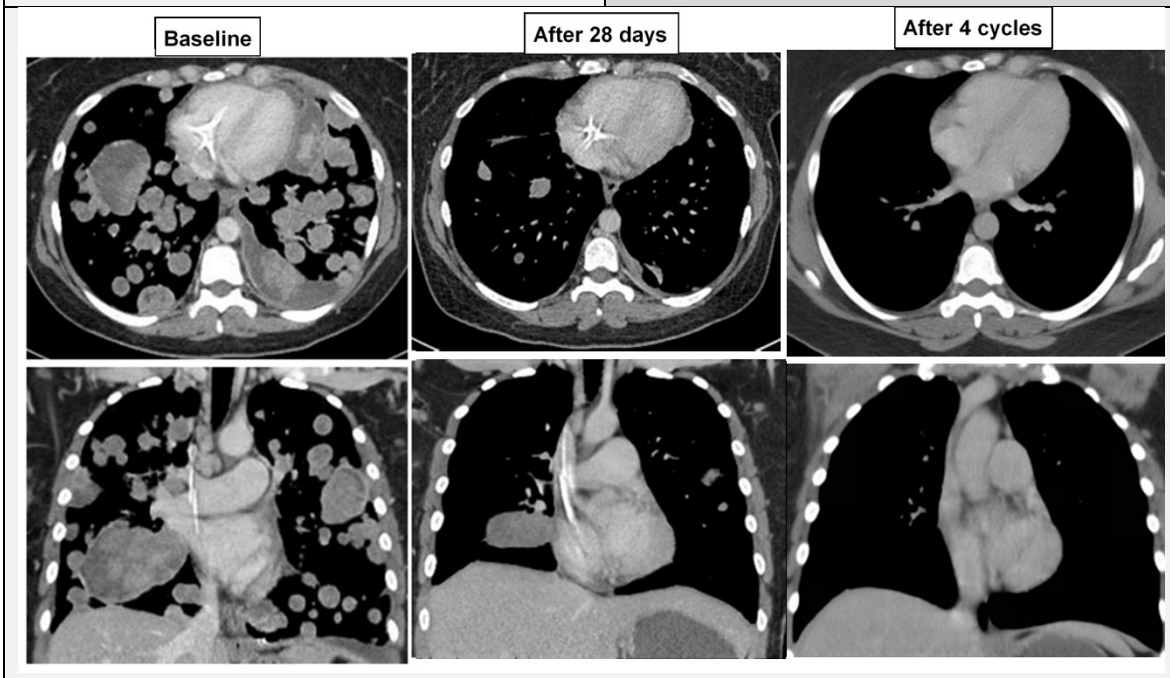
Sensitivity analysis	INV	IRC
Sensitivity analysis 1: TTP_{-1} calculated as the duration from prior therapy start date to PD or larotrectinib start date if date of PD is missing	**N=[redacted]**	**N=[redacted]**
< 1 n (%)	[redacted]	[redacted]
≥ 1 n (%)	[redacted]	[redacted]
1 to <1.33 n (%)	[redacted]	[redacted]
≥ 1.33 n (%)	[redacted]	[redacted]
Sensitivity analysis 1: includes all patients with at least 1 prior therapy regardless of disease setting	[redacted]	[redacted]
< 1 n (%)	[redacted]	[redacted]
≥ 1 n (%)	[redacted]	[redacted]
1 to <1.33 n (%)	[redacted]	[redacted]
≥ 1.33 n (%)	[redacted]	[redacted]

INV=investigator assessed; IRC=independent review committee assessed; PD=progressive disease;

Case studies

Figure 20. Case 1: Adult LMNA-NTRK1 Fusion Sarcoma (69)

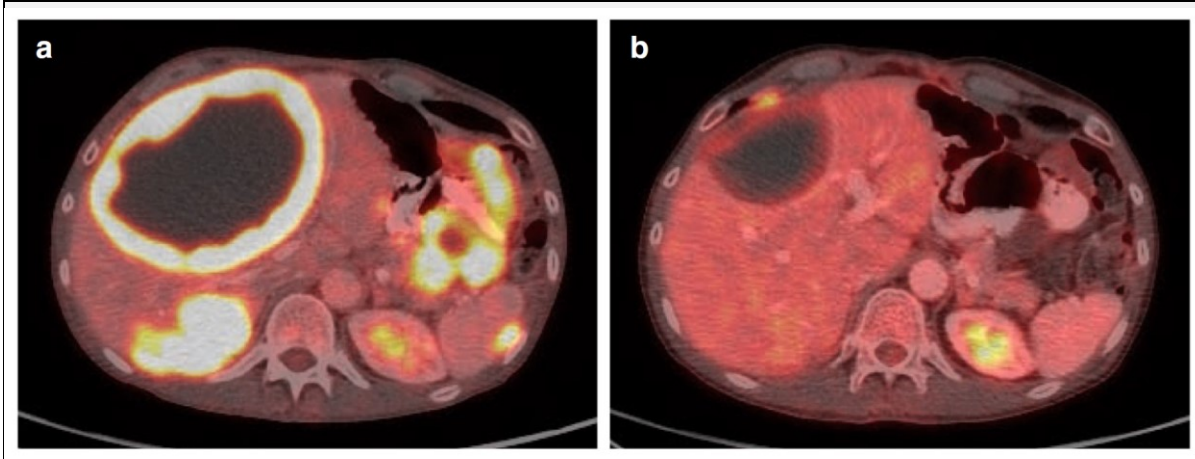
41-year-old female presented with a firm mass in her left groin (10cm) [undifferentiated soft tissue sarcoma]. Staging scans showed multiple bilateral 4-13mm pulmonary nodules consistent with metastatic disease. LMNA-NTRK1 fusion confirmed. After diagnosis, began an aggressive treatment plan of sorafenib with chemotherapy (epirubicin, ifosfamide with mesna), pre-operative radiation and limb-sparing surgery. The tumour became progressively more painful during the five weeks of systemic therapy. Then extension of the tumour was noted cranially within the psoas muscle, precluding the safe administration of effective radiation doses due to predicted bowel toxicity. Patient therefore came off protocol and proceeded to surgical resection. Resection of the primary tumour achieved negative margins and review of the pathologic specimen confirmed 90% tumour necrosis. A restaging chest CT obtained 9 weeks after initial scans showed worsening metastatic disease, with the largest nodule now measuring 18 mm. The patient's post-operative course was complicated by a polymicrobial wound infection requiring repeated wound debridement and prolonged antibiotic therapy. Repeat chest CT was obtained before resumption of chemotherapy and demonstrated dramatic progression over the prior 9 weeks, with multiple pulmonary nodules > 3 cm, the largest nearly 7 cm, and a large left pleural effusion. In February 2015, after placement of a pleural drain and initiation of supplemental home oxygen, the patient received doxorubicin 75 mg/m^2 once, while awaiting enrolment on the larotrectinib trial.	**Enrolled onto Adult phase I study.** **Baseline CT scan showed continued tumour progression with multiple large pulmonary metastases in both lungs, although the pleural effusion had resolved following placement of the pleural drain. On clinical presentation the patient had significant exertional dyspnoea and required 5litres/min of supplemental oxygen to maintain an oxygen saturation of 90%. Baseline laboratory values were notable for an elevated CA125 tumour marker level.** **Treated with larotrectinib 100mg b.d.** **Cycle 1: Patient was seen weekly and improvement in dyspnoea was noted, along with normalisation of CA125 levels. A CT scan was performed prior to the start of cycle 2 day 1, which demonstrated a marked improvement in multiple pulmonary metastases and was deemed a partial response by RECIST 1.1.** **Additional CT scans on cycle 5 day 1 (after 4 months of larotrectinib) demonstrated almost complete tumour disappearance of the largest tumours.** **Clinically, the patient had significantly improved exertional dyspnoea and was no longer requiring supplemental oxygen with an oxygen saturation of 97% on room air.** **As at July 30 2018 - Treatment ongoing, response >38.7 months.**

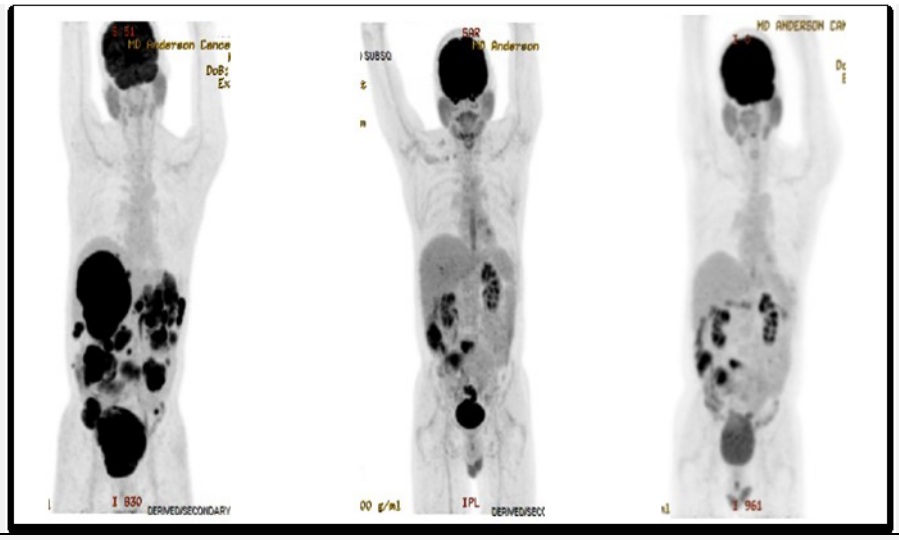


Inoperable metastatic STS is usually treated palliatively, with a median overall survival of approximately 1 year and a 5-year survival rate of less than 20% (70).

Figure 21. Case 2: Adult ETV6-NTRK3 GIST (71)

55-year-old male presented with T3N0M1 small intestine gastrointestinal stromal tumour (GIST). Confirmed ETV6-NTRK3 fusion. Originally diagnosed in May 2003. Five resection / debulking operations. Progressed on five lines of systemic therapy including: • Imatinib • Sunitinib • Sorafenib • Nilotinib • regorafenib	**Enrolled onto Adult phase I study.** **At the time of study entry, the patient had significant pain.** **Treated with larotrectinib starting dose 150mg b.d.** **Patient noted immediate improvement in his symptoms. Tumour response seen at end of week 8 by PET / CT (see image). Following 4 months of therapy, the patient had an ongoing PR (44%) according to RECIST 1.1 criteria.** **At July 30 2018 – Patient alive; Duration of response 26.3 months.**





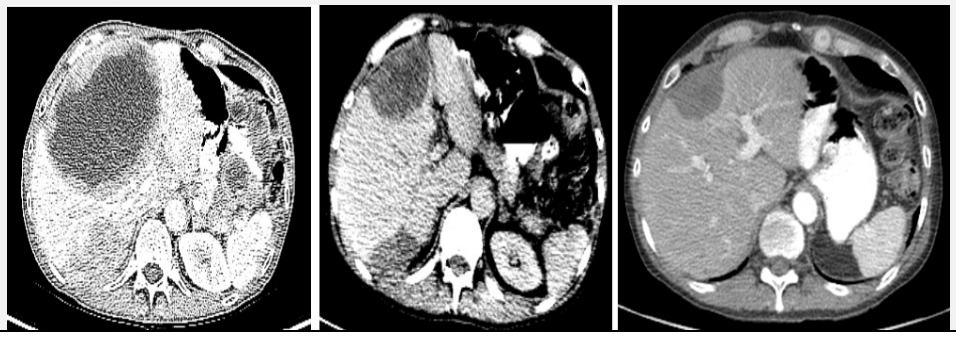
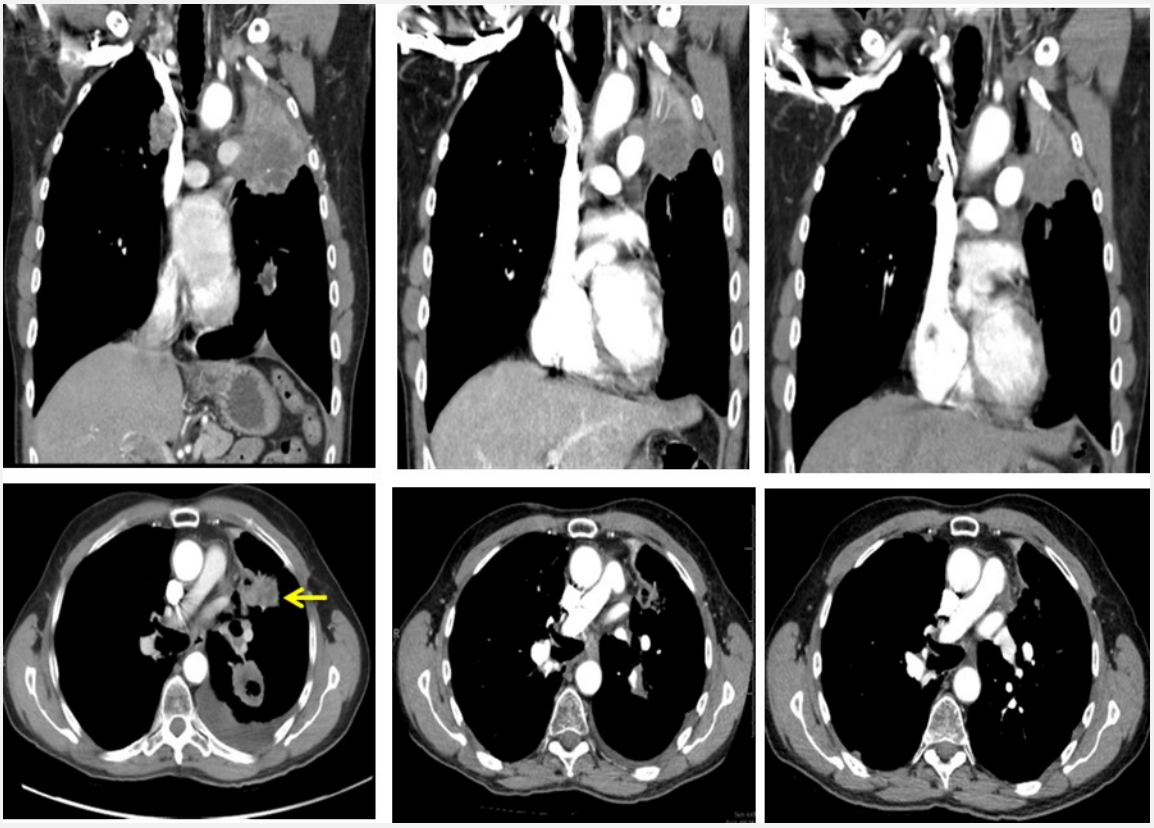


Figure 22. Case 3: Adult ETV6-NTRK3 Mammary Analogue Secretory Carcinoma of the Salivary Gland (MASC)(72)

66-year-old male initially diagnosed with MASC in 1998 (stage III disease). Confirmed ETV6-NTRK3 gene fusion. Underwent complete resection (with negative margins) in 1998 and 2001. Tumour became metastatic in 2006. Progressive disease despite several lines of treatment: • radiotherapy • dasatninb for ~ 9 months (achieved PR, discontinued due to toxicity) • radiotherapy • GDC-0941+erlotinib for ~29 months (stable disease initially, then disease progressed) • ABBV-399 for ~7 months (stable disease initially, then disease progressed) • radiotherapy	**Enrolled onto Adult phase I study.** **Treated with larotrectinib 100mg b.d.** **Confirmed partial response within 2 months of starting larotrectinib (see images below).** **As at July 30 2018 - Treatment ongoing, response >30.6 months**

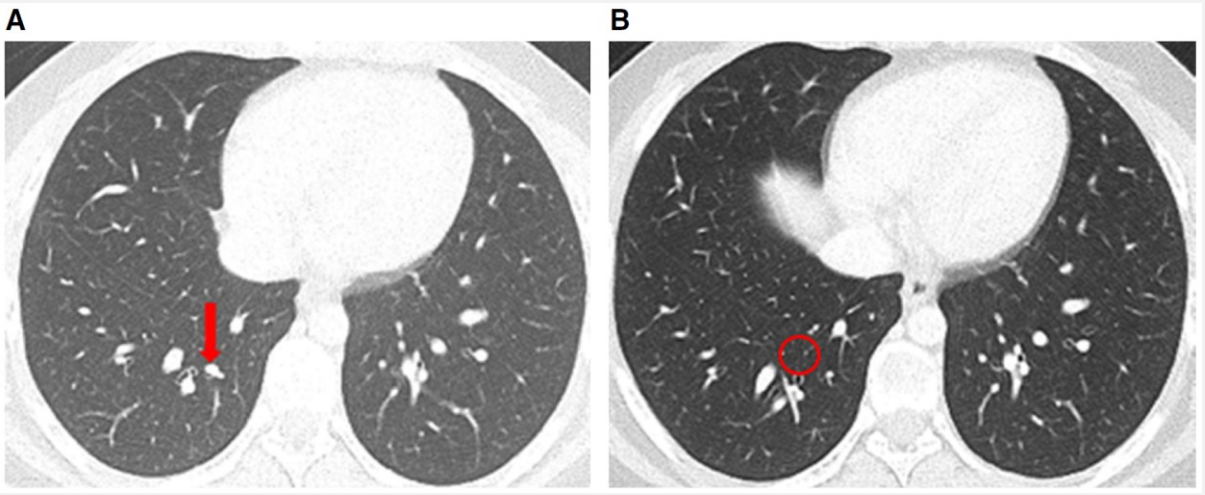


Generally, the clinical course of conventional MASC is characterised by moderate risk of local recurrences (15%) and lymph node metastases (20%) and low risk of distant metastases (5%) (62). Distant dissemination and tumour-related deaths are often preceded by local recurrence, the risk of which is higher after simple enucleation than after parotidectomy. High grade-transformed MASC is a much more aggressive neoplasm that follows an accelerated clinical course, resulting in local recurrences, cancer dissemination, and death.

Figure 23. Case 4: Paediatric TPM3-NTRK1 Papillary Thyroid Carcinoma (PTC) (73)

A previously healthy 12-yr-old boy presented with a large, painless right neck mass without constitutional symptoms. A total thyroidectomy confirmed PTC involving the right lobe and microscopic involvement of the isthmus and left lobe. Surgical margins were positive with multiple involved lymph nodes. Received 100 mCi of I131 Radioactive iodine (RAI) postoperatively and was maintained euthyroid with thyroid hormone replacement. Relapsed ~one year later – within the neck, thyroid bed and paratracheal lymph nodes, along with new bilateral pulmonary nodules. Despite cervical debulking surgery and additional 150 mCi of I131 RAI, follow-up I131 imaging demonstrated progression of multiple bilateral, small pulmonary nodules with active uptake in several neck and thyroid bed lymph nodes, confirmed by repeat neck and chest CT two months following repeat RAI revealed progression of innumerable pulmonary metastatic nodules. Analysis of sequencing data results from tumour sample revealed a somatic TPM3-NTRK1 fusion.	**In September 2016, patient enrolled onto paediatric phase I/II study (SCOUT) and treated with equivalent starting dose of 150mg b.d.** **After cycle 2, repeat CT chest demonstrated overall reduction in the number of pulmonary nodules (see images).** **Subsequent CT chest after four and six cycles revealed further interval improvement with excellent functional status (Lansky 100) without any therapy-related toxicity.** **At 10-months post initiation of therapy patient was reported to have excellent functional status with almost complete resolution of prior innumerable pulmonary nodules.** **As at July 30 2018 - Treatment is ongoing, stable disease >20 months**

(A) Axial CT chest image demonstrating a large (6-mm) pulmonary nodule (red arrow) prior to larotrectinib therapy initiation. At baseline, the patient had progressive innumerable, small pulmonary nodules demonstrated throughout the lungs bilaterally. (B) Axial CT image following four cycles of larotrectinib demonstrating resolution of the prior pulmonary nodule (red circle) documented at baseline.



Compared to adults, children with PTC are more likely to present with advanced local disease as well as widespread metastases. The treatment of locally invasive paediatric PTC is complete surgical excision followed by radioactive iodine (RAI) ablation (74). Although overall survival is excellent, the recurrence rate remains high at 40% and some children have multiple relapses, requiring multiple invasive surgeries and repeated RAI. High cumulative lifetime doses of RAI carries a significant risk of late effects, including secondary malignancies and pulmonary fibrosis (75).

Figure 24. Case 5: Paediatric ETV6-NTRK3 Secretory breast carcinoma (SBC) (76)

A 14-year-old girl originally presented in 2010 aged 8 years with a lump in the left breast. She underwent a lumpectomy with an initial diagnosis of fibroadenoma. One year later, she presented with a recurrent ipsilateral breast mass and underwent a second lumpectomy, her diagnosis revised to secretory breast carcinoma.

Treatment:

- 2 cycles of 5-FU, doxorubicin, and cyclophosphamide.
- Disease recurred locally (~ 1 year later). Simple mastectomy with axillary lymph node dissection, followed by 4 cycles of carboplatin and docetaxel.
- Chest wall recurrence (in 2014). Local resection followed by 2 cycles of vinorelbine and gemcitabine.
- One year later – a second left chest wall recurrence with bilateral lung metastases. Reresection of the chest wall mass and 2 cycles of ifosfamide, doxorubicin, dacarbazine, and mesna.
- < one year later, patient presented with a recurrent fungating mass in the left chest wall. Treated with 2 cycles of carboplatin and paclitaxel, with no clinical benefit.

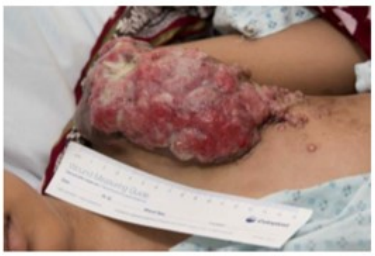
Having exhausted all treatment options and after the tumour was confirmed as having ETV6-NTRK3 gene fusion, patient was treated under a single patient use protocol (compassionate use).

Prior to larotrectinib initiation, the patient reported significant pain at the recurrent chest wall tumour site. Physical examination revealed a large fungating chest mass with multiple satellite lesions scattered over her chest wall (see images) and CT scan revealed numerous pulmonary metastases as well as bone metastases involving the sternum and vertebrae.

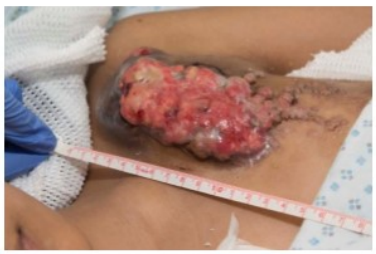
Treated with larotrectinib 100mg b.d.

Cycle 1 / cycle 2: Marked improvement of tumour-related pain within 3 days of starting therapy. Significant and rapid reduction in size of left chest mass within 1 week of therapy, with near-complete resolution after 2 months therapy. CT scan showed near-complete resolution of the pulmonary metastases.

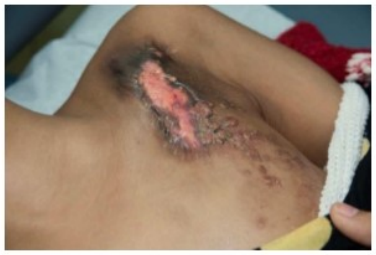
Response ongoing for approximately 4 months (to date of manuscript writing).



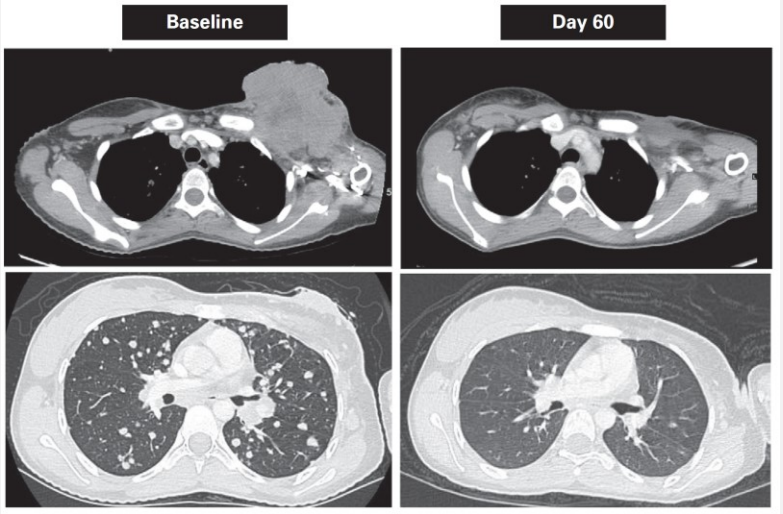
Baseline



Day 6



Day 20



B.2.7 Subgroup analysis

Supportive analyses (point estimate of the ORR [and 95% CI], based on IRC assessment was to be performed to assess ORR for consistency across selected subgroups (15):

- age at enrolment (1 month to < 2 years, 2 to < 6 years, 6 to < 12 years, 12 to < 18 years, 18 to < 65 years, ≥ 65 years)
- paediatrics (age at enrolment < 18 years)
- adults (age at enrolment ≥ 18 years)
- sex (male, female)
- race (White, Black, Asian, other)
- ECOG performance status at baseline (0–1, 2, 3)
- NTRK fusion (NTRK1, NTRK2, NTRK3)
- NTRK fusion partner (e.g. ETV6-NTRK3)
- any other known oncogenic alterations present (yes, no, unknown)
- primary cancer diagnosis (according to standardised term)
- cancers considered pathognomonic for NTRK fusions (IFS, MASC)
- number of metastatic sites of disease (0, 1–2, ≥ 3)
- number of prior systemic regimens or treatment courses (0, 1–2, ≥ 3)
- best overall response to most recent prior systemic regimen or treatment course (CR, PR, stable disease, progressive disease, unknown or unevaluable or not applicable)
- starting dose of larotrectinib and frequency of administration

A summary of analyses by key subgroups e.g. age, tumour diagnosis is presented in Appendix E.

B.2.8 Meta-analysis

Meta-analysis was not applicable for this submission – pooled data from the three trials has been presented.

B.2.9 Indirect and mixed treatment comparisons

Comparator identification

See appendix D for full details of the process and methods used to identify comparator clinical evidence.

Indirect comparisons were not conducted for this appraisal. The approach taken for selection of comparators and identification of evidence is described in full in section B.3.2 and Appendix D.

Due to there being no existing treatments for (and an absence of published data on) patients with NTRK fusion-positive cancer, evidence presented for the comparator arm considers a population where patients are not treated on the basis of NTRK gene fusion status, but in line with existing standard of care according to tumour location and the line of treatment.

To generate relevant comparator clinical evidence for this appraisal, systematic literature reviews (SLRs) were undertaken in tumour sites / locations known to harbour neurotrophic tyrosine receptor kinase (NTRK) gene fusions, reflective of those of patients so far investigated within larotrectinib clinical studies. The systematic reviews for clinical evidence were part of a broader review also including available economic, and patient-reported outcome (PRO)/health-related quality of life (HRQoL) evidence.

Selection of comparator technologies for each tumour site for the systematic literature review and economic modelling was based on:

- Inclusion / exclusion criteria associated with larotrectinib clinical development programme
- Previous NICE HTA guidance

- ESMO guideline recommendations
- EMA-approved treatments
- Targeted literature search

and where appropriate, supported by the opinions of clinical experts. Due to the heterogeneity of tumour types included in the larotrectinib clinical development program, the comparators for larotrectinib vary based on tumour type. Over 75 comparators were considered.

Results

The clinical efficacy and safety of currently available interventions used for the treatment of tumour types with NTRK gene fusions are variable. Across the included studies, median OS ranged from 2.3 months in pancreatic cancer, to not reached for thyroid carcinoma, GIST, and certain soft tissue sarcomas. Median PFS was generally less than 12 months across included tumour types, while overall response rates ranged from 0% to more than 95%. Pronounced variability in the percentage of patients experiencing serious adverse events (SAEs) was evident, ranging from less than 10%, to 100% in the included trials. Treatment-related SAEs were reported in patients with all evaluated tumour types. For several tumour types (including secretory breast cancer, salivary gland cancer, myopericytoma, and spindle cell sarcoma), there are no standard treatments used with locally advanced/metastatic disease as illustrated by the paucity of evidence-based literature identified. For other tumor types, chemotherapeutic interventions for locally advanced/metastatic disease provide little benefit as shown by the overall mortality data (GIST), lack of CR (NSCLC), and substantial toxicity (CRC). Summaries of the findings from the clinical SLR are presented in the tables below.

Table 25. Overview of Treatment Efficacy in Tumour Histologies where NTRK gene fusions occur [a] (Adapted from (77))

Tumour Type	Treatment Line	Number of patients N	ORR (%)	Median PFS (Months)	Median OS (Months)
NSCLC	Second	18–628	2.7–28.9	2.5–5.8	4.7–15.2
	Second or further	49–613	4.2–25.5	2.3–10.3	4.6–not reached
CRC	First or further	205	NR	NR	16.2–16.4
	Second	8–614	11–47.7	0.3–10.5	4–17
	Second or further	24–505	0–28	1.4–7.3	5–14.3
	Third	91–124	8–9	12.9–13.2	NR
	Third or further	57–534	0–13	1–4.8	5.3–11.4
	NR	33–534	43.6–67.5	3.8–8.6	5.2–9.9
Melanoma	First or further	47–555	4–39.9	2.7–6.6	11.5–not reached
	Second	NR	12–29	NR	NR
	Second or further	72–272	0–32	3.1–35 weeks	8.6–16.4
	NR	179–361	4–28	NR	11–14.7
Pancreatic cancer	Second	23–24	NR	3.9–4	2.3–9.1
	NR	10–11	NR	NR	5.2–7.2
Less frequent tumour types					
Thyroid cancer (anaplastic, follicular,	First	12–75	21.4–61.1	1.6–11	6–56

Tumour Type	Treatment Line	Number of patients N	ORR (%)	Median PFS (Months)	Median OS (Months)
or papillary tumor histology)	First or further	10–80	0–69	1.7–7.4	4.0–12
	Second or further	20–26	0	1.9	3.9–12.3
	NR	19–417	0.5–64.8	2.1–18.3	3.5–not reached
Gliomas	First or further	55–119	NR	NR	5.2–11.3
	Second	32–40	63	3.8–5.8	6.9
	Second or further	14–40	95.2	3–28	7–28.3
	Third or further	9–31	NR	2.9–12.4	12
	NR	20–61	25–53	NR	11–13.8
Biliary	First	41–206	50–81.4	3.7–8	7.7–11.7
STS[b]	First	12–48	17.2–44.4	NR	NR
	First or further	6–175	13.2-86	2.4–15.4	11–46.9
	Second or further	5	NR	6.5	8.9
	NR	7–103	0–66.7	1.92–not reached	9–not reached
GIST	First	19–22	NR	NR	49–not reached
	First or further	141–473	45	18–27.2	46.8–not reached
	Second	41–312	0–7	1.5–30	33 – 37
	Second or further	118–243	0–10	6–22.9 weeks	39–72.7 weeks

Tumour Type	Treatment Line	Number of patients N	ORR (%)	Median PFS (Months)	Median OS (Months)
	Third or further	17–188	1.5–53	0.9–7.4	7.5–not reached
	NR	25–835	29.9–57.1	0.8–20	9.7–not reached
Bone sarcoma[c]	First	4–180	NR	4.7	5.8–18
	First or further	4–340	10–25.6	3.5–9.3 years	7–20
	Second	116	8	NR	NR
	NR	3–73	0–33	2–12.5	3–87
Rare tumour types					
Salivary gland	First	42	31	6	10
	First or further	57	70.2	8.9	39.7
	Second	18	5	3.5	4
	NR	5–42	0–100	5–7	8.5–18
Appendix	First	54–109	44–56	6.9	11.7–not reached
	First or further	11–54	NR	7.6	2.5 years
	Second	45	NR	2.8	NR
	NR	5–567	39–85	4–44.4	16–not reached
IFS/IM	First	6–20	71–83	NR	NR
	NR	8–9	88.9–93	NR	NR

AE=adverse event; BSC=best supportive care; CMN=congenital mesoblastic nephroma; CR=complete response; CRC=colorectal cancer; EFS=event-free survival; GIST=gastrointestinal stromal tumour; IFS=Infantile fibrosarcoma; IM=infantile myofibromatosis; N=number; NR=not reported; NSCLC=non-small cell lung cancer; NTRK=neurotrophic tyrosine receptor kinase; ORR=overall response rate; OS=overall survival; PFS=progression-free survival; RCT=randomised controlled trial; SAE=serious adverse event; STS=soft tissue sarcoma.

[a] ORR, median PFS, or median OS efficacy data were not available for CMN and secretory breast carcinoma.

[b] STS includes inflammatory myofibroblastic tumor, myopericytoma, spindle cell sarcoma, and peripheral nerve sheath tumour.

[c] Bone sarcomas includes data for chondrosarcomas only.

Table 26. Overview of Treatment Safety in Tumour Histologies where NTRK gene fusions occur [a] (77)

Tumour Type	Treatment Line	Number of patients (N)	Patients with Any SAE (%)	Patients with Treatment-Related SAEs (%)
Common tumour types				
NSCLC	Second	17–628	23.5–52	4.7–24
	Second or further	107–613	<5–67	1–54
CRC	First or further	203	NR	79–94
	Second	43–614	7–83.5	62–79
	Second or further	90–505	36–79	14–54
	Third or further	57–534	9–69	15–54
	NR	33–83	32.5–42.4	NR
Melanoma	First or further	47–555	1	10.1–37
	Second or further	71–272	36.1–53.5	1–34
	NR	179–181	NR	13–26
Less frequent tumour types				
Thyroid	First	26–56	NR	7.7–28.6
	First or further	11–80	0–29	13–37.3
	Second or further	26	NR	4–35
	NR	18–417	22.9–100	6.1–30.3
Biliary	First	12–206	68.8–75	NR
STS[b]	First	14–34	NR	26.5–28.6
	First or further	14-29	71	5–57
	NR	11–12	8.3–9	NR
GIST	First	70	NR	70
	First or further	147–473	37–38	21–63
	Second	312	NR	5–20
	Third or further	17–199	18–49	0–83
	NR	20–50	16–40	NR
Bone sarcoma[c]	NR	3–73	11–100	NR
Rare tumour types				
Salivary gland	First	42	NR	4–24
	First and further	57	89.5	4–60
	Second	18	NR	11–33
	NR	16-42	NR	6–16
Appendix	NR	17–155	12–39	1.3–22

BSC=best supportive care; CMN=congenital mesoblastic nephroma; CRC=colorectal cancer; GIST=gastrointestinal stromal tumour; IFS=Infantile fibrosarcoma; IM=infantile myofibromatosis; N=number; NR=not reported; NSCLC=non-small cell lung cancer; NTRK=neurotrophic tyrosine receptor kinase; SAE=serious adverse event; STS=soft tissue sarcoma.

[a] SAE and treatment-related SAE safety data were not reported for the following tumor types: pancreatic, gliomas, IFS/IM, CMN, and secretory breast carcinoma.

[b] STS includes inflammatory myofibroblastic tumor, myopericytoma, spindle cell sarcoma, and peripheral nerve sheath tumor.

[c] Bone sarcomas include data for chondrosarcomas only.

B.2.10 Adverse reactions

Summary of larotrectinib safety and tolerability

The safety profile of larotrectinib is characterised by adverse events (AEs) that can be monitored and are manageable and reversible, and was comparable across adult and paediatric patients, and tumour types.

Larotrectinib was well tolerated in patients with NTRK fusion-positive cancer from a pooled safety analysis across three clinical studies. Most drug-related AEs ([redacted]%) were grade 1 or 2, and included [redacted]. AEs leading to dose modification occurred most commonly in the first [redacted] months of treatment. Permanent discontinuation due to an AE considered to be treatment-related occurred in [redacted] Long-term follow-up of patients with > 2 years exposure (n=[redacted], as of 30 July 2018) has not indicated new or cumulative toxicities.

Introduction to adverse event data

The safety profile of larotrectinib for the treatment of adult and paediatric patients who had locally advanced or metastatic NTRK fusion-positive solid tumours is based on the analysis of adverse events (AEs) that occurred in 3 clinical studies (Studies LOXO-TRK-14001, LOXO-TRK-15002 [NAVIGATE], and LOXO-TRK-15003 [SCOUT]).

As described in Section B.2.2 (Table 4), LOXO-TRK-14001 is a phase 1 clinical trial that included a cohort of adult patients with advanced solid tumours enrolled to receive larotrectinib at 50 to 200 mg/day (once daily [q.d.] or twice daily [b.d.]) in the dose-escalation phase, then 100mg b.d. in the expansion phase of the study; LOXO-TRK-15002 is a phase 2 clinical trial that included a cohort of patients aged ≥ 12 years with an advanced cancer bearing an NTRK gene fusion enrolled to receive larotrectinib at 100 mg b.d.; and LOXO-TRK-15003 is a phase 1/2 clinical trial that included a cohort of

paediatric patients with advanced solid or primary central nervous system (CNS) tumours enrolled to receive larotrectinib at dosing based on the adult equivalent of 100 or 150 mg b.d. then 100 mg/m^2 b.d. (with a maximum of 100 mg b.d.; actual doses administered ranged from [redacted] to 1[redacted] mg/m^2 b.d.).

The population for safety analysis within this submission is a pooled analysis, comprising 'all patients with NTRK fusion-positive cancer from LOXO-TRK-14001, NAVIGATE and SCOUT studies, who have received ≥ 1 dose of larotrectinib, as of 30 July 2018, regardless of whether evaluable for efficacy' (n=137; 82 patients from NAVIGATE, 10 from LOXO-TRK-14001 and 45 patients from SCOUT). This population aligns with the decision problem and the safety inputs within the economic model. Baseline characteristics were similarly distributed to those of the population evaluable for efficacy.

Median time on treatment for patients in the safety analysis population is [redacted] months (Min [redacted], Max [redacted]) and mean time on treatment is [redacted] months.

Summary of adverse events

Adverse events were classified using MedDRA (Medical Dictionary for Regulatory Activities) Version 18.1.

The majority of patients (n=[redacted]; [redacted]%) experienced at least 1 treatment-emergent AE (TEAE) considered to be related to larotrectinib. In most cases, drug-related TEAEs were grade 1 or grade 2. [redacted] patients ([redacted]%) experienced a grade 3 or 4 TEAE considered to be drug-related. [redacted] patients ([redacted]%) experienced a TEAE that led to drug discontinuation; however, a larotrectinib-related cause was determined in only [redacted] of these patients ([redacted]%). [redacted] patients ([redacted]%) missed, skipped or delayed a dose of larotrectinib due to an AE. Adverse events that led to a dose reduction occurred in [redacted] patients, and usually occurred in the first [redacted] months of treatment.

A summary of TEAEs of any cause that occurred in at least 10% of patients is presented in Table 27. The most frequent TEAEs included [redacted].

Table 27. TEAEs (all-cause and drug-related) occurring in ≥10% of patients with NTRK fusion-positive cancer in the pooled safety analysis of larotrectinib clinical trials (safety analysis set) (46)

Primary system organ class Preferred term	Larotrectinib, NTRK cancer safety analysis set (N=137)	
	All causality N (%)	**Larotrectinib-related** N (%)
Number of patients with at least one TEAE	[redacted]	[redacted]
Blood and lymphatic system disorders		
Anaemia	[redacted]	[redacted]
Neutrophil count decreased	[redacted]	[redacted]
Leucocyte count decreased	[redacted]	[redacted]
Lymphocyte count decreased	[redacted]	[redacted]
Gastrointestinal disorder		
Constipation	[redacted]	[redacted]
Nausea	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]
Vomiting	[redacted]	[redacted]
Abdominal pain	[redacted]	[redacted]
General disorders and administration site conditions		
Fatigue	[redacted]	[redacted]
Pyrexia	[redacted]	[redacted]
Oedema peripheral	[redacted]	[redacted]
Infections and infestations		
Upper respiratory tract infection	[redacted]	[redacted]
Investigations		
Alanine Aminotransferase increased	[redacted]	[redacted]
Aspartate Aminotransferase increased	[redacted]	[redacted]
Weight increased	[redacted]	[redacted]
Blood creatinine increased	[redacted]	[redacted]
Musculoskeletal and connective tissue disorders		
Myalgia	[redacted]	[redacted]
Pain in Extremity	[redacted]	[redacted]
Arthralgia	[redacted]	[redacted]
Back pain	[redacted]	[redacted]
Muscular weakness	[redacted]	[redacted]
Nervous system disorders		
Dizziness	[redacted]	[redacted]
Headache	[redacted]	[redacted]
Renal and urinary disorders		
Urinary tract infection	[redacted]	[redacted]
Respiratory, thoracic and mediastinal disorders		
Cough	[redacted]	[redacted]

Primary system organ class Preferred term	Larotrectinib, NTRK cancer safety analysis set (N=137)	
	All causality N (%)	**Larotrectinib-related** N (%)
Nasal congestion	[redacted]	[redacted]
Dyspnoea	[redacted]	[redacted]

AE=adverse event; ALT=alanine aminotransferase; AST=aspartate aminotransferase; NTRK=neurotrophic tyrosine kinase; TEAE=treatment-emergent adverse event;

Drug-related TEAEs

The most common drug-related TEAEs (≥ 10% patients) were [redacted].

Treatment-emergent serious adverse events (TESAEs)

Grade 3 or 4 TEAEs that occurred in at least 2% of patients are presented in Table 28. The most frequent grade 3 or 4 TEAEs were [redacted]. Grade 4 TEAEs considered to be related to larotrectinib occurred in [redacted] patients ([redacted]).

Table 28. Grade 3 or 4 TEAEs (all and drug-related) occurring in ≥2% of patients with NTRK fusion-positive cancer in the pooled analysis of larotrectinib clinical trials (safety analysis set) (46)

	Larotrectinib, NTRK cancer safety analysis set (N=137)	
Primary system organ class Preferred term	**All causality** N (%)	**Larotrectinib-related** N (%)
Number of patients with at least one grade 3 or 4 TEAE	[redacted]	[redacted]
Anaemia	[redacted]	[redacted]
Neutrophil count decreased	[redacted]	[redacted]
Hypophosphataemia	[redacted]	[redacted]
Alanine Aminotransferase increased	[redacted]	[redacted]
Hypokalaemia	[redacted]	[redacted]
Lymphocyte count decreased	[redacted]	[redacted]
Sepsis	[redacted]	[redacted]
Weight increased	[redacted]	[redacted]
Hyponatraemia	[redacted]	[redacted]
Aspartate Aminotransferase increased	[redacted]	[redacted]
Cellulitis	[redacted]	[redacted]
Hypocalcaemia	[redacted]	[redacted]

AE=adverse event; ALT=alanine aminotransferase; AST=aspartate aminotransferase; NTRK=neurotrophic tyrosine kinase; TEAE=treatment-emergent adverse event;

Adverse events of special interest

Adverse events which could be specifically related to the mechanism of action of larotrectinib or are known to occur in drug treatments from a similar class of action are of particular interest. For larotrectinib, these included any neurologic adverse events, liver function abnormalities, and neutropenia and leucopenia.

Neurologic AEs: Dizziness n=[redacted] paraesthesia n=[redacted] and peripheral sensory neuropathy n=[redacted] Neurologic reactions leading to dose modification included dizziness (2%) and gait disturbance (<1%). None of these led to treatment discontinuation. In all cases, patients with evidence of anti-tumour activity who required a dose reduction were able to continue dosing at a reduced dose and/or schedule(78).

ALT and or AST were reported as increased in approximately [redacted] patients (ALT: n=[redacted]%; AST: n=[redacted]% mostly of grade 1 severity and occurring within the first [redacted] cycles of treatment. Increases in ALT and AST leading to dose modifications occurred

in [redacted] ([redacted]%) patients and [redacted] ([redacted]%) patients, respectively. During [redacted], [redacted] ([redacted]%) patients permanently discontinued the treatment due to Grade 3-4 ALT and AST increases.

Neutropenia and / or leucopenia occurred in approximately [redacted]% patients, generally grade 1 or 2, and mainly within the first [redacted] cycles of treatment.

Adverse events leading to premature permanent discontinuation of study drug

[redacted] patients permanently discontinued treatment due to TEAEs, in [redacted] considered to be due to larotrectinib (increased ALT and AST), and [redacted] patients whose TEAEs were not considered drug-related (disease progression n=[redacted]; intestinal perforation n=[redacted] and jaundice n=[redacted], small intestinal obstruction n=[redacted]).

Deaths

There were [redacted] deaths in the pooled safety analysis of patients with NTRK fusion-positive solid tumours – [redacted] as a result of disease progression, [redacted] due to intestinal perforation and [redacted] due to small intestinal obstruction, none of which were associated with larotrectinib treatment (46).

Subgroup analyses(78)

The safety profile in the paediatric population (< 18 years) was consistent in types of reported AEs to those observed in the adult population. The majority of adverse reactions were Grade 1 or 2 in severity and were resolved without larotrectinib dose modification or discontinuation. The adverse reactions of vomiting ([redacted]% versus [redacted]% in adults), leucocyte count decrease ([redacted]% versus [redacted]% in adults), neutrophil count decrease ([redacted]% versus [redacted]% in adults), and transaminase elevations (ALT [redacted]% versus [redacted]% in adults and AST [redacted]% versus [redacted]% in adults) were more frequent in paediatric patients compared to adults. The majority of these AEs were reported associated with concurrent viral infection and bone marrow injury from prior systemic chemotherapy. Elevations in liver enzymes in children < 1 year of age may be due to immaturity of liver function.

Long-term treatment

[redacted] patients (as of 30 July 2018 data cut-off) have so far had [redacted] exposure with larotrectinib treatment. Long-term follow-up of patients has not indicated new or cumulative toxicities.

Overview of the safety of the technology in relation to the decision problem

It is anticipated that the safety profile of larotrectinib within the population defined in the decision problem in routine clinical practice in the UK will be similar to that of the pooled safety analysis described above. Patients included in the pooled safety analysis had a locally advanced or metastatic NTRK fusion-positive solid tumour which was refractory to standard therapy or for which there were no satisfactory treatment options, or where surgical resection was likely to result in severe morbidity. This was a well-defined population which matches that of the decision problem population. There are no data available on the characteristics of the NTRK fusion-positive population within clinical practice in England, however, the age range of the larotrectinib safety population was sufficiently wide ([redacted] years) to suggest it has broad applicability with regard to paediatric and adult patients. Additionally, demographic baseline characteristics of the larotrectinib safety population with NTRK fusion-positive tumours generally reflect the UK cancer population, including 'Race' with [redacted]% of the Larotrectinib NTRK group being 'White' and 73% of UK cancer patients also described as 'White' (National Cancer Intelligence Network and Cancer Research 2009 (79)).

B.2.11 Ongoing studies

The 3 single-arm clinical studies that comprise the larotrectinib clinical development programme, reported in this submission, are ongoing: LOXO-TRK-14001 (Phase I in adult patients), LOXO-TRK-15002 (Phase II in adult and adolescent patients, NAVIGATE), and LOXO-TRK-15003 (Phase I/ II in paediatric patients, SCOUT).

Table 29. Ongoing studies

Study	Final CSR anticipated
LOXO-TRK-14001	[redacted]
LOXO-TRK-15002	[redacted]
LOXO-TRK-15003	[redacted]

* based on additional patients in common cancer types to be enrolled with 12 months follow up

As well as the final CSRs, updates (including survival analysis) are currently planned [redacted].

For all studies (14001, 15002, 15003), long-term follow up (LTFU) assessments will occur every 3 months. LTFU may be conducted by telephone.

As part of the FDA commitments, Bayer will also submit the final report, including datasets, from the first 55 patients (primary analysis set) with NTRK-fusion solid tumours enrolled across Study LOXO-TRK-14001 (NCT02122913), SCOUT (NCT02637687), and NAVIGATE (NCT02576431), to further characterise the duration of response in patients who achieved a complete or partial response to larotrectinib. All responding patients will be followed for at least 2 years from the onset of response and duration of response will be assessed by independent central review.

Further to this, a non-interventional study has been proposed and the protocol is currently being finalised and approved with the FDA. Post authorisation measures are also being discussed with EMA.

Lastly, there are two RWE studies that are ongoing:

[redacted]

[redacted]

B.2.12 Innovation

In line with the changing approach to cancer treatment, larotrectinib is considered innovative and a 'step change' in the management of NTRK fusion-positive cancer.

1. **Larotrectinib provides a specific treatment for NTRK fusion-positive solid tumours where previously no treatment was available** - Prior to the introduction of larotrectinib, patients with solid tumours harbouring an NTRK gene fusion were not considered as a separate treatable population. Instead, patients with solid tumours, irrespective of NTRK status, received treatment per treatment guideline recommendations for the specific tumour histology. Larotrectinib represents a paradigm shift in the way cancer is treated, enabling cancer treatment to be delivered according to causation (in this case, the presence of NTRK gene fusion) as opposed to tumour location e.g. lung, prostate, thyroid, as has been done traditionally. If approved by the EMA, **larotrectinib will be the first histology independent therapy approved in Europe.**

2. **Innovative design to selectively target NTRK fusion cancer: a precision medicine** – Larotrectinib is a first-in-class, orally bioavailable, potent and highly selective inhibitor of TRKA, TRKB, and TRKC, rationally designed to avoid activity with off-target kinases using crystallography-informed structure-based design. It was selected for clinical development, in part, for its unusually high selectivity; it has low nanomolar potency against all 3 TRK family members in enzyme and cellular assays, and 100- to 1,000-fold selectivity relative to other kinases.

   There are several advantages to selective kinase inhibitors compared with non-selective kinase inhibitors. Kinase inhibitors that are designed to selectively block the kinase involved in aberrant tumour signaling have shown increased potency compared with non-selective kinase inhibitors (80). In addition, less off-target effects can minimise side effects associated with therapy (80). This marks a departure from non-targeted chemotherapy, generally associated with significant toxicities, and which may not represent an adequate standard therapy for certain

patient populations (e.g. very young paediatric patients or adult patients who may be elderly or frail).

Use of targeted therapies has been shown to provide maximum benefit and have the potential to improve patient quality of life (QoL) (27). It is also expected to reduce the overall cost for the healthcare system, as patients may ultimately avoid treatment unlikely to benefit them and potentially cause harm (27).

3. **Treatment of adults and children within one indication -** Larotrectinib has been shown to be generally safe and effective across a broad range of tumours including rare tumours and rare subsets of more common tumours, and in paediatric and adult patients ranging in age from [redacted] years.

4. **A step towards delivering 'Personalised medicine' in cancer patients** - Personalised medicine is based on comprehensive genomic and diagnostic characterisation, meaning different subtypes of patients within a given condition can be identified, and treatment can be tailored to the underlying cause. The availability of larotrectinib enables delivery of personalised medicine to cancer patients harbouring NTRK gene fusions.

5. **In a rare disease (~1% of solid tumours), larotrectinib provides improved patient outcomes where previous therapies have failed or no standard therapy is available** – The histology-independent mechanism of action of larotrectinib has been clearly demonstrated in terms of the compelling efficacy in multiple tumour types with NTRK gene fusions. The population considered in this submission generally have had poor response to prior treatment with non-targeted therapies and/or have no satisfactory treatment options. Treatment with larotrectinib exhibited rapid (median time to response: [redacted] for patients with non-CNS solid tumours; [redacted] in primary CNS tumours), substantial antitumour activity (ORR: [redacted]% (n=[redacted]/93) in non-CNS solid tumours; [redacted]% in primary CNS tumours) with durable disease control that appears to be independent of NTRK fusion partner, tumour type and patient age. A post hoc analysis comparing

patient response to larotrectinib to that of their most recent prior systemic anticancer therapy showed that [redacted]% patients had GMI ≥ 1.33, which even at a conservative level, is considered a sign of drug activity. Results were consistent across key subgroups (see Appendix E – subgroup analyses) and sensitivity analyses.

In recognition of the significance of larotrectinib the US Food and Drug Administration (FDA) granted larotrectinib Priority Review and Breakthrough Therapy designation. Larotrectinib also received Orphan Drug designation, an incentive to assist and encourage the development of drugs for rare diseases. This was backed up in November 2018, when the FDA granted accelerated approval to larotrectinib for the 'treatment of adult and paediatric patients with solid tumours that have a neurotrophic receptor tyrosine kinase (NTRK) gene fusion without a known acquired resistance mutation, are metastatic or where surgical resection is likely to result in severe morbidity and have no satisfactory alternative treatments or that have progressed following treatment'. Accelerated approval, enables the FDA to approve drugs for serious conditions to fill an unmet medical need using clinical trial data that is thought to predict a clinical benefit to patients.

Also, the EMA has issued a positive recommendation for an accelerated assessment procedure for larotrectinib, only granted to medicinal products which are of major interest from the point of view of public health and, in particular, from the viewpoint of therapeutic innovation.

In terms of unmeasured benefit, the value that an oral oncology medication brings for treating paediatric patients with advanced cancer, in terms of impact on schooling and the further impact on parents should not be underestimated. This compares favourably with treatment regimens requiring daily visits to the hospital as well as admissions to manage adverse events and was highlighted in the clinical validation interviews.

B.2.13 Interpretation of clinical effectiveness and safety evidence

2.13.1 Principal findings from the clinical evidence: clinical benefits and harms

In a disease setting where patients with NTRK fusion-positive cancer have no satisfactory treatment options remaining, larotrectinib provides compelling efficacy, with manageable risks.

Patients with NTRK fusion-positive cancer are a rare patient population, for whom there are currently no approved targeted therapies. The prevailing standard of care in advanced cancer, regardless of the presence of an NTRK gene fusion is typically based on care standards for the tumour site of origin. Initial treatments generally include surgery and radiotherapy; and, for thyroid cancers, radioactive iodine. Systemic therapy options (including chemotherapy and treatment with biologics) are then considered. For many of these patients, ongoing salvage treatments with existing alternatives is often not considered beneficial due to known toxicities of available treatments or co-morbidities of the patient that predict for a deterioration in quality of life with ongoing therapy e.g. limb amputation in infantile fibrosarcoma. Thus, patients with advanced NTRK fusion cancer, who have exhausted (or have no) satisfactory treatment options, represent an area of high unmet medical need.

Larotrectinib has been shown to be effective across multiple tumour types in patients with an identified NTRK gene fusion cancer, who have exhausted all satisfactory treatment options.

Evidence for the compelling targeted efficacy of larotrectinib is based on the pooled interim data of 102 patients from three currently ongoing trials; a dose-finding phase 1/2 study in adults with or without NTRK gene fusions (LOXO-TRK-14001) (4), a phase 2 basket trial in adolescent and adult patients with NTRK fusions (LOXO-TRK-15002 [NAVIGATE]), and a dose-finding phase 1/2 study in paediatric patients with NTRK fusions (LOXO-TRK-15003 [SCOUT]) (5). Due to the rarity of tumours with NTRK gene fusions, the trials were designed to enrol patients with diverse tumour types, thereby testing the ‘histology-independent’ mode of action of larotrectinib.

Treatment of NTRK fusion-positive cancer with larotrectinib exhibited rapid, substantial antitumour activity with durable disease control that appears to be independent of NTRK isoform, tumour type and patient age. There was no effect in patients without TRK fusion cancer, irrespective of tumour type, demonstrating the targeted nature of the therapy.

In the pooled analysis of patients with NTRK fusion-positive tumours, the primary endpoint of overall response rate (ORR) was [redacted]% (n=[redacted]/93) in patients with non-CNS primary tumours and [redacted]% in patients with primary CNS tumours. [redacted] patients with non-CNS tumours achieved a complete response (CR) on larotrectinib. Stable disease was reported in a further [redacted]% (n=[redacted]/93) patients with non-CNS primary tumours and [redacted]% (n=[redacted]/9) patients with primary CNS tumours. A comparison of larotrectinib response rates by tumour type versus response rates to standard of care treatments [obtained from the literature] Table 25 suggests that the ORR observed across the larotrectinib-treated population may exceed that achievable with available currently authorised treatments. It must also be borne in mind that the proposed indication for larotrectinib places the product in a setting where no satisfactory treatment options remain, where one would usually expect worse responses to therapy than the previous line of therapy (see also Table 23 presenting the GMI analysis).

Results for the secondary endpoints in the pooled analysis are supportive of the primary endpoint. Disease control rate (confirmed CR, surgical CR, PR, or stable disease lasting 16 weeks or more following the initiation of larotrectinib) was [redacted]% (n=[redacted]/93) for non-CNS primary tumours, and [redacted]% (n=[redacted]/9) in patients with primary CNS tumours.

Median time to response (TTR) was short at [redacted] for patients with non-CNS primary tumours and [redacted] in patients with primary CNS tumours, with most patients ([redacted]%) responding within [redacted] months or less. A short time to response is considered of value in the treatment of any metastatic tumour since tumour shrinkage may reduce tumour symptoms and enable rapid onset of patient benefit. This was borne out in the health-related quality of life assessment results from patients in the NAVIGATE and SCOUT studies, where patients with NTRK fusion-positive cancer experienced a rapid and sustained improvement in health-related quality of life (HRQoL) compared to baseline,

which occurred within [redacted] months of the start of treatment. For the majority of patients, scores reached or exceeded the minimal important difference (MID) threshold that translates into a meaningful improvement to the patient. In patients with critical visceral disease or patients at risk of debilitating events (e.g. paralysis due to spine involvement) a short TTR is even more important. In addition, a short TTR would allow a quick understanding of whether the therapy is effective or not.

After a median duration of follow-up [redacted] months (25th, 75th percentiles; [redacted]), the median duration of response (DoR) was [redacted] in non-CNS primary tumours. This is consistent with the high proportion of patients still in response ([redacted]%; n=[redacted]/67). Responses are durable, with [redacted]% of responding patients having a response lasting 6 months or more (95% CI: [redacted]), and [redacted]%, 12 months or longer (95% CI: [redacted]). Based on the present median duration of response follow-up time of [redacted] months, a median response duration of at least [redacted] months may be expected, which would generally be considered a clinically relevant duration of response, regardless of line of treatment in any/most metastatic solid tumours.

Median duration of progression-free survival (PFS) is [redacted]4 months (95% CI: [redacted]) (median duration of follow-up [redacted] months) in patients with non-CNS primary tumours. In primary CNS tumours, median PFS is [redacted] months (95% CI: [redacted]). Overall survival (OS) data are immature with the median duration of OS [redacted]. In patients with non-CNS primary tumours, the 1-year OS rate is [redacted]% (95%: [redacted]) and is [redacted] in patients with primary CNS tumours.

Along with compelling efficacy, larotrectinib has also demonstrated a tolerable safety profile, with the majority of treatment-related TEAEs being of grade 1 or 2, easily monitored and managed, and reversible. The safety profile of larotrectinib is comparable across adult and paediatric patients, and tumour types.

These data demonstrate that it is now possible to treat a patient based on the type of mutation (gene fusion) their tumour contains, regardless of where the cancer originated. This is considered a therapeutic advance when compared with traditional chemotherapy

which tends to be of limited benefit in many tumours, non-specific with respect to molecularly defined targets, generally associated with significant toxicities, and often unsuitable for certain patient populations (e.g. very young paediatric patients or adult patients who may be elderly or frail).

In the context of the population indicated for larotrectinib / defined in the decision problem, larotrectinib offers a targeted therapy, for which, given the beneficial treatment effect seen in a diverse patient population, any known risks are considered acceptable. This is especially considering the alternative for patients at this stage, which could be non-targeted, more toxic systemic therapies or amputation or disfiguring surgery.

2.13.2 A discussion of the strengths and limitations of the clinical evidence base for the technology.

A key strength of the evidence base is that larotrectinib is demonstrated to provide compelling efficacy, balanced with an acceptable risk profile, in a population with significant unmet medical needs, who having exhausted all (or have no) standard therapy options, or were about to undergo disfiguring surgery such as amputation. The results were also confirmatory of the pharmacological action of larotrectinib, enabling the target population i.e. people with NTRK fusion-positive tumours, to be clearly defined.

In addition, disease assessment across the three contributing studies was performed using standard recognisable methods (i.e. computed tomography (CT), positron emission tomography (PET), and/or magnetic resonance imaging (MRI)) and tumour response criteria (i.e. RECIST version 1.1 for non CNS solid tumours and RANO for CNS tumours) for assessing cancer and monitoring response to treatment. Furthermore, assessment for key endpoints (i.e. response, time to response, time to best response, duration of response, disease control rate, PFS) in the pooled analysis were performed by a centralised independent review committee.

The evidence base also provides information on the efficacy of larotrectinib in *both* paediatric and adult patients (range [redacted] years of age (ePAS2: [redacted] years; SAS3: [redacted] years)), facilitating a more timely approval and use in the paediatric

population, which in the past has lagged behind use of new medicines in adults. This is particularly important if no suitable or standard treatments are available, or patients face an alternative of amputation or disfiguring surgery.

Health Technology Appraisal processes are presently more typically geared to assessing clinical and cost effectiveness of therapies based on evidence from prospective, randomised, well-controlled, double-arm clinical trials, and, in the case of cancer therapies, consider only one tumour site per assessment. Using solely these criteria, the evidence base for larotrectinib may be considered to have several limitations given that clinical evidence was not obtained through randomised, or comparative studies, and the target of larotrectinib (NTRK gene fusions) occurs in a diverse range of solid tumours, requiring a histology-independent approach to treatment.

Although the gold standard for generating the most reliable evidence of a drug's efficacy and safety is randomised controlled trials (RCTs), there are situations where randomised trials are not feasible or ethical, particularly for rare diseases (3, 4). Patients with NTRK fusion-positive solid tumours are a rare patient population, and a traditional disease-specific study design is not feasible owing to insufficient patient enrolment. Also, patients have no other satisfactory therapy options and no specified standard of care at this point in their clinical management, making it difficult and potentially unethical to select an appropriate comparator. Clinical evidence for larotrectinib is thus based on single arm ‘basket’ studies which enrol patients who have the same molecular feature i.e. NTRK gene fusion, across anatomically and histologically diverse solid tumours. A benefit of this approach was that the ‘histology-independent’ mode of action of larotrectinib could be tested. Basket trials are considered best suited to assess the efficacy of targeting genomic alterations that occur at low frequencies across a wide variety of tumour types (3).

Use of a non-standard study design also means there is little experience in the most appropriate method of pooling and controlling for data obtained from single-arm basket studies. Nevertheless, the patients included in the pooled analysis share the same characteristics as would be expected in the larotrectinib-eligible population within clinical

practice and defined within the decision problem i.e. have a documented NTRK gene fusion and have exhausted all standard treatment options.

Other limitations of the use of basket trials is the lack of a comparator arm to inform a head-to-head comparison, and relatively small numbers of patients with each tumour type. Currently, there are no approved targeted therapies for patients with NTRK fusion-positive cancer. The major therapeutic advantage that larotrectinib will provide is in a disease setting where patients have no satisfactory treatment options remaining. They will have failed to respond to standard of care, did not tolerate it or do not have any standard of care for treatment. This represents a therapeutic setting of high unmet medical need. In this context, defining an appropriate 'comparator' and demonstrating a therapeutic advantage over existing therapies is difficult since there would be no further treatment options remaining for patients. A further challenge arises due to the histology-independent nature of larotrectinib treatment and the fact that it can treat multiple tumour types harbouring NTRK gene fusions. Thus, salvage therapies and standard of care treatments will vary across tumour types, meaning there isn't one common standard of care therapy that can be identified as 'the comparator' in this patient population.

In this submission, given the afore-mentioned challenges, the company deemed the most appropriate 'standard of care comparator' to be based on a mixed basket reflecting the most appropriate option after patients have exhausted all satisfactory treatment options [redacted] according to tumour type (weighted by patient enrolment per tumour location). The reference case stipulates that data to inform the comparator arm of the model must be sourced systematically. In adhering to the reference case, systematic literature reviews have been conducted across each tumour location. However, there is no recommended process as to how this data should be aggregated to inform a comparative assessment. Difficulties in comparing larotrectinib to standard of care in this way include the likely heterogeneity between study populations. Also, with larotrectinib patients having exhausted all other treatments, they could be expected to have poorer outcomes than published best supportive care data.

Hence, the comparison may not accurately reflect a comparable ‘standard of care’ in larotrectinib-eligible populations in clinical practice.

The gold standard for demonstrating clinical benefit of cancer therapies is overall survival, however analysis of overall survival requires longer follow-up than other endpoints. The primary endpoint in the pooled analysis for larotrectinib efficacy was Overall Response Rate (ORR). ORR is often used as a surrogate endpoint in (accelerated) approval of treatments intended to treat serious or life-threatening diseases and that generally demonstrate an improvement over available therapy or provide therapy where none exists. Disease response is a real-world tool used daily in the clinic for ongoing assessment of patients, where significant and prolonged reduction of tumour burden can be clinically meaningful. In refractory tumours where no available therapy exists, single-arm trials can be used to assess ORR - the approval of imatinib for the treatment of gastrointestinal stromal tumours (GISTs) was based on the demonstration of a higher ORR.

Finally, as the three larotrectinib studies are ongoing and the submission is based on the latest dataset from these studies, efficacy data, particularly in relation to PFS and OS are immature (due to the low number of events), and the safety database is limited, especially in relation to long-term safety. Median PFS of [redacted] months in non-CNS primary tumour patients; and [redacted]% of patients are still alive at 12 months or more. These results are supportive of the positive durable response rates and are considered a substantial benefit in this population with limited therapeutic options.

The fact that several aspects of the larotrectinib submission may be unprecedented should not detract from the clinical results achieved with this targeted treatment in a clearly defined patient population with high unmet medical need, and no satisfactory alternative treatment available.

Given the current level of uncertainty, Bayer proposes that whilst data mature, larotrectinib is made available in a timely manner through the Cancer Drugs Fund.

Relevance of the evidence base to the decision problem

Population:

[redacted] (pending definitive EC decision).

The presence of an NTRK gene fusion enables a clear delineation of eligible patients from the entirety of the population. The pooled analysis, forming the basis of clinical evidence for this submission, is entirely comprised of patients confirmed to have NTRK fusion-positive tumours.

No UK patients were enrolled in the studies. Participants in the three studies included in the pooled analysis came from 34 different study centres, most of which were in the United States (US) (n=24). Remaining sites were situated in Denmark, France, Germany, Ireland, Italy, Korea and Spain. The mean age of patients in the pooled analysis is [redacted] years. Due to diversity of tumours harbouring an NTRK gene fusion, and hence the mix of paediatrics and adults studied, age ranged widely from [redacted] years. Of the patients in the analysis, [redacted]% were described as 'White'. There is no information on typical epidemiological features / baseline characteristics for a population with NTRK gene fusion tumours in the UK. Based on the diverse nature of the tumours and age of the patients, the likelihood is that such diversity will also be reflected in any larotrectinib-eligible patients within clinical practice in England. There is no reason to suggest that the tumour - and age agnostic efficacy (and safety) demonstrated in the larotrectinib studies would not be generalisable to the population found in clinical practice in England.

Comparator: Standard of Care.

There are no approved targeted therapies for patients with TRK fusion cancer in the UK. Patients eligible for larotrectinib will have no satisfactory treatment options remaining. They will have failed to respond to or could not tolerate the standard of care, or for some tumours, there may not be any standard of care treatment. Disfiguring surgical procedures or further salvage treatments would be considered unsatisfactory due to

limited benefit, known toxicities, or unsuitability in very young paediatric patients or adult patients who may be elderly or frail.

In light of the evidence base for larotrectinib coming from single-arm / basket studies and the absence of a definitive comparator, as discussed above in section B.2.13.2, a mixed basket reflecting the most appropriate option after patients have exhausted all satisfactory treatment options [redacted] according to tumour type (weighted by patient enrolment per tumour location) was considered the most relevant / closest matching comparator to larotrectinib patients in the proposed clinical setting. In the absence of any data after the final line of approved active treatment, we use a proxy such as the last line of active treatment. On this basis, larotrectinib-eligible patients in a real clinical setting, who will have exhausted satisfactory treatment options, could be expected to have poorer outcomes than the data used in the comparator arm.

Intervention: Larotrectinib.

Larotrectinib is a precision medicine, which specifically targets the protein product of the neurotrophic tyrosine kinase receptor (NTRK) fusion genes, irrespective of the location or histology of the tumour.

The proposed dosing of larotrectinib is:

- Adults: 100 mg taken orally, twice daily, until the patient is no longer clinically benefiting from therapy or until unacceptable toxicity occurs.

- Paediatric population: Dosing in paediatric patients is based on body surface area (BSA). 100 mg/m2 taken orally, twice daily with a maximum of 100 mg per dose until the patient is no longer clinically benefiting from therapy or until unacceptable toxicity occurs.

The pooled analysis included patients from dose-finding studies. Therefore, some patients may not initially have received the proposed licenced dose (maximum [redacted]% patients; dose range 100-150mg b.d. or 9.6mg/m^2 - 120mg/m^2). This did not result in inferior responses to larotrectinib and as soon as the recommended dose had been

established in trials, all patients received it. Patients in clinical practice in England receiving the licenced dose of larotrectinib, would therefore be expected to respond to treatment in a similar way to those studied.

Relevance of the outcomes assessed in clinical trials to clinical benefits experienced by patients in routine clinical practice

Endpoints assessed in the pooled analysis include overall response rate (ORR), median duration of response (DoR), progression-free survival (PFS), and overall survival, all commonly used efficacy endpoints in oncology clinical trials. The latest results for these endpoints are presented in section B 2.6 and summarised above (section B 2.13.1).

Patients with NTRK fusion-positive tumours suitable for larotrectinib treatment will typically be experiencing complications and symptoms due to tumour enlargement and metastatic spread, and possibly a reduced quality of life, with, prior to the introduction of larotrectinib, no further satisfactory treatment to ameliorate disease progression and associated effects. Outcomes in the larotrectinib analysis therefore focus on disease response (including duration), the effect of treatment on clearing, slowing or halting disease, amelioration of symptoms, extending survival and health-related quality of life, all of which are directly relevant to patients within clinical practice.

All efficacy and safety assessments were standard variables and methods for clinical studies in oncology. They are widely recognised as valid, reliable, accurate and relevant to clinical practice and in regular use within the NHS. Disease assessment was performed by computed tomography (CT), positron emission tomography (PET), and/or magnetic resonance imaging (MRI) and tumour response criteria were based on Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1 for non-CNS solid tumours and Response Assessment in Neuro Oncology Criteria (RANO) for CNS tumours.

The primary endpoint in the pooled analysis for larotrectinib efficacy was Overall Response Rate (ORR). Disease response is a real-world tool used daily in the clinic for ongoing assessment of patients, where significant and prolonged reduction of tumour burden can be clinically meaningful.

In summary, larotrectinib represents a significant breakthrough for the small population of patients with advanced TRK fusion cancer who have no other satisfactory treatment options. In such a patient population, larotrectinib is demonstrated to be an effective and safe targeted treatment, exhibiting rapid, substantial antitumour activity with durable disease control that appears to be independent of NTRK isoform, tumour type and patient age.

Given the current level of uncertainty, Bayer proposes that whilst data mature, larotrectinib is made available in a timely manner through the Cancer Drugs Fund.

Larotrectinib as an end-of-life therapy

Evidence to support larotrectinib as an end-of-life therapy, in line with NICE criteria, is summarised in Table 31 and below.

The extensive comparator therapy SLR (outlined in section B.2.9) on a multitude of tumours known to harbour NTRK gene fusions, indicates a limited life expectancy with 'standard of care' treatments in patients who have received ≥ 1 prior therapy. With available treatments median PFS and OS varies across tumour types in patients with progressive, recurrent or metastatic disease. Median PFS was generally less than 12 months across included tumour types, considerably lower than that of larotrectinib (median PFS [redacted] months). On the basis that patients will be eligible for larotrectinib only if there are no other available satisfactory treatment options, and hence, as a subsequent line of therapy to those summarised in Table 30, results would suggest a likely life expectancy for larotrectinib-eligible patients to be within the 24 months NICE criterion.

Table 30. Overview of Treatment Efficacy in Tumour Histologies where NTRK gene fusions occur a (Adapted from (77))

Tumour Type	Treatment Line	Number of patients (N)	ORR (%)	Median PFS (Months)	Median OS (Months)
NSCLC	Second	18–628	2.7–28.9	2.5–5.8	4.7–15.2
	Second or further	49–613	4.2–25.5	2.3–10.3	4.6–not reached
CRC	First or further	205	NR	NR	16.2–16.4
	Second	8–614	11–47.7	0.3–10.5	4–17
	Second or further	24–505	0–28	1.4–7.3	5–14.3
	Third	91–124	8–9	12.9–13.2	NR
	Third or further	57–534	0–13	1–4.8	5.3–11.4
	NR	33–534	43.6–67.5	3.8–8.6	5.2–9.9
Melanoma	First or further	47–555	4–39.9	2.7–6.6	11.5–not reached
	Second	NR	12–29	NR	NR
	Second or further	72–272	0–32	3.1–35 weeks	8.6–16.4
	NR	179–361	4–28	NR	11–14.7
Pancreatic cancer	Second	23–24	NR	3.9–4	2.3–9.1
	NR	10–11	NR	NR	5.2–7.2

Less frequent tumour types					
Thyroid cancer (anaplastic, follicular, or papillary tumor histology)	First	12–75	21.4–61.1	1.6–11	6–56
	First or further	10–80	0–69	1.7–7.4	4.0–12
	Second or further	20–26	0	1.9	3.9–12.3
	NR	19–417	0.5–64.8	2.1–18.3	3.5–not reached
Gliomas	First or further	55–119	NR	NR	5.2–11.3
	Second	32–40	63	3.8–5.8	6.9
	Second or further	14–40	95.2	3–28	7–28.3
	Third or further	9–31	NR	2.9–12.4	12
	NR	20–61	25–53	NR	11–13.8
Biliary	First	41–206	50–81.4	3.7–8	7.7–11.7
STS[b]	First	12–48	17.2–44.4	NR	NR
	First or further	6–175	13.2-86	2.4–15.4	11–46.9
	Second or further	5	NR	6.5	8.9
	NR	7–103	0–66.7	1.92–not reached	9–not reached
GIST	First	19–22	NR	NR	49–not reached
	First or further	141–473	45	18–27.2	46.8–not reached
	Second	41–312	0–7	1.5–30	33 – 37

	Second or further	118–243	0–10	6–22.9 weeks	39–72.7 weeks
	Third or further	17–188	1.5–53	0.9–7.4	7.5–not reached
	NR	25–835	29.9–57.1	0.8–20	9.7–not reached
Bone sarcoma[c]	First	4–180	NR	4.7	5.8–18
	First or further	4–340	10–25.6	3.5–9.3 years	7–20
	Second	116	8	NR	NR
	NR	3–73	0–33	2–12.5	3–87
Rare tumour types					
Salivary gland	First	42	31	6	10
	First or further	57	70.2	8.9	39.7
	Second	18	5	3.5	4
	NR	5–42	0–100	5–7	8.5–18
Appendix	First	54–109	44–56	6.9	11.7–not reached
	First or further	11–54	NR	7.6	2.5 years
	Second	45	NR	2.8	NR
	NR	5–567	39–85	4–44.4	16–not reached
IFS/IM	First	6–20	71–83	NR	NR
	NR	8–9	88.9–93	NR	NR

AE=adverse event; BSC=best supportive care; CMN=congenital mesoblastic nephroma; CR=complete response; CRC=colorectal cancer; EFS=event-free survival; GIST=gastrointestinal stromal tumour; IFS=Infantile fibrosarcoma; IM=infantile myofibromatosis; N=number; NR=not reported; NSCLC=non-small cell lung cancer; NTRK=neurotrophic tyrosine receptor kinase; ORR=overall response rate; OS=overall survival; PFS=progression-free survival; RCT=randomised controlled trial; SAE=serious adverse event; STS=soft tissue sarcoma.

[a] ORR, median PFS, or median OS efficacy data were not available for CMN and secretory breast carcinoma.

[b] STS includes inflammatory myofibroblastic tumor, myopericytoma, spindle cell sarcoma, and peripheral nerve sheath tumour.

[c] Bone sarcomas includes data for chondrosarcomas only.

In terms of extension to life, the survival data, although immature and analysis ongoing, supports durability of larotrectinib effect and extension of life of greater than the 3 months specified by NICE. Larotrectinib represents a step-change in the management of patients with refractory locally advanced or metastatic NTRK fusion-positive solid tumours in that it is a treatment option for patients who have exhausted all other satisfactory treatment options. Larotrectinib should therefore be considered as an end-of-life therapy. Due to overall infrequency of NTRK fusion-positive solid tumours, the number of patients eligible for larotrectinib in England is limited (approximately [redacted] patients in England in total estimated to be potentially eligible for larotrectinib) – see budget impact analysis.

In order to reduce uncertainty in decision making, Bayer proposes that larotrectinib is made available in a timely manner via the cancer drugs fund.

Table 31. End-of-life criteria

Criterion	Data available	Reference in submission (section and page number)
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	Evidence derived from an extensive SLR of clinical efficacy of existing standard of care treatments (according to tumour location and the line of treatment) in tumours known to harbour NTRK gene fusions. Tumours included in the SLRs were reflective of those of patients investigated within larotrectinib clinical studies. NTRK status of patients included in studies in the SLR was not known.	Table 25 in Section B2.9, page 103;
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	In ePAS2 (30 July 2018), the median PFS is [redacted] months (95% CI: [redacted]) (Kaplan-Meier). Six-month and one-year PFS rates are [redacted]% (95% CI: [redacted]) and [redacted]% (95% CI: [redacted]) respectively. Median PFS in the SAS3 patient group is [redacted] months (95% CI: [redacted]) with 6-month PFS rate at [redacted]% (95% CI: [redacted]). The median duration of overall survival is [redacted] after a median follow-up of [redacted] months (25th: [redacted], 75th [redacted]) for ePAS2 ([redacted] patients alive; [redacted]%) and a median follow-up of [redacted] (25th: [redacted], 75th [redacted]) for SAS3 ([redacted] patients alive; [redacted]%). One-year OS rate for ePAS2 is [redacted]% (95%: [redacted]) and [redacted] in SAS3.	Section B2.6; Page 81-85

NTRK= neurotrophic tyrosine receptor kinase;

B.3 Cost effectiveness

Given the current level of uncertainty, Bayer proposes that whilst data mature, larotrectinib is made available in a timely manner through the Cancer Drugs Fund

B.3.1 Published cost-effectiveness studies

Identification of studies

In addition to searches to inform the methodology and assumptions employed in the cost-effectiveness model, a number of steps were taken to identify published cost-effectiveness analyses relevant to the submission:

1. A systematic literature review was conducted to identify published cost-effectiveness analyses considering the treatment of patients with TRK-Fusion cancer (see Appendix G.1)

   - The systematic literature review (SLR) was conducted on 5th May 2019 which confirmed that there were no published cost-effectiveness models in NTRK fusion-positive cancer. Key sources searched were Medline® and Medline in process, EMBASE®, Cochrane library and Econlit® databases. A total of 108 citations were identified in the search.

2. A series of systematic literature reviews (SLRs) were completed for all solid tumours that are known to harbour neurotrophic tyrosine receptor kinase (NTRK) gene fusions that were studied in the larotrectinib clinical trial programme. Studies on treatments (approved and in development) for each tumour site were identified and the economic evidence was synthesised. Details of the methods used to generate relevant cost-effectiveness evidence across tumour sites is presented in Appendix G.2.

Description of identified studies

1. Cost-effectiveness SLR: TRK-Fusion cancer

No published cost-effectiveness studies on the treatment of NTRK fusion-positive cancer were identified during title and abstract screening. This is not surprising given that treatment for NTRK fusion-positive cancer has only very recently become available (Appendix G.1).

2. Cost effectiveness SLR: Tumour site specific

To identify published cost-effectiveness analyses a series of systematic literature reviews was conducted which included each of the tumour sites enrolled into the larotrectinib clinical trial programme in solid tumours that are known to harbour neurotrophic tyrosine receptor kinase (NTRK) gene fusions.

In total, 98 publications were identified across tumour sites. Results of these tumour specific systematic literature reviews were informative for assessment of the model structure, and assumptions used in model development. For full details of this series of SLRs, see Appendix G.2.

In addition to the SLRs, a review of NICE technology appraisals related to oncology treatments based on single arm trial data, as well as a review of NICE technology appraisals that have previously considered multiple histologies was conducted. This was undertaken to identify appropriate analogues for the larotrectinib economic evaluation and inform the model structure and methodology.

B.3.2 Economic analysis

Identified cost-effectiveness studies

No economic evaluations or cost-effectiveness publications considering a TRK-Fusion population were identified.

The series of cost effectiveness SLRs conducted by tumour site, identified publications which provided comparator specific inputs and assumptions. Cost-effectiveness results were not suitable for informing decision making, however inputs and assumptions for relevant comparators were utilised in the model development.

The review of previous oncology NICE technology appraisals did not identify any existing approaches or available economic models that considered multiple tumour sites in a single-arm trial, however there were consistent findings that were incorporated into development of the de-novo model including:

- Partitioned survival models were most commonly used in the oncology disease setting for treatments sharing similarities in trial design to larotrectinib
- Methods for controlling for single arm trials were identified from NICE Technical Support Document 18 (81) and included matched-adjusted indirect comparisons (MAIC), simulated treatment comparisons (STCs) and use of historical control data.

The findings from the reviews were used to help inform model design and are discussed in the sections below.

Patient population - Larotrectinib

The patient population considered in the economic evaluation reflects the patients enrolled in the larotrectinib clinical trial programme and proposed marketing authorisation.

[redacted]

Patients enrolled in the larotrectinib clinical trial programme were heavily pre-treated (average of [redacted] previous systemic therapies), [redacted]% of patients received at least one and [redacted]% patients received >3 prior systemic therapies. The majority of patients had also previously failed surgical ([redacted]%) and radiotherapy treatment options ([redacted]%).

Patients enrolled who had not failed previous systemic therapies ([redacted]%) were not deemed suitable for conventional therapy, where for example the patient's disease stage or severity (i.e. risk of amputation) would have rendered approved therapies ineffective.

In summary the population enrolled in the larotrectinib arm of the economic model reflects patients that have exhausted satisfactory treatment options, where remaining treatment options would not be of clinical benefit.

Patient population – Comparator

In current practice patients are not treated specifically for TRK-Fusion cancer. Patients with TRK-Fusion cancer are currently treated within a broader group of patients characterised by the site of their cancer and the stage of their disease.

The comparator population in the economic model reflects established management without larotrectinib for patients with locally advanced or metastatic disease who have no satisfactory treatment options.

In accordance with the reference case, the economic evaluation considers the relevant comparator(s) technologies displaced by adoption of the new technology. Using clinical guidance and evidence from the literature, this has been defined as a mixed basket of last-line standard of care (SOC) approaches to therapy, including chemotherapies and best-supportive-care. The selection of comparator evidence, as well as additional considerations when simultaneously assessing multiple comparator sources is described in detail later in this section.

Model structure

In establishing the model structure, the preliminary step was a review of the current literature to inform potential or previously used approaches (as described above). Results from the modelling methodology review and the tumour site specific cost-effectiveness SLRs (described in section B3.1) found no precedent for modelling histology independent treatments.

[redacted]

Given the lack of precedence for modelling histology independent treatments, the model methodology was aligned to the NICE Reference Case and prior accepted approaches in economic modelling for oncology treatments where possible. Transparency was also considered important when considering potential approaches, given additional complexities with modelling multiple tumour sites, it was important that the model structure allowed uncertainty and alternative assumptions to be assessed.

The economic model is a cohort state transition model with a survival partition approach. This technique is commonly used in oncology modelling, and is appropriate in capturing progressive, chronic conditions which are described with clinical outcomes requiring an ongoing, time-dependent risk, such as progression and death (82, 83).

Unlike a Markov model, cohort partition models do not require the explicit estimation and use of transition probabilities. Instead, the number of patients in each health state is calculated directly from the treatment and comparator's progression-free survival (PFS) and overall survival (OS) curves. Additional assumptions are only made to estimate the extrapolated portion of the curves. This ensures that the fitted PFS and OS match the

published trial data, and does not require the model to assume that there is a definite relationship between PFS and OS as would be required in a Markov model to calculate transition probabilities between “progressed” and “dead” health states.

It was determined that using patient level data directly from the trial would be more transparent and require fewer assumptions than potential simulation methods.

Both the larotrectinib and comparator arms of the model follow the same health states, however to account for differences in conventional standard of care across tumour sites in the comparator arm health states are stratified by tumour site.

Larotrectinib arm

In each cycle of the model, larotrectinib patients are assigned to one of three mutually exclusive health states according to the proportion of patients who are ’progression-free‘, ’progressed‘, or ’dead‘ (Figure 25).

Figure 25. Larotrectinib arm of partitioned survival model

During development of the model, there has been an on-going assessment to determine whether patients enrolled in the trial can be modelled as sub-groups.

The study design and patient numbers do not allow for any robust conclusions to be drawn about efficacy and safety at individual tumour sites. Stratification of the data is limited by the small number of events in time-to-event outcomes such as overall survival where there have been [redacted] events (15%) over a median follow-up of [redacted] months. Assessment of co-variate models are presented for OS in Appendix L.

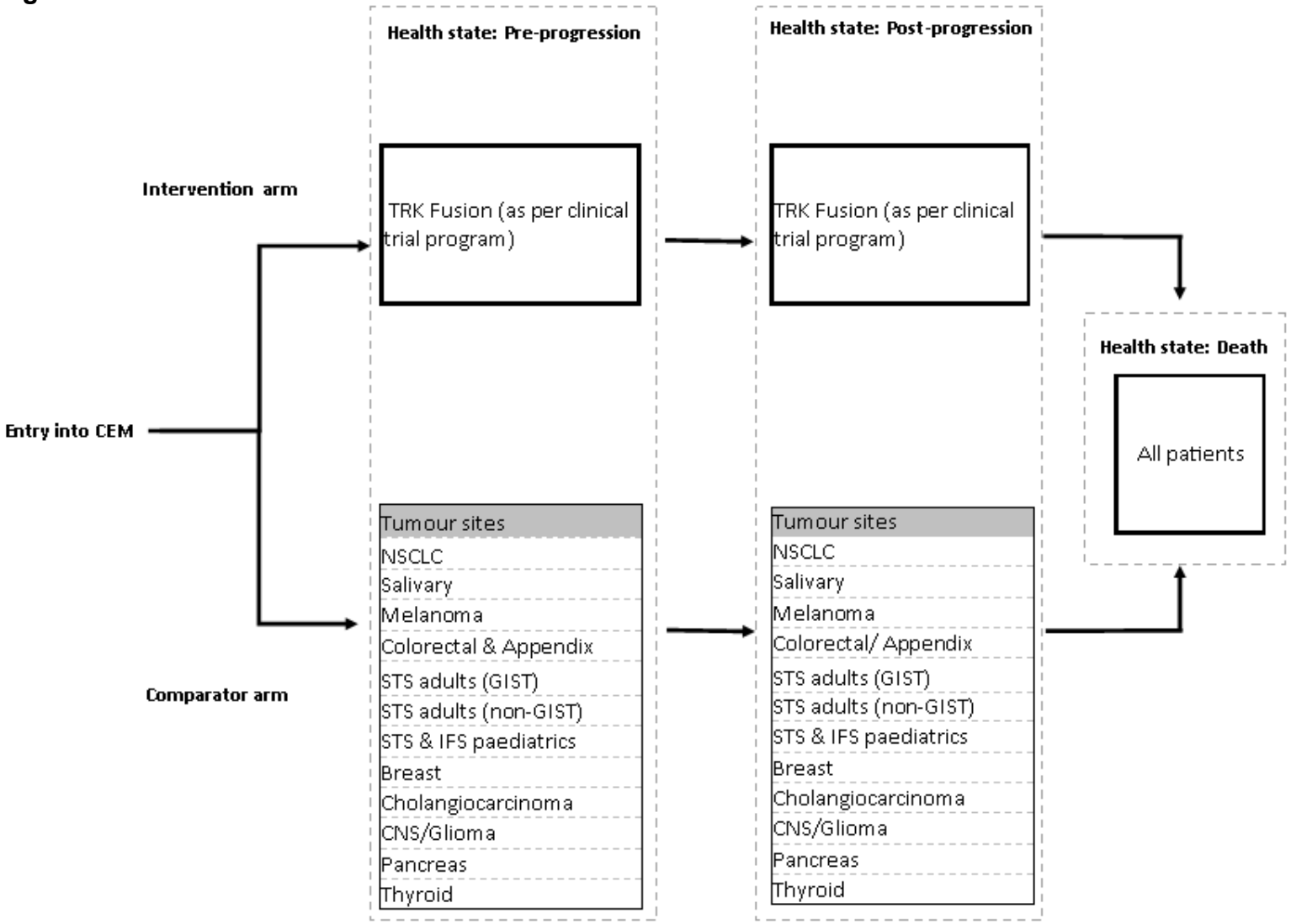
Comparator arm

The comparator arm is also represented by the three health states in Figure 25. However patients are currently treated on the basis of their tumour site and stage of disease, with published evidence and past decision making also organised by tumour site. Examination across these tumour sites show past decision making reflects, different treatments, prognosis, quality-of-life, costs and resource use which all need to be accounted for in the economic evaluation. Therefore a decision was taken to stratify the comparator arm by tumour site reflecting clinical practice.

The comparator arm of the economic model is stratified into 12 model engines reflecting the tumor sites enrolled in the larotrectinib clinical trial programme. Each considers the health outcomes, quality-of-life and costs of patients currently treated in the absence of larotrectinib. It is from these populations and comparators that the eligible larotrectinib population will be drawn. In this sense each tumour site enrolled in the clinical trial programme has its own control reflecting conventional practice.

Figure 26 outlines the model structure, and information on the selection of comparator evidence is presented later in this section, with detailed information presented in Appendix M.

Figure 26. Economic model schematic



Each of the comparator engines independently generates its own results (health outcomes, utilities and costs) for a given tumour-site. These results are weighted based on the number of patients enrolled into the larotrectinib trial to form a balanced control (contributions of each comparator engine are presented in Table 32). Once weighted, the pooled results of the comparator arm can be assessed versus the outcomes derived from the larotrectinib arm of the model and an incremental analysis can be performed.

The benefit of this approach, and stratifying the comparator arms into multiple engines, is that the comparator arm can be constructed and informed based on results from the systematic literature reviews and past NICE technology appraisals. Modelling each tumour site independently as a model engine, avoids the need to synthesise this data into one engine and the loss of transparency and additional assumptions this incurs.

Table 32 presents the larotrectinib patient population by patient per tumour site enrolled into the clinical trial programme. Each of the tumour groups presented reflects an independent model engine, for the comparator therapies, calculating health outcomes and quality-adjusted life-years. Based on the number of patients in each tumour site, the contribution of each comparator engine to the weighted comparator arm of the model results is calculated.

The economic model is designed in accordance with the requirements of the NICE guidance (84), and the ISPOR-SMDM guidelines (85). The economic model was developed in Microsoft Excel® (Microsoft Corporation, Redmond, WA, USA).

Table 32 Tumor site weightings in the economic model

Tumour-site groupings in CEM	Patients per tumour site	Calculated contribution of each comparator engine (rebased to 100%)
STS paediatrics/IFS	[redacted]	[redacted]
Salivary	[redacted]	[redacted]
Cholangiocarcinoma	[redacted]	[redacted]
STS adults (GIST)	[redacted]	[redacted]
STS adults (non-GIST) Bone sarcoma	[redacted]	[redacted]
Thyroid	[redacted]	[redacted]
Colorectal/Appendix	[redacted]	[redacted]
NSCLC	[redacted]	[redacted]
Melanoma	[redacted]	[redacted]
Pancreas	[redacted]	[redacted]
CNS/Glioma	[redacted]	[redacted]
Breast	[redacted]	[redacted]
Total	**102**	**100.0%**

Calculations in the economic model

Within each cycle of the model, patients can either:

- Stay in their current health state
- Move to progressive disease (from the progression-free health state)
- Move to death (from either progression-free or progressed disease health states)
- Patients are not allowed to move backwards in the model.

The proportion of patients in the ’progression-free‘ health state is equal to the survival function value for PFS, while the proportion of patients in the “dead” health state is equal to 1 less the survival function value for OS. Lastly, the proportion of patients in the ’progressed’ health state is equal to the survival function of OS – PFS.

Progression-free and overall survival

The proportion of patients in each health state, and movement between health states, are determined by the survival functions for PFS and OS derived from the larotrectinib clinical trial data.

In the comparator arm, PFS and OS data are modelled independently for each of the tumour sites included in the model. PFS and OS were derived from the literature, either from the results of the SLR, or if identified the clinical data used for decision making from the technology in the relevant NICE technology appraisal. The Guyot method (86) was used to digitalise KM data and parametric survival curves were fitted, to estimate comparator specific health outcomes over a lifetime perspective, following the approach outlined in NICE TSD 14 (87) (further information is presented in section B.3.3 and Appendix M).

Health state utility values

Each health state is associated with a corresponding utility value.

Utility values for larotrectinib are informed by EQ-5D-5L and PedsQL estimates taken directly from the patients enrolled in the larotrectinib clinical trial programme (as described in section B.3.4). For decision making these are mapped to EQ-5D-3L in line with the current positioning statement from NICE (Appendix N).

In the comparator arm health state utilities are applied independently per health state in each comparator engine. Details on the identification and selection of health state utility values are presented in section B.3.4.

Healthcare resource use (HCRU)

Each health state is associated with corresponding healthcare resource use. This varies by tumour-site under conventional care, and this can be seen upon review of past NICE technology appraisals.

In each comparator engine HCRU is applied independently per health state. Sources of HCRU are informed by both published estimates derived from the systematic literature review and previous NICE technology appraisals. Costs applied are based on NHS Reference costs 2017-2018 (88). Details of source identification and selection can be found in section B.3.5.

Recognising that there are currently no published estimates of HCRU or prior clinical experience of histology independent treatments in England, healthcare resource use for larotrectinib is assumed to be equal to that of the weighted comparator arm for each health state. This assumption was considered to potentially overestimate resource use for larotrectinib in validation interviews with clinical experts, who expected resource use to be lower with a targeted therapy (section B.3.10).

Treatment costs

Treatment costs are applied at the start of the treatment cycle with the half-cycle correction turned-off. This assumption accounts for all treatment wastage due to a patient discontinuing treatment during a cycle for any reason, reflecting clinical practice.

Cycle length

The model uses a 7 day cycle length. This cycle length was selected to accommodate the evidence sources used in the model where treatment and assessment of outcomes regularly occur over a set number of weeks. Costs (other than direct treatment costs) and utilities are applied with a half-cycle correction. Health outcomes and costs are accrued and summed for each arm of the economic model.

Time horizon

A lifetime horizon is used in the economic model. For model engines considering adult patients only this was determined to be 40 years, for paediatric populations and pooled populations (adult and paediatric patients) this was determined to be 80 years

Alternative approaches explored for the cost-effectiveness model

Whilst developing the economic model alternative modelling methods were explored especially for controlling for the larotrectinib trial data. Data from the larotrectinib trial programme was generated from single arm basket studies, meaning an in-trial comparison was not possible.

Conventional approaches such as those outlined in NICE TSD 18 (81) were considered but are not feasible for larotrectinib.

- Naïve comparison (unanchored): A conventional indirect treatment comparison (ITC) is not possible due to the absence of a control arm. Guidance suggests an alternative approach in this instance would be to use a comparable evidence source and conduct an unanchored ITC. However no published data source was identified in the SLRs that adequately reflects the cohort of patients enrolled in the larotrectinib clinical trial programme (see Appendix D).

- Matched-adjusted indirect comparison (MAIC): If a comparable source could be found, differences in study population may be adjusted via propensity score matching. However published evidence identified in the SLR was cosigned to a

particular tumor site (Appendix D). In every case matching to a tumour specific source would result in the loss of the vast majority of larotrectinib patients, before covariates other than tumour site could be considered. Due to the low number of events in the larotrectinib arm such an approach would not be suitable for decision making.

- Simulated treatment comparison: This approach is not feasible due to the absence of comparable published data identified through the SLRs.

Based on the results of the model methodology review (described in Section B.3.1) alternative approaches were identified; these approaches and their merits are explored below:

Comparison versus non-responders

A comparison versus non-responder approach uses the non-responders from the larotrectinib clinical trial programme (n=[redacted], [redacted]%) as a proxy for patients not receiving an active treatment. The inherent assumption here is while exposed to larotrectinib, non-responding patients (stable or progressive disease) do not register a treatment response and therefore are not considered to be exposed to a treatment effect.

This approach is limited by the relatively small number of non-responders in the larotrectinib trial, and the inability to balance underlying prognostic factors such as tumour site within the sample. These issues aside this approach has previously been criticised as it requires the clinical assumption of equivalence between the responders and non-responder groups prior to treatment exposure. In addition it assumes the substitution of outcomes from the non-responder population to a non-exposed population.

Finally as well as reflecting a non-exposed patient, outcomes must also be considered representative of patients in England not receiving larotrectinib. A summary evaluation of the approach is presented in Table 33 below, exploratory scenario analyses are presented in section B.3.6.3.

Table 33. Assessment of larotrectinib vs. non-responder control

Responder versus non-responders	Benefits: • Data is collected within the clinical trial programme, meaning all patients will have met the same pre-specified inclusion/exclusion criteria Limitations: • Number of non-responders is small (n=[redacted]), and is not large enough to adjust for differences in the patient groups • Non-responders may be inherently different to responders, this status could be linked to prognosis independent to treatment exposure • Cost-effectiveness analysis considers the incremental benefit versus currently used treatments. Patients do not receive a comparator treatment • Level of treatment effect in comparator arm is unknown, some patients may receive a treatment benefit (but not register a partial or complete response) others may not receive any benefit

Comparison with previous line of therapy

Comparison with prior line of therapy assesses a patient's health outcomes with larotrectinib (PFS/ORR) versus outcomes obtained using their previous line of therapy (TTP/ORR). Being trial based, and a self-comparison, results in the patient population being well controlled in terms of demographic factors and to a large extent clinical factors, with only time-dependent variables such as stage of disease expected to change. This makes the analysis conservative with a bias against the later line treatment and provides strong evidence as to the treatment effect of larotrectinib on these outcomes. The clinical results of this comparison are presented in section B.2.6 and Appendix Q.

Comparison with prior line of therapy, whilst informative from a clinical perspective, is difficult to incorporate into an economic evaluation.

Firstly the analysis is highly conservative, the condition of a patient and their prognosis is expected to deteriorate following disease progression. The analysis can attempt to account for this difference by restricting the analysis population to those that received prior treatment to a metastatic setting.

Censored larotrectinib patients who are currently progression-free also contribute to the analysis, however their PFS value is constrained by their trial follow-up. As the clinical trial for larotrectinib has on-going recruitment, follow-up times are heavily skewed. For these patients the censored PFS value is compared against an absolute value of time-to-treatment progression from their prior therapy. This biases against larotrectinib.

The major limitation of this analysis is it cannot provide overall survival data for the patient's prior therapy line (as they had to have survived following disease progression to receive larotrectinib). This makes a comparison of cost-effectiveness heavily reliant on further assumptions.

In terms of evaluating costs, patients in the previous-line of therapy received active treatment, however the treatments received vary substantially and may not have been reflective of treatments received in clinical practice in England.

A scenario analysis is presented in section B.3.6.3, exploring the use of this analysis in the economic evaluation. An assessment of this approach is presented below (Table 34)

Table 34: Assessment of self-control comparison with previous line of therapy

Self-control comparison with previous line of therapy	Benefits: • Methodology forms an internal control and is naturally suited to a basket trial • Naturally conservative (baseline status likely to decline over the course of disease) Limitations • Many larotrectinib patients remain progression-free (and censored) these are compared against unrestricted TTP from the previous line, biasing this result against the later line treatment. • No comparative overall survival • Comparison depends on patients previous treatment, this varies substantially, and may deviate from that used in clinical practice

Intervention technology and comparators

Larotrectinib is an orally bioavailable, adenosine triphosphate (ATP) competitive, potent and highly selective TRK inhibitor that was rationally designed to avoid activity with off-

target kinase. The target for larotrectinib is the TRK family of proteins inclusive of TRKA, TRKB, and TRKC that are encoded by NTRK1, NTRK2, and NTRK3 genes, respectively. **Importantly, if approved by EMA, larotrectinib will be the first histology independent drug approval in the EU.**

Comparator technologies

There are no existing treatments for patients with NTRK gene fusion cancer. The relevant comparators (treatments used in the absence of larotrectinib) are typically dependent upon both the site of the tumour, and the line of treatment. The positioning adopted in the economic model and evidence sources is aligned with the anticipated marketing authorisation and discussed earlier in this section.

Selection of comparator evidence source was guided by available data on current standard-of-care, which is limited in some of the rarer tumour sites. Results from the clinical validation suggest that in tumour sites where chemotherapy is not considered efficacious, it would be displaced by an available TRK inhibitor.

Approach to selection of comparator technologies and model inputs

SLRs were conducted to identify clinical, economic and health-related quality of life evidence for each of the tumour sites represented in the larotrectinib clinical trial programme known to harbour neurotrophic tyrosine receptor kinase (NTRK) gene fusions. These looked to exhaustively document the comparator evidence base. These reviews are described in sections B.2.9, B.3.1, B.3.4 and B.3.5 and the corresponding appendices.

A large number of model inputs are needed to populate an economic model for a conventional oncology appraisal which typically considers a single tumour site. For a histology independent treatment this number is multiplied, as the stratified comparator arm must reflect all tumour sites enrolled and the respective care currently provided in clinical practice.

Accordingly, in-line with a conventional technology appraisal an assessment must also determine the model assumptions. For oncology models these often include endpoints not fully captured within the clinical trial. For histology independent treatments these assessments must again extend to all tumour sites within the stratified comparator arm.

Approach to comparator selection and source data

Due to the number of tumour sites modelled in the stratified comparator arm, one comparator has been selected per each tumour site. This reduces the number of potential combinations of comparator treatments, which become unmanageable (over 100 combinations) as soon as multiple comparators per tumour site are considered. The partitioned survival structure was selected to be transparent and to allow other evidence sources to be explored. Following a clinical validation an alternative source was implemented for one of the comparators (STS non-GIST, pazopanib) as a scenario analysis (section B.3.10, Appendix M for further details).

Data selection for comparators followed a pre-specified algorithm and hierarchy starting with the series of tumour-site specific SLRs. In the absence of available data for a particular tumour-site, assumptions and groupings based on clinical rationale were explored. This is explained in detail in the following pages.

Step 1 – **Systematic literature review: relevant NICE technology appraisals**

Where identified in the SLR, previous NICE technology appraisals that met the inclusion criteria were selected as sources for each tumour site in the economic model.

Whilst this approach departs from that used in conventional oncology appraisals, it was considered a pragmatic approach, given the number of tumour-sites and accompanying assumptions needed to populate each of the tumour site engines in the stratified comparator arm.

Given the process and scrutiny undertaken in each technology appraisal to select the Committee's preferred inputs and assumptions these sources were determined to be

most suitable for decision making in England, and allowed the data and assumptions used in the model to reflect the Committee's preferred assumptions. This minimises uncertainty, and allows incorporation of input from the wide range of stakeholders who contributed to previous appraisals. If more than one previous NICE TA was eligible, further criteria were assessed to select the most appropriate TA (presented below).

***Step 2* - Systematic literature review: clinical publications**

For tumour sites where there was no available or suitable previous NICE technology appraisal results from the clinical SLR were assessed to confirm the most appropriate evidence source that had published data to populate the CEM. If several publications were eligible, further criteria were considered to select the most appropriate publication (presented below).

Step 3 – Targeted literature searches: Expanding the original SLR scope

The SLRs outlined in the previous steps targeted the population enrolled into the larotrectinib clinical trial programme. For tumour sites characterised by multiple subtypes (e.g. approximately 50 for soft-tissue sarcoma), the SLRs did not pick up relevant publications and thus expansion was required. The objective here was to align each as closely to the enrolled trial population as possible. As an example, TA185 (used for inputs for "adult soft tissue sarcoma (nGIST))" was not identified in the SLR as the searches were specific to six subtypes. However the NICE recommendation for TA185, is reflective of patients enrolled in the larotrectinib clinical trial programme. Decisions such as this were validated through clinical interviews (section B.3.10).

Targeted searches included searches for UK, European, American treatment guidelines in addition to scientific publications. For scientific publications, targeted searches were performed in PubMed combining disease terms (e.g., 'cholangiocarcinoma') and known comparator terms (e.g., 'gemcitabine cisplatin') in combination with terms describing the evidence of interest (e.g., 'controlled trial', 'cost'). The results were screened following the same process as the SLRs, first with title and abstract screen, and then full-text review if

deemed relevant. If several publications were eligible, further criteria were considered to select the most appropriate publication (see section presented below)

Step 4 *- Tumour groupings*

Finally if no relevant published studies were identified that could inform the economic model, the final step was grouping tumour sites. This occurred in the rarer cancer sites, for example where published survival data for a pre-treated population was not available. Groupings were based on discussions with oncologists and a review of the literature to assess the validity (section B.3.10). Comprehensive details are presented in Appendix M.

Methods to deal with multiple sources

When several sources were identified, further elements were taken into consideration to select the most appropriate source:

1. Where multiple past NICE Technology Appraisals were identified, the following criteria were considered :

    - Appraisals with the later line of therapy (e.g. last line of systemic therapy) reflecting the point at which a patient would have exhausted all satisfactory treatment options.

    - Extent to which the publication used to inform the Technology Appraisal matches the inclusion criteria as applied in the larotrectinib clinical trial programme.

    - Treatment having a positive recommendation by NICE indicating acceptance for routine commissioning and use in UK clinical practice

    - Date of publication prioritising more recent appraisals where multiple treatments have been assessed in the same line of therapy for the same tumor type.

2. If a publication is considered, the appropriateness of the identified publications in the SLR and targeted searched were judged based on:

- Reporting of outcomes that can inform the health economic model inputs such as KM plots on progression free survival and overall survival.
- Extent to which the publication matches the inclusion criteria as applied in the larotrectinib trial protocol
- Published source with the later line of therapy (e.g. last line of systemic therapy) reflecting the point at which a patient would have exhausted all satisfactory treatment options.
- Date of publication prioritising more recent publications

Information regarding source data for each tumour site is presented in the following sections with detailed information for each comparator presented in Appendix M.

B.3.3 Clinical parameters and variables

Overview

The sections below detail the data sources and assumptions used to populate the larotrectinib and comparator arms of the economic model.

Time to event data

The economic model partitions patients into the progression-free (PF), progressive disease (PD) and death health states by means of time-to-event data. For larotrectinib transitions between health states are determined by progression-free survival (PFS) and overall survival (OS) data taken directly from the larotrectinib clinical trial programme. This data was immature and required extrapolation, a summary of survival analyses conducted is presented in this section, with a comprehensive report presented in Appendix L.

For comparators, data were taken from relevant technology appraisals (TAs), or publications identified to represent the efficacy of each of the tumour locations in the pooled comparator. Where immature, this data was extrapolated, information on survival analysis for comparators in the economic model is presented in Appendix M.

Parametric modelling

The assessment of appropriate parametric models used to inform the partitioned survival analysis followed the recommended approach by the NICE Decision Support Unit (DSU), as well as recommendations from published literature (87) (89, 90). The parametric models assessed assume that survival times for patients follow a given theoretical distribution (exponential, Weibull, Gompertz, lognormal, log-logistic and generalised gamma) (89). Each was tested for their goodness of fit in accordance with the NICE DSU Technical Support Document (87).

Larotrectinib

A systematic approach was taken to determine the appropriate parametric models for time to event data according to the algorithm outlined in the NICE DSU guidance (87). Following the guidance, all standard parametric distributions based on one and two

parameter survival functions were assessed. For each parameter the following steps were taken:

Step 1. Visual inspection of KM estimates and log-cumulative hazard plots to assess the types of hazards observed in the dataset

Consideration of the observed hazard rates over time is important when considering suitable parametric models, as different parametric models incorporate different hazard functions. In the case of non-linearity, alternative modelling methods were explored. The visual assessment of log cumulative hazard plots also allowed for assessment of the proportional hazards (PH) assumption (in the case of 2 or more treatment groups). However, as this is not relevant given the single arm trial design, assessment of the PH assumption was not performed.

Step 2. Assessment of model fit to observed data by considering how closely it followed the KM curve visually

The standard log-cumulative hazard plots were transformed to test the suitability of distributions based on the methods described by Collett et al., 2003 (91). It is important to note that this method of assessment is uncertain and can be inaccurate if censoring is heavy and observed data points are clustered at certain points along the KM curve, as is the case with the larotrectinib survival data. Hence the use of this approach for assessing the suitability of parametric models was supplemented with additional tests and benchmarking.

Step 3. Tests of Internal and External Validity

To test internal validity of the model fittings, the Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC) statistical tests were performed to assess the relative fit of alternative parametric models to the observed data. For larotrectinib given the immaturity of time-to-event data, interpretation of these results is not overly informative. Evaluation of clinical plausibility of the distributions models was completed for each model based on the amount of time it would take for 10% and 1% of patients to remain alive or progression free, as relevant.

Step 4. Final model selection (base case and scenarios)

Model selection was based on the appropriateness of fit to the observed data and the plausibility of the extrapolated portions of the curves. If there was more than one plausible model, alternative models were considered in scenario analysis (see section.B.3.6).

Details regarding the results of the above steps to evaluate the goodness of fit of the survival models fitted to time-to-event data for larotrectinib are available in Appendix L, which are also summarised below.

Model selection summary

Step 1: Assessment of hazards

Survival analysis was initially conducted on an earlier data-cut. Visual inspection of log-cumulative hazards saw some variation over time, leading to an exploration of more complex survival models (presented in Appendix L). It was not clear whether the change in the hazard rates was driven by a low number of events, or reflected an underlying change in the risk of events. The assessment found that whilst complex models could fit the observed data, there was currently no rationale to select these over the models recommended in NICE TSD 14. Assessments of more complex models was not conducted for the dataset used in this appraisal.

Step 2: Visual assessment to observed trial data

The standard parametric models fitted to the observed data did not appear to be clearly distinguishable when comparing visual fit to the KM data across the trial follow-up period. However, when looking over an extended time horizon (up to 400 months), projections varied substantially across models (Figure 27).

Figure 27: Parametric model fittings for larotrectinib for PFS and OS across time horizons

Step 3: Assessment of internal and external fit.

For all time to event parameters, the AIC/BIC values were closely clustered, with the difference too small to inform selection of one model over another. Given the lack of published data on patients with TRK-Fusion cancer, it was difficult to benchmark the projections against external data sets, such as other oncology studies or natural history studies.

An assessment of clinical acceptability determined that when using the lognormal, log logistic, Gompertz and generalised gamma distributions, patients overall survival exceeded current UK life expectancy (based on published all-cause mortality rates (92)). This could suggest that due to the efficacy of larotrectinib people no longer die of their cancer, but rather other causes. However in considering the base case, and immaturity of the OS data, a conservative approach was adopted in considering only

the [redacted] and [redacted] model. Both of these models underestimate PFS and OS versus the observed trial data (section B.3.10).

Both the [redacted] and [redacted] distributions could plausibly be used to model the time to event inputs. Applying the [redacted] is a more simplistic approach as it relies on one parameter rather than two; however, it assumes a constant hazard throughout lifetime as it does not account for the change in survival hazards with aging (92).

The [redacted] distribution provides closer estimates to the later points of the KM than the [redacted] (whilst still slightly underestimating the observed data) and also tends to be cited as more appropriate for modelling the change in hazards with aging (93). The [redacted] was therefore selected as the base case for modelling survival of larotrectinib patients for PFS and OS.

All other distributions were tested in scenario analysis. For those survival models that exceeded current life expectancy, the upper limits of the OS parametric models were adjusted by the background all-cause mortality reflective of the UK population to ensure the projected cohort did not remain alive longer that the general population. This is described in additional detail under “Age-Specific Mortality” on page 166. Scenarios considering cost-effectiveness results using these extrapolations are presented in section B.3.6.3.

Figure 28 and Figure 29 below provide a comparison of the Kaplan Meier (KM) curves for PFS, OS with the fitted survival models. Table 35 and Table 36 provide the parametric model coefficients and fit statistics for the respective distributions considered to be plausible fits.

It is recognised that survival data from the larotrectinib clinical trial programme is currently immature, driven by low event numbers, therefore the estimates of overall survival are subject to uncertainty. Bayer is making this submission with a view to consideration for access via the Cancer Drugs Fund with a view to continue collection of survival data. Further details of the currently planned ongoing data collection can be found in section B.2.11.

Figure 28. Larotrectinib PFS curves based on the larotrectinib clinical trial programme patients: KM versus extrapolated curves

Table 35. Parametric model coefficients and fit statistics for larotrectinib PFS survival models

Distribution	Shape	Scale	AIC	BIC
Exponential	[redacted]		328.56	330.17
Weibull	[redacted]	[redacted]	329.24	332.46

AIC – Akaike information criterion; BIC – Bayesian information criterion

Figure 29. Larotrectinib OS curves based on the larotrectinib clinical trial programme patients: KM versus extrapolated curves

Table 36. Parametric model coefficients and fit statistics for larotrectinib OS survival curves

Distribution	Shape	Scale	AIC	BIC
Exponential	[redacted]		161.036	161.675
Weibull	[redacted]	[redacted]	162.756	164.034

AIC – Akaike information criterion; BIC – Bayesian information criterion

Pooled comparator

Data collection for the clinical parameters reflecting the pooled comparator followed the algorithm outlined in section B.3.2.

For tumour locations where a relevant NICE recommended treatment option was available, the approach taken was to simulate the clinical outcomes, health-related quality-of-life and costs of the NICE appraisals that reflected the Appraisal Committee's preferred assumptions.

To determine these, an assessment of all relevant documents including the Manufacturer Submission, ERG report, and Final Appraisal Determination documents

were reviewed. The methodology and the relevant data used to inform the Committee's preferred assumptions were then extracted.

To check the internal validity of this process, the individual comparator outcomes in the model were benchmarked to the results as reported in the relevant results section of the NICE Technology Appraisal. Once benchmarking demonstrated that the modelling replicated the results, the model was calibrated to incorporate the methods agreed by the NICE Appraisal Committee.

For comparators where there was no NICE recommended treatment option available, data collection for clinical inputs for these comparators was based on the clinical, health-related quality-of-life and economic SLRs, as described in section B.3.2.

Pooled standard of care: Extrapolation of health outcomes

For tumour sites where evidence on health outcomes was based on previous NICE Technology Appraisals extrapolation of time-to-event data (PFS and OS) followed the Appraisal Committee's preferred assumptions in the respective NICE TA.

Survival modelling was conducted independently for each of the comparators included in the model. The parametric model coefficients were extracted directly, if available, and the survival models were fitted. For some previous appraisals, data from Kaplan-Meier plots was used directly, where this occurred the approach was replicated. Where the KM data was not complete the AC incorporated extrapolations to derive the survival estimates for the tails. The same process was replicated in this analysis. This information was either taken directly (where presented) or digitized from the KM curve to replicate this approach.

In several instances, the methods used to model time-to-event outcomes in previous appraisals were more complex than application of a simple fitted distribution. Specifically, for comparators in breast, non-small cell lung cancer (NSCLC), thyroid, and melanoma tumours, a more complex model such as a piecewise or spline model was determined by the NICE Committee to be the preferred method. These more complex methods were replicated as closely as possible, however, exact implementation was on occasion limited by the data available (further information outlining these methods is provided in Appendix M).

Where clinical data was sourced from the literature, an extrapolation and goodness of fit assessment (i.e., AIC/BIC) was performed using extracted KM curves from the clinical efficacy source. These curves were digitized and reconstructed using the Guyot method (86). This process included an initial step to digitise the KM curves from the published sources to create a dummy patient level data set based on a combination of the number at risk at each time point and the respective survival distribution. Following the Guyot method, the dummy patient level data set was then fit to different distributions to obtain the shape and scale parameters necessary to generate the parametric curves allowing for extrapolation. Each curve was assessed for quality of fit using AIC and BIC and visual inspection. The models selected in the base case for each tumour site are shown below, along with the criteria used for model selection. The impact of fitting alternative survival models for the different tumour sites are presented as scenario analyses (section B.3.10). A summary of model selection is presented in Table 37 with details regarding the methods performed for each comparator to extrapolate time to event data presented in Appendix M.

Table 37. Base case model selection by tumour site

Tumour Type	Model Selection	Rationale
NSCLC	[redacted]	Approach used in TA374
Salivary	[redacted]	PFS: no KM available; assumed exponential with parameter calculated from median survival OS: best statistical fit
Melanoma	[redacted]	Approach used in TA357
Colorectal/Appendix	[redacted]	Approach used in TA405
GIST	[redacted]	Approach used in TA488
Non-GIST/Bone sarcoma	[redacted]	Approach used in TA185
STS paediatrics/IFS	[redacted]	Best statistical fit
Breast	[redacted]	Approach used in TA423
Cholangiocarcinoma	[redacted]	Best statistical fit
Glioma	[redacted]	Best-statistical fit and fit by visual inspection
Pancreas	[redacted]	Approach used in TA440
Thyroid anaplastic, follicular and papillary	[redacted]	Approach used in TA535

NSCLC, non-small-cell lung cancer; GIST, gastrointestinal stromal tumours; STS, soft tissue sarcoma; IFS, infantile fibrosarcoma; KM, Kaplan-Meier

Although each tumour location has been modelled independently, results reflecting the pooled survival curves were also generated to illustrate the cohort values. The curves in Figure 30 demonstrate pooled PFS and OS curves for the comparator, using the trial-based weighting for each base case comparator parametric fit to generate average curves over time.

Figure 30. Pooled comparator PFS and OS survival models

All-cause mortality

Age- and sex-specific all-cause mortality rates for the general UK population were also calculated for each cycle. In any period and for any treatment where modelled OS suggested lower mortality than the general population, all-cause mortality hazard rate based on the UK Office of National Statistics was used instead of the study based estimate (92). This corrects for the long tails for some of the parametric fits for larotrectinib. However, these adjustments to mortality were not triggered for any of the comparators.

The average age for both adult and paediatric cohorts and the male-to-female ratio were based on the patient characteristics of the larotrectinib clinical trial programme patients. The age was tracked by modelling cycle to determine the background mortality hazard rate within each cycle, which was used where the background mortality hazard was observed to be greater than the specific survival curve.

Adverse events

Treatment emergent grade 3-4 adverse events (AEs) that occurred in ≥5% of patients in the relevant treatment arm were included within the economic assessment. The 5% threshold was based on common assumption used in NICE technology appraisals (93-95). Treatment emergent adverse event rates for larotrectinib were taken from the larotrectinib clinical trial programme safety population (n=137) (see section B.2.10).

Adverse events for the comparators were taken from the respective sources that informed clinical efficacy. For tumour sites sourced from previous NICE TA, AE rates were derived from the publicly available appraisal documents, with a few exceptions. For NSCLC, only pooled AE rates were provided in the TA. Therefore, the original rates from the study population of interest were reviewed and included in the model (96). For CNS/glioma and cholangiocarcinoma, rates from the relevant clinical trial were used (97, 98).

AEs that occur in ≥5% of patients in the relevant treatment arms were extracted for each tumour site comparator. Review of the rates across larotrectinib and the comparator tumour locations revealed that applying the 5% criterion to the individual tumour locations may bias the results, overestimating the AE rates and impact for the comparator arm. For example, an adverse event may have occurred in >5% of the source publication, however when pooled this reflected <5% of the adverse events in the pooled comparator population.

The AE rates from the comparator sources were weighted based on the tumour distribution in the larotrectinib clinical trial programme (Table 32 in Section B.3.2). Only AEs with a weighted rate in the full comparator sample of 5% or higher were included for model calculation. This is also conservative, as different treatments have different adverse events, weighting across multiple tumour sites reduces the chance of a given adverse event meeting the 5% threshold.

Table 38 below outlines the AEs that occurred in ≥ 5% of patients in the treatment arm of the source documents. The base case only included AEs with a rate of 5% of higher for larotrectinib or the pooled comparator after weighting by tumour location (AEs in bold). To explore the impact of the alternative approach for AE inclusion, a scenario

analysis was conducted where the inclusion of AEs was based on unweighted rates, i.e., all AEs in Table 38 below.

Table 38. Adverse events for larotrectinib and the pooled comparator treatment (grade 3 or higher; inclusion determined by weighted rates)

Treatment emergent (Grade 3+)	AE rates	
	Larotrectinib	Pooled comparator
Abnormal liver function		[redacted]
Anaemia	[redacted]	[redacted]
Alanine aminotransferase increased		[redacted]
Anorexia		[redacted]
Diarrhoea		[redacted]
Fatigue		[redacted]
Febrile neutropenia		[redacted]
Increase alkaline phosphatase level		[redacted]
Increase creatinine level		[redacted]
Increase in total bilirubin		[redacted]
Infection		[redacted]
Leukopenia		[redacted]
Lymphocyte count decreased (lymphopenia)		[redacted]
Nausea		[redacted]
Neutropenia	[redacted]	[redacted]
Pulmonary embolism		[redacted]
Thrombocytopenia		[redacted]
Vomiting		[redacted]

AE: adverse event, Source: Larotrectinib clinical trial programme safety population (n=137); See Appendix M for comparator sources

B.3.4 Measurement and valuation of health effects

The impact of treatments on health-related quality of life (HRQoL) was also tracked throughout the model time horizon by assigning differential health state utility values (HSUVs) to progression-free disease, progressed disease, and death. Heath impact of adverse events (AEs) were incorporated as utility decrements (disutilities) per event. HRQoL results are presented as quality-adjusted life years (QALYs).

Health-related quality-of-life data from larotrectinib clinical trial programme

Health related quality-of-life data from the larotrectinib clinical trial program trials LOXO-TRK-15002 (aged 12 and older) and LOXO-TRK-15003 (aged 1 month to 21 years) were available, these data were used in a mapping exercise to generate utilities for larotrectinib in the base case cost-effectiveness analysis.

HRQoL was assessed in the LOXO-TRK-15002 trial using the EQ-5D-5L questionnaire every 8 weeks during the first year of follow-up, and every 12 weeks after one year of follow-up. In line with the NICE position statement on the use EQ-5D-3L (99), data from these trials were used in a mapping exercise to derive EQ-5D-3L utilities (using the UK value set) as described below and in more detail in Appendix N.

In the LOXO-TRK-15003 study, HRQoL was assessed using the PedsQL for all patients (the PedsQL Infant Scales [PedsQL IS] for infants aged 1-24 months, and PedsQL Generic Core Scales [PedsQL GCS] for children over 2 years old) as part of pre-treatment screening and then on day 1 of every 28-day cycle until the patient discontinues treatment. In case of disease progression and treatment continuation, HRQoL assessments were still implemented in each follow-up visit if still on treatment. These values were also mapped to EQ-5D-3L to create a pooled set of utility values.

Mapping

Derivation of Utility Values

EQ-5D-5L responses obtained in the LOXO-TRK-15002 study were used to derive utility values for patients over 12 years of age. To ensure estimates were relevant to the UK population, a crosswalk developed by Van Hout *et al.,* 2012 was used to derive

mapped utility values, as recommended by NICE for data gathered using the EQ-5D-5L (100, 101).

A targeted literature review was conducted to identify mapping algorithms that could translate assessments from the PedsQL (PedsQL IS, PedsQL GCS) to EQ-5D-3L. The search yielded a single relevant publication by Khan et al. (2014), where data from a cross-sectional survey among 11-15 year old children in four English secondary schools was used to generate a series of models to support mapping from PedsQL GCS to EQ-5D-3L (102). No publications were identified that provided mapping algorithms for transforming PedsQL IS to EQ-5D scales. Therefore, for patients with assessments obtained using the PedsQL IS version only (and not the PedsQL GCS), utility values were not derived due to the lack of an available mapping algorithm.

Patient Disposition

The total eligible study population included all subjects from the LOXO-TRK-15002 and LOXO-TRK-15003 studies who had at least one HRQoL assessment (by means of EQ-5D-5L or/and PedsQL GCS questionnaire) available. Patients from the LOXO-TRK-15003 trial who did not use the PedsQL GCS scale, namely patients age 2 and under, were removed given the lack of available mapping algorithm. Figure 31 below describes the starting patient sample, reasons for attrition and final patients and assessments included in this analysis.

Figure 31. Disposition of patients and assessments included in utility analysis

After confirming these inclusion criteria, a total of [redacted] patients were included in the sample and contributed a total of [redacted] assessments: [redacted] were paediatric patients from LOXO-TRK-15003 (20 patients) and [redacted] were adult patients included in LOXO-TRK-15002 (53 patients). The higher average number of assessments per patient in the LOXO-TRK-15003 study reflects that this study collected information each cycle (as opposed to every other cycle, as seen in 15002).

The only missing data in the final sample of included patients was information on the school functioning domain for [redacted] patients in LOXO-TRK-15003 ([redacted] patients under age 5 and [redacted] aged 12). For a total of [redacted] assessments ([redacted] progression-free, [redacted] progressed), the missing school function score was estimated by the mean of scores from the other available dimensions of the same subject. Note that all assessments were included in the analysis with the exception of the baseline assessment (i.e., cycle 1) when the effects of larotrectinib may not yet be felt by the patient.

1. Statistical analysis

Statistical analysis followed recommendations included in Technical Support Documents from NICE (103). All analyses were conducted using SAS Enterprise Guide 7.3 (SAS Institute Inc., Cary, NC).

EQ-5D-3L utility values were analysed using various regression modelling techniques (Ordinary Least Square (OLS) and Mixed Model Repeated Measures (MMRM)) to estimate the mean utility score for patients designated as either progression free (and on treatment) or progressed disease. The selection of the preferred model was based on:

- Model reflecting the repeated nature of measurements; and
- Selection based on AIC measurements.

The MMRM model was selected as it reflects the repeated nature of measurements as it accounts for the autocorrelation of patient utility values. Using this model, the potential relationship (lower variability) between HRQoL assessments reported by the same patient (i.e., responses reported by the same patient can show a lower variability than with scores from other patients) was taken into account.

Due to the small number of TEAEs these could not be reliably evaluated within the regression models as a unique covariate. Instead these are implicitly captured through the derived HSUVs. To be conservative the economic model also applies TEAE disutilities from the literature for larotrectinib, as on occasion it could not be verified whether the comparator HSUVs had been adjusted for AEs. This assumption avoids any introduction of bias in favour of larotrectinib.

EQ-5D-3L utilities for larotrectinib

Table 39 provides the trial-based EQ-5D-3L utility values applied to larotrectinib in the base case for the progression-free and progressive disease health states,.

Table 39. EQ-5D-3L utility values applied in the base case for larotrectinib

Health state	Utility value: mean (standard error)	95% confidence interval
Progression-free disease, receiving larotrectinib treatment, with or without TEAEs	[redacted]	[redacted]
Progressive disease, receiving larotrectinib treatment, with or without TEAEs	[redacted]	[redacted]

Health-related quality-of-life studies

A series of systematic literature reviews (SLRs) was completed for solid tumours that are known to harbour neurotrophic tyrosine receptor kinase (NTRK) gene fusions that were studied in the larotrectinib clinical trial programme. Studies on treatments (approved and in development) for each tumour site were identified and the available patient-reported outcome (PRO)/ health-related quality of life (HRQoL) evidence was extracted. Details of the methods used to generate relevant HRQoL evidence across tumour sites is presented in Appendix H.

Health-related quality-of-life data used in the cost-effectiveness analysis

In the base case, the larotrectinib-specific health utility values were used in the model.

The comparator-specific utilities values were taken from the following sources:

- For tumour sites with previously published NICE technology assessments, HSUVs were extracted directly from the submissions, leveraging the Committee's preferred assumptions on the values to use for the analysis. This was true for the following tumour sites: NSCLC, melanoma, colorectal, GIST, adult soft tissue sarcoma (nGIST) (also used as proxy for bone sarcoma), breast, CNS/glioma, pancreas and thyroid (93, 94, 98, 104-110).

- For cholangiocarcinoma, published health-state utility values could not be identified from the literature. Cholangiocarcinoma patients were assigned the weighted average of health state utilities for other tumour sites.

- For the remaining tumour sites (salivary, STS paediatric), targeted literature searches were conducted to identify appropriate utility information (111, 112).

The health state utility values for STS paediatric patients could not be identified and were set equivalent to the general STS population (112).

- For NSCLC, the health state utilities in the committee papers were adjusted by adverse reactions (104).However, an attempt to back-calculate was not successful, so this analysis uses the original health state utility values for progression-free from the cited study in the publicly available NICE appraisal documents (113). This approach was taken to avoid double-counting utility decrements due to adverse reactions.

- For breast cancer, the progression-free utility value was also adjusted for adverse reactions and response rates (109). Back-calculation was successful, so the value used for progression-free utility in this analysis only represents adjustment for response. Therefore, there was no double-counting when applying adverse reaction utility decrements.

A summary of the health state utility values used in the base case economic analysis is presented in Table 40.

Table 40. Summary of utility values for cost-effectiveness analysis

Tumour site	Model Health State	Utility value: mean (standard error)	95% confidence interval	Reference in submission (section and page number)	Justification
Larotrectinib	Progression-free	[redacted]	[redacted]	Section B.3.4, page 173	Trial-based EQ-5D-3L utility mapping study
	Progressed disease	[redacted]	[redacted]		
Pooled comparator	Progression-free	[redacted]		Calculation, see Appendix M	Weighted average of tumour-specific health utilities
	Progressed disease	[redacted]		Calculation, see Appendix M	Weighted average of tumour-specific health utilities

Adverse reaction disutilities

Grade 3 or 4 reported in at least 5% or more of larotrectinib patients or the pooled comparator arm were included in the model.

The model includes utility decrements for these AEs. These decrements vary by tumour site, as tumour-specific decrements were preferred. Utility decrements reported in the publicly available NICE appraisal documents and the SLR were preferred. In the absence of this data a systematic approach was taken, based on the following steps:

1. Use disutility values as reported in the committee papers by tumour site
2. Use estimates from other TAs for the same tumour site
3. Use information from a targeted literature review for the same tumour site
4. Identify a proxy from another tumour site and/or a previously used source

Utility decrements reported for the same tumour site were preferred over use of utility decrements from other tumour site or making assumption for event proxies. The utility decrements for adverse reactions for larotrectinib were assumed to be the maximum disutility for the event across all tumour sites to conservatively account for the utility decrements. Table 41 presents the AE disutilities for AEs included in the base case.

Table 41. Summary of utility decrement values for cost-effectiveness analysis

Adverse reaction	Utility decrement: mean (standard error)	95% confidence interval	Reference in submission (section and page number)	Justification
Anaemia	Larotrectinib: [redacted]	Pancreas: [redacted]	[redacted] (93)	Used highest utility decrement across tumour sites
	Cholangiocarcinoma: [redacted]		[redacted] (110)	Assumed the same decrement as colorectal
	Colorectal: [redacted]		[redacted] **(110)**	**Recommended by committee from TA405**
	Melanoma: [redacted]		[redacted] (110)	Assumed the same decrement as colorectal
	Pancreas: [redacted]		[redacted] **(93)**	**From TA440**
	STS paediatrics: [redacted])		[redacted] (114)	Assumption made in STS CEA
Neutropenia	Larotrectinib: [redacted]	Breast: [redacted]	[redacted] (113)	Used highest utility decrement across tumour sites
	Breast: [redacted]		[redacted] **(109)**	**Value is from the TA and confirmed by ERG**
	Cholangiocarcinoma: [redacted]		[redacted] (113)	Assume same as NSCLC
	STS paediatrics: [redacted]		[redacted] (114)	Assumption made in STS CEA

Abbreviations: AG: assessment group, CEA: cost-effectiveness analysis, ERG: evidence review group, NSCLC: non-small cell lung cancer, STS: soft tissue sarcoma, TA: technology assessment

Bold: Directly from the TA/committee papers of the TA used to populate the comparator data

B.3.5 Cost and healthcare resource use identification, measurement and valuation

Methodology for identifying resource use and health state costs followed methods outlined for model input collection, see section B.3.2 and Appendix I. Unit costs were sourced from national drug tariff and fee schedules and were presented in current value or inflated to 2017/18 GBP. Assumptions around healthcare resource use were explored in the clinical validation interviews – see section B.3.10.

Intervention and comparator costs and resource use

Larotrectinib drug costs

To allow use across the adult and paediatric populations, larotrectinib is available in different presentations (100mg capsules, 25mg capsules and oral solution). With a proposed per mg list price of [redacted], the proposed list price per package varies by presentation to reflect the specific dose intensity per unit and package size.

The larotrectinib modelled pooled cohort is formed of [redacted]% paediatric and [redacted]% adult patients, based on the larotrectinib clinical trial programme. The paediatric patient's treatment formulation is split across 100mg capsules, 25mg capsules and oral solution. Presentations of larotrectinib used in the economic model reflect those received in the larotrectinib clinical trial programme and are presented within Table 42.

Individual patient data from the clinical trial programme for the paediatric proportion of patients are included within the modelled engine, tracking the age of each patient in order to determine switching to adult formulation and dosing and update the proportional split of the overall cohort across all formulations.

Table 42. Larotrectinib drug cost (expected list price)

Larotrectinib formula	Formulation	Pack size	Total mg per pack	Expected cost per pack	Average dose per day (mg)	Expected cost per day	Proportion of cohort in cycle 0
Adults	100mg capsules	56 capsules	5,600	[redacted]	[redacted]	[redacted]	[redacted]
Paediatric	100mg capsules	56 capsules	5,600	[redacted]	[redacted]	[redacted]	[redacted]
	25mg capsules	56 capsules	1,400	[redacted]		[redacted]	[redacted]
	Bottle of solution	1 bottle	2,000	[redacted]		[redacted]	[redacted]

Comparator drug and administration costs

For comparator treatments, drug acquisition costs of generic compounds were sourced from the electronic market information tool (eMIT) (115), with the remainder of drug acquisition costs being sourced from the British National Formulary (BNF) (116). The least expensive cost per mg of drug was used to represent unit cost, and drug wastage was not considered for comparators in the base case. A summary of the intervention and comparator costs are presented in Table 43.

For some tumour sites in the stratified comparator arm, some drugs were administered through intravenous therapy (IV) route and were dosed according to average body surface area (BSA). The average BSA were taken as reported in the relevant NICE TA, where applicable, or a published clinical trial informing the efficacy inputs if this information was not available in the appraisal. The BSA range used in the model is [redacted]. For tumour locations where no NICE TA-based or literature-based BSA was available, the average BSA of larotrectinib patients was used. The average adult BSA of [redacted] m^2 from the larotrectinib trial was used for salivary cancer. Similarly, the average paediatric BSA of [redacted] m^2 from the larotrectinib trial was used for STS paediatric patients.

Table 43. Treatment drug cost for intervention and comparators included in the model

Drug	Dosing schedule	Admin route	Dose per treatment cycle, mg	mg per pack	Expected pack cost	Expected cost per day/ Admin[1]	Source
Larotrectinib							
Adult	Average dose of [redacted] mg per day	Oral	[redacted]	5600	[redacted]	[redacted]	
Paediatrics	Average dose of [redacted] mg per day	Oral	[redacted]	1400 - 2000	[redacted]	[redacted]	
NSCLC	No active treatment	-	-	-	-	[redacted]	(104)
Salivary							(117)
Cisplatin	80 mg/m^2 on day 1 every 21 days	IV	[redacted]	100	£52.86	[redacted]	
Vinorelbine (50mg)	25 mg/m^2 on days 1 and 8 every 21 days	IV	[redacted]	500	£38.91	[redacted]	
Melanoma							(105)
Dacarbazine	1000mg/m^2 every 21 days	IV	[redacted]	1000	£47.79	[redacted]	
Paclitaxel	175mg/m^2, every 21 days	IV	[redacted]	150	£10.48	[redacted]	
Carboplatin	AUC 5, every 21 days	IV	[redacted]	600	£17.54	[redacted]	
Temozolomide	200mg/m^2 on days 1-5 every 28 days	Oral	[redacted]	700	£20.26	[redacted]	
Palitaxel+carboplatin	Paclitaxel 175mg/m^2 and carboplatin AUC 5, every 21 days	IV	[redacted]	-	-	[redacted]	
Colorectal/Appendix	No active treatment	-	[redacted]	-	-	[redacted]	(106)
GIST	No active treatment	-	[redacted]	-	-	[redacted]	(107)
Non-GIST/Bone sarcoma	No active treatment	-	[redacted]	-	-	[redacted]	(108)
STSp/IFS							(118)
Irinotecan	50 mg/m^2 per day for 5 days at weeks 1, 4, 13, 25, 34, 46, 49.	IV	[redacted]	20	£130.00	[redacted]	
Vincristine	1.5mg/m^2 on day 1 of weeks 1, 2, 4, 5, 13, 14, 25, 26, 34, 35, 46, 47, 49, 50.	IV	[redacted]	5	£133.30	[redacted]	
Breast							(109)
Vinorelbine (IV)	30 mg/m^2 weekly for 6 months	IV	[redacted]	500	£38.91	[redacted]	

Vinorelbine (Oral)	60mg/m^2 weekly for first 3 administrations, 80mg/m^2 weekly for subsequent administrations until progression or for a maximum of 6 months	Oral	[redacted]	20	£43.98	[redacted]	
Gemcitabine	1250mg/m^2 two times per 21-day cycle for 6 months	IV	[redacted]	2000	£16.01	[redacted]	
Doxetaxel	100mg/m^2 once per 21-day cycle for 6 months	IV	[redacted]	80	£11.95	[redacted]	
Paclitaxel	175mg/m^2 once per 21-day cycle for 6 months	IV	[redacted]	150	£10.48	[redacted]	
Doxorubicin	75mg/m^2 once per 21-day cycle for 6 months	IV	[redacted]	200	£16.80	[redacted]	
Cholangiocarcinoma							(98)
Gemcitabine (2000mg)	1000 mg/m^2 on days 1 and 8 every 3 weeks	IV	[redacted]	2000	£16.01	[redacted]	
Cisplatin	25 mg/m^2 on days 1 and 8 every 3 weeks	IV	[redacted]	100	£52.86	[redacted]	
Glioma/CNS							(97)
Lomustine	110mg/m^2 day on 1 every 6 weeks	Oral	[redacted]	800	£780.82	[redacted]	
Pancreas							(93)
5-fluorouracil	2000mg/m^2 administered 4 times over 6 weeks	IV	[redacted]	5000	£3.30	[redacted]	
Leucorovin	200mg/m^2 administered 4 times over 6 weeks	IV	[redacted]	1000	£17.55	[redacted]	
Thyroid	No active treatment	-	[redacted]	-	-	[redacted]	(94)

[1] Treatment drug costs for all tumour locations are costs per day except for STSp/IFS, for which costs per administration are presented given the irregular dosing schedule.

Administration costs for comparators were calculated based on the administration procedure(s) required in each treatment cycle and the number of administrations. Drugs administered orally were assumed to incur no administration cost. The reference cost for simple chemotherapy delivery was used for all IV chemotherapies (88). Drugs were assumed to require an administration procedure every treatment cycle. A simple parenteral chemotherapy administration was applied as it was most commonly used in the TAs (93, 105, 108, 109) and was a conservative approach for estimating comparator costs. Procedural code to source National Health Service (NHS) reference costs for a simple chemotherapy administration is presented in Table 44

Table 44. Drug administration costs

Administration type	Code	Unit cost
Deliver simple parenteral chemotherapy at first attendance	SB12Z	£228.99
Oral therapies are assumed to be associated with no administration cost.		

Table 45 provides the calculated total treatment cost per modelling cycle for larotrectinib and the comparator. When the comparator treatment featured a mix of drugs, the drug and administration costs of each component drug were weighted according to the distribution available in the source documents.

Table 45. Treatment cost (drug + administration) per cycle (week) by tumour location

Drug	Cost per cycle (week)
Larotrectinib (expected list price)	[redacted]
Comparators with no active treatment	
NSCLC	[redacted]
Colorectal/Appendix	[redacted]
GIST	[redacted]
Thyroid anaplastic, follicular and papillary	[redacted]
Non-GIST/Bone sarcoma	[redacted]
Active treatments	
Melanoma	[redacted]
Breast	[redacted]
Gliomas	[redacted]
Pancreas	[redacted]
Salivary	[redacted]
STS paediatric/IFS	[redacted]
Cholangiocarcinoma	[redacted]

[1] Price shown is based on the baseline split between adult and paediatric patients
[2] STS paediatric treatment dosing is irregular from week-to-week. See Table 43 for further details on dosing
For tumour locations with no active treatment, both treatment arms receive current standard management and larotrectinib is an add-on therapy. Thus, comparator arm treatment cost is £0. For tumour locations where active treatments are used as proxies, costs of the specific active treatments are calculated.
NSCLC, non-small-cell lung cancer; GIST, gastrointestinal stromal tumours; STS, soft tissue sarcoma; IFS, infantile fibrosarcoma.

Treatment duration

In the base case, larotrectinib treatment was assumed to continue until disease progression. A scenario explored the application of time to treatment discontinuation (TTD) curve from the larotrectinib clinical trial programme.

Where the comparator data source was a previous NICE technology appraisal, the methods for calculating treatment duration were adopted. Where a clinical publication was identified, relevant data on treatment duration was extracted, and followed either a fixed treatment schedule, or point estimate (as published). In the case of a maximum treatment duration, treatment costs were capped by the fixed schedule or the maximum duration for patients who have not progressed. These tumour locations are: salivary cancer ([redacted] 21-day cycles) (117), STS paediatrics ([redacted] days, based on the fixed treatment schedule which spanned over

50 weeks) (118), breast cancer ([redacted] months) (109), and cholangiocarcinoma (maximum treatment duration of [redacted] weeks) (98).

Health-state unit costs and resource use

A series of systematic literature reviews (SLRs) was completed for a multitude of solid tumours that are known to harbour neurotrophic tyrosine receptor kinase (NTRK) gene fusions. Studies on treatments for each tumour site were identified and the available clinical, economic, and patient-reported outcome (PRO)/health-related quality of life (HRQoL) evidence was extracted. Details of the methods used to generate relevant cost and resource utilisation evidence across tumour sites is presented in Appendix I.

Larotrectinib

There are no published estimates of healthcare resource use for the patients with TRK Fusion cancer. Given the lack of UK clinical experience outside of a clinical trial setting for treatments for TRK-Fusion cancer (and histology independent treatments in general), primary research would have not been able to adequately inform health care resource use for the population enrolled in the trial.

Health state costs for larotrectinib were assumed equal to the weighted average of the comparators costs, using the tumour site distribution in the larotrectinib clinical trial (section B.3.2).

This approach was validated by UK clinicians interviewed as part of the clinical validation (section B.3.10). All clinicians interviewed considered this an appropriate assumption given the data available, and expected this would likely be conservative, and overestimate health care resource use for larotrectinib.

Comparators

For the comparator arm, as per the other model inputs, healthcare resource use was modelled independently for each tumour site. Where a NICE TA was available, the approach selected was to use the HCRU inputs used to inform the Committee's preferred assumptions.

Data collection for HCRU inputs for the tumour locations without a NICE TA was based on the SLR output where possible and otherwise broader targeted searches were

conducted for published articles, where no evidence was found in the SLR. Data sources used for each tumour site are presented in Table 46.

Table 46. Data source used for HCRU

Type of data source	Tumour location
NICE TA (committee recommendation)	STS GIST, STS non-GIST/Bone sarcoma, Thyroid, Colorectal/Appendix, Salivary, NSCLC, Breast, Melanoma, Pancreas, Glioma/CNS
SLR	None
TLR	STS paediatrics/IFS
Weighted average of comparators with available data	Cholangiocarcinoma
NICE, National Institute for Health and Care Excellence; TA, technology appraisal; SLR, systematic literature review; TLR, targeted literature review; NSCLC, non-small-cell lung cancer; GIST, gastrointestinal stromal tumours; STS, soft tissue sarcoma; IFS, infantile fibrosarcoma	

The cost of the progression-free or progressive disease health states could take the form of a one-off cost at state initiation (start-up costs) and recurrent costs per cycle. The cost of death was applied as a one-off lump-sum.

Aggregated costs with a paucity of information around individual components or costs from alternative sources were unable to be updated from national databases. In these cases, costs were inflated using the PSSRU reported inflation indices. Based on PSSRU recommendations, the ‘New Health Services Index using Consumer Price Index’ values were used to inflate prices between 2014-2018 with the ‘Hospital and Community Health Services Index' (HCHS) prior to 2014.

Where a source for a tumour site provided an aggregate cost, or HCRU details for start-up costs, these were implemented in the model. If the source did not present a start-up cost, start-up costs were assumed null, remaining consistent with the replication of methodology used within the specific TA or evidence source.

All tumours except for glioma/CNS and cholangiocarcinoma reported per cycle health state costs or detailed HCRU. The glioma NICE TA (110), only reported a magnetic resonance imaging (MRI) procedure at baseline, after 2 treatment cycles, and at 6-month follow-up. While this was cost spread over time, it could not be implemented on

a per-cycle basis. Thus, the model calculated the total cost of 3 MRIs and applied it as a one-off cost to glioma.

No HCRU source was identified for cholangiocarcinoma . Therefore, the health state costs for cholangiocarcinoma were based on a weighted average of all other tumour sites

End of life costs could not be identified for glioma, STS non-GIST, STS paediatric, and cholangiocarcinoma patients. The targeted literature review identified a modelling study for end of life cost among cancer patients, which was used to inform this input (119). The health care cost and social care cost from this modelling study were included to align with the resources accounted for in other TAs for consistency. Because the larotrectinib HCRU was the weighted average of all comparators, absolute values for comparator tumour location was not expected to impact the incremental result between intervention and the comparators.

Table 47 lists the resource use associated with the comparator arm and the unit costs. Table 48 shows the calculated health state costs for the tumour locations. Detailed data on resources applicable to each tumour location and frequency of use within the health states are presented in Appendix M.

Table 47. Healthcare resource use components and associated unit costs

Healthcare resource	Details	Unit cost	Unit cost source
Outpatient/inpatient visits			
Oncologist visit	Code 370 WF01A Follow Up Attendance	£160.00	(88)
GP visit (home/surgery)	10.3b GP: per surgery consultation lasting 9.22 minutes	£37.00	(120)
Community nurse/health specialist visit	Weighted average of community-based nurses band 5 to 8a	£80.75	(120)
Plastic surgeon visit	Plastic surgery code 160. Total outpatient attendances	£107.00	(88)
Dental visit	Dental medicine specialties; code 450; total outpatient attendance	£122.00	(88)
Depression management	Occupational therapist, adult, one to one; A06A1	£81.00	(88)
Nutritional supportive care visit	Specialist Nursing, Enteral Feeding Nursing Services, Adult, Face to face; N16AF	£110.00	(88)
Speech therapy visit	Speech and Language Therapist, Adult, One to One; A13A1	£96.00	(88)
Diagnostic tests			
CT scan (one area)	Computerised Tomography Scan of One Area, with Contrast RD22Z	£132.00	(88)
CT scan (three areas)	Computerised Tomography Scan of Three Areas, with Contrast RD26Z	£130.00	(88)
MRI scan	Weighted average of all MRI codes (RD01A to RD07Z)	£145.72	(88)
Ultrasound	Ultrasound Scan with duration of 20 minutes and over, with Contrast; RD43Z	£47.46	(88)
Full blood count	Haematology; DAPS05	£2.51	(88)
Liver function test	Clinical biochemistry; DAPS04	£1.11	(88)
Bone scan	Nuclear Bone Scan of Two or Three Phases, 19 years and over; RN15A	£226.85	(88)
ECG	Simple Echocardiogram, 19 years and over Direct Access; RD15A	£65.18	(88)
Chest X-ray	Direct Access Plain Film (DAPF)	£31.49	(88)
Total protein	Clinical biochemistry; DAPS04	£1.11	(88)
Urinalysis	Microbiology; DAPS07	£7.59	(88)
Clinical/laboratory test	Clinical biochemistry; DAPS04	£1.11	(88)
Coagulation panel (PT/PT-INR, PTT)	Haematology; DAPS05	£2.51	(88)
Haematologic growth factor transfusions (first cycle)	NICE guideline NG24 Blood Transfusion (2015)	£170.14	(121)

Haematologic growth factor transfusions (subsequent cycle)	NICE guideline NG24 Blood Transfusion (2015)	£162.00	(121)
Pain management	Calculation (total of combination drugs below)	£104.00	(88)
Co-codamol 8/500 caplets	Co-codamol 8/500 caplets - 30 pack - cost per tablet - 8 tablets per day	£0.03	(122)
Tramadol 50mg capsules	Tramadol 50mg capsules - 100 pack - cost per tablet - 8 tablets per day	£0.03	(122)
Paracetamol 500mg caplets	Paracetamol 500mg caplets - 100 pack - cost per tablet - 8 tablets per day	£0.02	(122)
Morphine sulphate 10mg I/R	Morphine sulphate 10mg I/R - 56 pack - cost per tablet - 18 tablets per day	£0.09	(116)
Dexamethasone 2mg tabs	Dexamethasone 2mg tabs - 50 pack - cost per tablet - 2 tablets per day	£0.86	(122)
End of life care			
Palliative resection	Weighted average of costs of single intervention for malignant GI tract disorder (code: FD11D, FD11E, FD11F)	£3,844.47	(88)
Palliative radiotherapy	Weighted average of adult medical specialist palliative care attendance costs (code: SD01A, SD02A, SD03A, SD04A)	£150.92	(88)
Terminal care inpatient	Respiratory Neoplasms without Interventions, with CC Score 13+; DZ17S	£3,051.42	(88)
Terminal care hospice	Assumed 25% increase on hospital inpatient care	£3,814.28	NICE TA374 (104)
Terminal care hospice/palliative unit	Hospital Specialist Palliative Care Support, 19 years and over; SD03A	£117.84	(88)
Excess bed day	Non-elective excess bed days	£429.45	(88)
Macmillan nurse	Assumed 66.7% of community nurse cost	£53.83	NICE TA374 (104)
Drugs and equipment	Marie Curie report figure increased for inflation	£240.00	(111, 120)
GP, general practitioner; CT, computed tomography; MRI, magnetic resonance imaging; ECG, electrocardiogram; PT, prothrombin time; INR, international normalized ratio; PTT, partial thromboplastin time			

Table 48. Health state costs by tumour location

Tumour locations	Progression-free, start-up	Progression-free, per cycle	Progressed, start-up	Progressed, per cycle	Death/End-of-life
Larotrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Comparators with no active treatment					
NSCLC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Colorectal/Appendix	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GIST	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Thyroid anaplastic, follicular and papillary	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Active treatments accepted as a positioned last-line comparator					
Non-GIST/Bone sarcoma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Melanoma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Breast	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Glioma	[redacted]				[redacted]
Pancreas	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Salivary	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
STS paediatric/IFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cholangiocarcinoma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Health state costs are based on the source NICE TA or literature.
Start-up cost is the one-time cost of health resources required for assessment and/or treatment initiation when patients enter a health state. Start-up cost is assumed £0 if the source does not mention any HRU details or aggregate health state cost.
Glioma TA reported monitoring cost over the treatment period by a fixed schedule that did not fit a per-cycle calculation. Thus, the total costs were applied as a one-off cost to glioma health states.
Round 2015 was used to inform end-of-life cost for tumour locations that did not have this data in the TA or literature sources.
NSCLC, non-small-cell lung cancer; GIST, gastrointestinal stromal tumours; STS, soft tissue sarcoma; IFS, infantile fibrosarcoma.

Adverse reaction unit costs and resource use

The cost of treating an adverse event was assumed not to vary based on the patient's tumour site. This approach has been applied in previous NICE TAs, where the cost of treating AEs were based on reported costs for other tumour locations (106).

Healthcare resource group (HRG) codes were used for adverse events. When multiple codes for the same adverse reaction were identified, the codes were weighted using the 'Activity' information from the NHS Reference Costs (88). If the evidence source only provided HCRU details rather than the codes, the HCRU terms were searched in the NHS reference schedule to identify relevant codes.

In the event that the evidence source estimated the AE cost based on activity codes that were no longer used or there was no information on HCRU or cost at all, assumptions were made either for HCRU (then codes are identified) or to equate to the cost of another adverse event that was considered to have similar cost impact, in the cases where applicable methods are reported in Appendix M.

In regard to the costs included in the economic model base case, the HRG codes for anaemia were based on the ERG report of TA405 because this TA provided the most comprehensive information on coding and represented an ERG perspective (106).

Neutropenia was costed in a previous TA using the code XD25Z (93, 106, 109), with a cost of lower than £200. Specifically, the ERG report of the colorectal cancer TA405 recommended not to include hospital stays when estimating the cost of neutropenia (106). However, code XD25Z is no longer available in the NHS cost reference. We assumed this would be represented by an outpatient visit (code 300), which was associated with a similar cost to that of XD25Z.

Table 49 provides the details for AE costs included in the economic model base case.

Table 49. NHS reference costs of adverse reactions

Adverse event	Cost per event	Coding details	Source
Anaemia	[redacted]	[redacted]	NICE TA405 (106)
Neutropenia	[redacted]	[redacted]	NHS reference costs (88)

B.3.6 Summary of base-case analysis inputs and assumptions

B.3.6.1 Summary of base-case analysis inputs

A summary of inputs used within the model cross-referenced to the detailed sections above outlining the processes for identifying each specific input and the source is presented in Table 50. A full list of larotrectinib and comparator-specific inputs that inform the base case analysis are further detailed in Appendix M.

Table 50. Inputs summary and global inputs table

Variable	Input	Detailed section	Source
Model characteristics			
Discount rate - costs and outcomes	3.50%	-	NICE reference case
Mean body surface area (Where used to inform dosing)	Larotrectinib adults: [redacted] Larotrectinib paediatrics: [redacted] Breast: [redacted] Cholangiocarcinoma: [redacted] CNS/glioma: [redacted] Pancreas: [redacted]	For Larotrectinib BSA: Section B.3.5	Larotrectinib: Larotrectinib clinical trial programme Breast: TA423(109) Cholangiocarcinoma: Roth et al., 2012 (123) CNS/glioma: TA23 (110) Pancreas: TA440(93)
Clinical inputs			
Treatment duration	Time (days)	Section B.3.3.	
Adverse events	Larotrectinib clinical trial programme, pivotal comparator studies	Section B.3.3.	
Progression-free and overall survival	Larotrectinib clinical trial programme, pivotal comparator studies	Section B.3.3.	
Health-related quality-of-life			
Health state utilities and AE disutilities	Larotrectinib clinical trial programme, pivotal comparator studies	Section B.3.4.	
Costs			
Expected list price of larotrectinib per pack; - (100mg x 56 capsules)	[redacted]	Section B.3.5.	Bayer

Variable	Input	Detailed section	Source
- (25mg x 56 capsules) - Oral solution (1 x 2000mg)	[redacted]		
Drug acquisition costs	Branded comparators from BNF; generics from eMIT	Section B.3.5.	(115, 116)
Drug administration costs per therapy	NHS reference costs; PSSRU	Section B.3.5.	(88, 120)
Health state resource costs	NHS reference costs; PSSRU	Section B.3.5.	(88, 120)
Adverse event costs	NHS reference costs; PSSRU	Section B.3.5.	(88, 120)
Other variables			
Time horizon	Lifetime; 40 years adults, 80 years paediatrics	Section B.3.3.	
Dose	**Larotrectinib:** [redacted]mg per day for adults [redacted]mg per day for paediatrics **Comparators:** Based on TA/pivotal study dosing	Section B.3.5.	Larotrectinib clinical trial programme
Perspective	National Health Service and Personal and Social Services perspective	-	NICE reference case
Willingness-to-pay threshold	£50,000	-	WTP threshold for EOL therapies
Comparator weighting	Larotrectinib clinical trial programme	Section B.3.3.	

B.3.6.2 Base-case assumptions

Base-case model assumptions are outlined in Table 51 below.

Table 51. Base-case model assumptions

Data challenges	Mitigation method/Assumptions	Justification
Treatment duration/discontinuation data was not always available for the comparator, and varied in how it was assessed/modelled.	Comparators with treatment duration data followed the specified schedule or approach adopted in the identified NICE TA, otherwise the treatment data was extracted from the clinical publication, and used a fixed schedule or point estimate. Where a fixed treatment schedule was used treatment was incorporated capped prior to progression.	Where available past NICE appraisals are used to inform decision making, or published estimates to align with clinical evidence identified. The assumption of treat-to-progression or capping treatment at progression reflects clinical practice in England.
Multiple sources for health state resource use and cost were identified for the comparator treatments, with varied approaches to data collection and/or reporting.	Tumour sites without end-of-life cost data were assigned with a value from a modelling study that aimed to quantify end-of-life cost among cancer patients (119)	This avoids bias in accounting for end-of-life cost and ensures a balanced cost input for the pooled comparator arm.
	The health state costs for larotrectinib were equated to the weighted average of comparator HCRU (based on the larotrectinib clinical trial programme)	Clinical experts advised this was an appropriate assumption given available evidence and would likely be conservative, over-estimating costs for larotrectinib.
Data is unavailable to understand the timing and duration of AEs for larotrectinib and comparators.	Impact of adverse events are modelled as a one-time upfront cost/disutility instead of a cumulative effect over time.	This removes the need for complicated and/or impossible to justify assumptions for temporality of AE impact by tumour site, and this approach has been used in past NICE submissions in oncology.

Data challenges	Mitigation method/Assumptions	Justification
Resource use/cost and disutility associated with AEs were not always available and were reported in different ways across the tumour locations.	COSTS: The NHS reference schedules were reviewed to identify the relevant codes for each AE. The reference costs of all relevant codes were weight averaged by the activity data to derive the costs per event.	Facilitates the use of publicly available data and has been used in past NICE submissions in oncology.
	DISUTILITY: NICE submissions for oncology indications were reviewed for AE disutilities. Selection of final disutilities prioritised data from past NICE appraisals, supplemented with published literature and assumptions supported by clinical opinion.	Facilitates use of tumour-specific values where possible and is consistent with submissions to NICE in oncology.
PFS was unavailable for one of the comparators in the model due to a lack of reporting of KM curves (salivary PFS).	An exponential distribution was assumed using the median PFS from the clinical study	This was the best available data to use and ensures consistency between PFS and OS, based on the same cohort. The exponential distribution is the only distribution that enables back-calculation into a curve based on the median survival point estimate, and tends to be a conservative selection and ensures consistency between the PFS and OS for this patient cohort.

AE, adverse event; KM, Kaplan-Meier; OS, overall survival; PFS, progression-free survival; TTD, time to discontinuation

B.3.6.3 Scenario analyses

To assess the impact of the base-case assumptions on the model findings, a number of key scenario analyses were conducted. These are described in Table 52 below.

Table 52. Conducted scenario analyses

Scenario analysis	Scenario description	Justification
Discount rate	Replace 3.5% discount rates for cost and outcomes with 1.5% rate	Investigate the long term uncertainty and impact of discounting
Utility	Replace larotrectinib utilities with alternative utility model: Revised patient pool excluding paediatrics under age 11	Investigate the uncertainty surrounding the utility values derived from the small patient numbers
Drug costs	For adults, base case will use actual trial dose, and scenario will test the full daily dose ([redacted] will be cost out as 200 mg)	Investigate the impact of 100% adherence to treatment dose
	Use of larotrectinib TTD curves	To test the impact of alternative treatment assumptions
Time horizon	10 years, 20 years	Investigate impact of using shorter time horizon
Health state costs	Replace tumour location specific health state costs with consistent costs for every tumour location	Investigate the impact of the inconsistency and uncertainty of health state costs across tumour locations
	Remove health state costs if not reported in the source documents	Investigate the outcomes if model follows the original sources exactly instead of making assumptions to fill data gaps
Survival	Different comparator and larotrectinib survival curves where possible (PFS, OS)	Investigate the uncertainty and sensitivity of alternative parametric fits to survival curves
	Alternative comparator survival data for STS non-GIST; pazopanib (following clinical validation)	

Scenario analysis	Scenario description	Justification
AEs	Alternative AE inclusion criteria; all AE with individual 5% rates reported in source publication	Investigate the uncertainty of adverse event rates for the pooled comparator.
NTRK prognosis	Results from the SLR conducted to consider evidence on NTRK prognosis	Used to explore how a prognostic effect of being NTRK positive may affect CE results.
Alternative modelling methods	Stratified responder/non-responder analysis, with non-responder representing the comparator arm	Investigate the uncertainty of the overall results using alternative survival modelling methods to represent efficacy.
	Use of GMI as relative risk applied to larotrectinib health outcomes to represent a previous line of therapy comparator. See section B.2.6.	

BSA, body surface area; GMI, growth modulation index; mg, milligram; OS, overall survival; PFS, progression-free survival

B.3.7 Base-case incremental cost-effectiveness analysis results

Larotrectinib was associated with higher LYs and QALYs compared to the pooled comparator (vs. for LYs, vs. for QALYs). This translated into an additional LYs and QALYs for larotrectinib versus the pooled comparator. Total costs in the base-case were higher with larotrectinib versus the pooled comparator vs) with an incremental cost of .

The incremental results for costs and health effects indicate that at the expected list price larotrectinib was associated with a cost per QALY of . Detailed results are presented in Table 53.

Table 53. Basecase cost-effectiveness results

Source of results	Larotrectinib	Comparators	Incremental
Treatment cost			
Routine care costs			
Adverse event			
End of life care			
Total costs			
Progression-free life years			
Progressed disease life years			
Total life years			
Progression-free QALYs			
Progressed disease QALYs			
Adverse events			
Total QALYs			
ICER			

B.3.8 Sensitivity analyses

Deterministic sensitivity analysis

In order to assess the uncertainty around the results, the model includes DSA whereby parameters are iteratively varied. The results of the DSA are presented using a tornado diagram. The parameters varied in the DSA are summarised in Appendix M.

Deterministic sensitivity analysis shows the cost-effectiveness of larotrectinib to be relatively stable when key parameters are varied across their standard error/reported upper and lower ranges.

The cost-effectiveness of larotrectinib is most sensitive to the scale and shape parameters used to extrapolate overall survival for larotrectinib, this not surprising given the immaturity and low event numbers as discussed in previous sections. The scale parameter used in the base case extrapolation of overall survival in the comparator arm for STS paediatric patients was the only other input that when varied led to a change in the ICER of >£2,500. Both of these inputs are assessed further with additional assumptions in section B.3.8

The top 20 parameters that the ICER was most sensitive to are presented within Table 54 and the tornado diagram in Figure 32 below.

Table 54. Deterministic sensitivity analysis

Variable	Lower bound ICER	Upper bound ICER
Base case	[redacted]	
OS weibull shape (p) - Larotrectinib	[redacted]	[redacted]
OS weibull scale (lambda) - Larotrectinib	[redacted]	[redacted]
OS log-normal scale (mu) - STS paediatrics	[redacted]	[redacted]
OS log-normal scale (mu) - CNS	[redacted]	[redacted]
Complete response utility - STS paediatrics	[redacted]	[redacted]
OS log-normal shape (sigma) - STS paediatrics	[redacted]	[redacted]
Progressed disease utility - Thyroid follicular and papillary	[redacted]	[redacted]
Progressed disease health state cost - Larotrectinib	[redacted]	[redacted]
PFS log-normal scale (mu) - STS paediatrics	[redacted]	[redacted]
Complete response utility – Salivary	[redacted]	[redacted]
Model mixed cohort start age (years)	[redacted]	[redacted]
Model adult start age (years)	[redacted]	[redacted]
Progressed disease utility - STS paediatrics	[redacted]	[redacted]
Progression free health state cost - Larotrectinib	[redacted]	[redacted]
PFS log-normal shape (sigma) - STS paediatrics	[redacted]	[redacted]
OS log-normal shape (sigma) - CNS	[redacted]	[redacted]
Death health state start cost - Larotrectinib	[redacted]	[redacted]
Complete response utility - Thyroid follicular and papillary	[redacted]	[redacted]
Progressed disease utility - CNS	[redacted]	[redacted]
Progressed disease utility - STS adults (non-GIST)	[redacted]	[redacted]

Figure 32. DSA Tornado diagram

Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) was performed to account for multivariate and stochastic uncertainty in the model. The uncertainties in the individual parameters for treatment effect, costs, and utilities were characterised using probability distributions and analysed using a Monte Carlo simulation using 1,000 simulations.

The following groups of parameter values were included in the PSA:

- Model characteristics (discount rate, time horizon, age)
- Parametric survival models
- Adverse event costs, disutilities
- Health state utilities
- Health state costs

Appendix M presents the specific parameters varied in the PSA. Disutilities, survival parameters, health state costs were assumed to follow a normal distribution. Utilities were assumed to follow a Beta distribution (124).

The PSA results produce mean values similar to those presented within the base case analysis (presented in Table 55), providing confidence in the base case results. However, the results show a large dispersion of the 1,000 individual iterations (convergence of health outcomes, costs and ICERs presented in Figure 33) and the calculated 95% confidence intervals, presented within the cost-effectiveness plane (Figure 35).

Figure 33. PSA convergence, QALYs, costs and ICER

[redacted]

Simultaneously varying all inputs across the larotrectinib and the 12 tumour site engines leads to a large potential range in costs and health outcomes between each iteration. Results show a large spread in costs and health outcomes. However the mean incremental costs, health outcomes and ICER converges closely with the base case results.

This suggests that whilst there are influential parameters (identified through deterministic sensitivity analysis) much of the variation in the economic analysis may be explained through the structure employed in the model. Multiple parameters can appear uncertain (through cumulative standard error), however estimates informed by multiple parameters may benefit from the increased accuracy that specific sources provide.

At the expected list price, the probability of larotrectinib being cost-effective at a £50,000 per QALY is [redacted]%. The probability of cost-effectiveness at different

willingness-to-pay thresholds is presented in Table 56 and graphically in a cost-effectiveness acceptability curve (CEAC) in Figure 34.

Table 55. Probabilistic sensitivity analysis results: Expected list price

Technologies	Total mean costs (£)	Total mean LYG	Total mean QALYs	Mean incremental costs (£)	Mean incremental LYG	Mean incremental QALYs	Mean ICER (cost/QALY)
Larotrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pooled comparator	[redacted]	[redacted]	[redacted]				
Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years							

Table 56. Probability of larotrectinib cost-effectiveness

Willingness to pay	Percentage cost-effective
£0	[redacted]
£50,000	
£100,000	
£150,000	
£200,000	
£250,000	
£300,000	
£350,000	
£400,000	
£450,000	
£500,000	

Figure 34. Larotrectinib cost-effectiveness acceptability curve (CEAC)

[redacted]

Figure 35. Cost-effectiveness plane

Summary of scenario analysis results

A number of scenario analyses were conducted to test structural and input assumptions in the model. The results are shown below in Table 57.

Results from the scenario analyses show larotrectinib to be associated with a similar degree of upwards and downwards uncertainty.

- The use of alternative survival functions for larotrectinib show a wider range of uncertainty, this is not surprising given the immaturity of survival data. For PFS [redacted]per/QALY), and OS [redacted]per/QALY) there is both upwards and downwards uncertainty. For OS there is a considerable amount of downwards uncertainty.
- The use of alternative survival models for the comparator arm for PFS and OS results in a small change in the ICER [redacted]per/QALY). These scenarios suggest the base case ICER is robust to alternative comparator assumptions and potentially data sources.

- Alternative modelling (versus responder, and versus prior line of therapy) are employed to explore methods for comparative assessment, each of these approaches, whilst limited are conducted independently, all result in relatively small changes to the base case ICER (below the range of [redacted] per/QALY)
- A scenario assessing the prognostic nature of NTRK is presented based on the results of an SLR (See Appendix D). This approach has limitations, and it is expected more information will be available to inform such an analysis in the future. The naïve analysis resulted in a reduction in the ICER to between £[redacted] per/QALY, dependent on the assumptions employed.
- In scenarios a TTD approach to treatment costing was explored and was found to be consistent with the base case when using the best fitting curve (£[redacted] per QALY). Using the exponential (the only other plausible model where treatment <1% after 80 years) the ICER drops to (£[redacted] per QALY)

Table 57. Scenario analyses results

Scenario number	Scenario category	Description	Larotrectinib			Pooled comparator			ICER
			Costs	QALYs	Life years	Costs	QALYs	Life years	
0		Base case results	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
1	Discount rate	Replace 3.5% discount rates for cost and outcomes with 1.5% rate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2	Utility	Replace larotrectinib utilities with alternative utility model: Revised patient pool excluding paediatrics under age 11	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
3	Drug costs	Full daily dose for larotrectinib adults (200mg)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
4		Larotrectinib time-to-discontinuation curve for time on treatment (Weibull)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
5		Larotrectinib time-to-discontinuation curve for time on treatment (Exponential)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
6	Time horizon	10 year time horizon	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
7		20 year time horizon	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario number	Scenario category	Description	Larotrectinib			Pooled comparator			ICER
			Costs	QALYs	Life years	Costs	QALYs	Life years	
8	Health state costs	Replace tumour location specific health state costs with consistent costs for every tumour location; weighted average of all tumour location sources	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
9		Remove health state costs if not reported in the source documents for each tumour location	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
10	Adverse events	Alternative AE inclusion criteria; all AE with individual 5% rates reported in source publication	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
11	Non-GIST survival source	Use survival data from alternative source (pazopanib)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12	Survival; alternative fits	Larotrectinib OS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
13		Larotrectinib OS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
14		Larotrectinib OS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
15		Larotrectinib OS - Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
16		Larotrectinib OS - Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario number	Scenario category	Description	Larotrectinib			Pooled comparator			ICER
			Costs	QALYs	Life years	Costs	QALYs	Life years	
17		Larotrectinib PFS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
18		Larotrectinib PFS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
19		Larotrectinib PFS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
20		Larotrectinib PFS - Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
21		Larotrectinib PFS - Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
22		Salivary OS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
23		Salivary OS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
24		Salivary OS - Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
25		Salivary OS - Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
26		Melanoma OS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
27		Melanoma OS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
28		Melanoma OS - Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario number	Scenario category	Description	Larotrectinib			Pooled comparator			ICER
			Costs	QALYs	Life years	Costs	QALYs	Life years	
29		Melanoma OS - Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
30		Colorectal OS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
31		Colorectal OS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
32		Colorectal OS - Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
33		Colorectal OS - Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
34		Colorectal PFS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
35		Colorectal PFS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
36		Colorectal PFS - Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
37		Colorectal PFS - Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
38		STS GIST OS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
39		STS GIST OS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
40		STS GIST OS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario number	Scenario category	Description	Larotrectinib			Pooled comparator			ICER
			Costs	QALYs	Life years	Costs	QALYs	Life years	
41		STS GIST OS - Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
42		STS GIST PFS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
43		STS GIST PFS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
44		STS GIST PFS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
45		STS GIST PFS - Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
46		STS non-GIST OS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
47		STS non-GIST OS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
48		STS paediatrics OS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
49		STS paediatrics OS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
50		STS paediatrics OS - Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
51		STS paediatrics PFS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
52		STS paediatrics PFS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario number	Scenario category	Description	Larotrectinib			Pooled comparator			ICER
			Costs	QALYs	Life years	Costs	QALYs	Life years	
53		STS paediatrics PFS - Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
54		Cholangiocarcinoma OS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
55		Cholangiocarcinoma OS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
56		Cholangiocarcinoma OS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
57		Cholangiocarcinoma OS - Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
58		Cholangiocarcinoma PFS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
59		Cholangiocarcinoma PFS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
60		Cholangiocarcinoma PFS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
61		Cholangiocarcinoma PFS - Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
62		CNS/Glioma OS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
63		CNS/Glioma OS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
64		CNS/Glioma OS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario number	Scenario category	Description	Larotrectinib			Pooled comparator			ICER
			Costs	QALYs	Life years	Costs	QALYs	Life years	
65		CNS/Glioma OS - Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
66		CNS/Glioma PFS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
67		CNS/Glioma PFS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
68		CNS/Glioma PFS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
69		CNS/Glioma PFS - Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario analysis: Alternative modelling methods

In response to discussions held during the NICE scoping phase of the appraisal, alternative approaches for controlling for the larotrectinib clinical trial data have been explored. The merits of these approaches are discussed in section B.3.2. Each of these approaches relies on a number of assumptions and are intended to be exploratory.

Non-responder control analysis

This scenario leveraged the results from the responder/non-responder stratified survival analysis of patients in the larotrectinib clinical trial programme outlined in Appendix L.1.4.

This analysis considers those patients in the trial who did not respond to therapy, to be representative of patients who did not receive active therapy. This approach has previously been used in economic evaluations, however no previous examples were identified for histology independent treatments or basket trials (125). There are inherent limitations with this analysis, and strong assumptions needed to be made to incorporate the analysis into the model:

- Low numbers of events, especially important as the population was stratified by all patients and non-responders this substantially reduces the confidence in the overall survival analysis
- Uncertainty in the projected survival curves given the relatively short, variable follow-up in the larotrectinib clinical trial programme
- The differences in the distribution of tumour sites/disease severity between responders and non-responders could not be accounted for.
- The assumption that the non-responders would represent a control arm. Patients on larotrectinib may not respond for a variety of reasons and may be inherently different to those patients that do respond.

Methodology

The full clinical trial cohort, including responders and non-responders to larotrectinib, were applied to the larotrectinib arm while outcomes for non-responders (stable or progressive disease) alone were applied to the comparator arm.

The larotrectinib arm remained consistent with the base case, using the Weibull fit model for PFS and OS. A single model (Weibull) was used for non-responders with response status included as a covariate. The Weibull model was selected for PFS and OS based on clinical plausibility and to keep assumptions (where possible) as per the base case. See Appendix L.1.4 for full details on the survival analysis conducted.

This analysis kept all remaining base case assumptions constant (e.g. utilities, health state costs, intervention costs, and AE rates) except for the use of the non-responder survival data.

Results of the non-responder analysis show the cost per QALY decreased, driven by a small decrease in survival outcomes for the non-responder defined comparator arm compared to the pooled comparator arm from the base case (Table 58).

The survival and cost results for the comparator arm shifted minimally from the base case, suggesting that larotrectinib non-responders could represent patients on current standard management. Total LYs for the larotrectinib arm were equal to the base case at [redacted] years, with QALYs remaining at [redacted]. As such, the responder/non-responder scenario produced an ICER of £[redacted], slightly lower than the base case ICER of £[redacted].

Table 58. Alternative modelling methods: using non-responding patients as a control

Technologies	**Total costs (£)**	**Total LYs**	**Total QALYs**	**Incremental costs (£)**	**Incremental LYs**	**Incremental QALYs**	**ICER (£/QALY)**
Larotrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Comparator	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Previous line of therapy naïve comparison

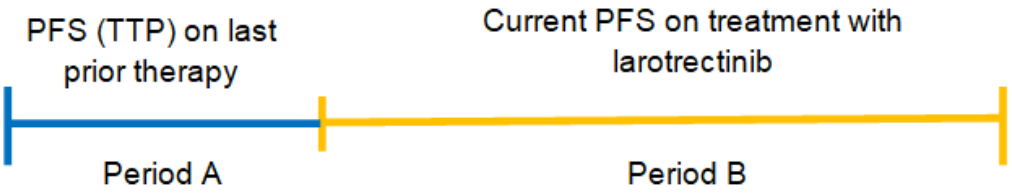
An alternative method for controlling for the larotrectinib arm was considered using the results of the Growth Modulation Index (GMI) presented in section B.2.6. The GMI compares patient's progression-free survival when treated with larotrectinib versus their time-to-progression on their previous line of therapy.

An overview and assessment of this method is presented in Section B.3.2. Results of GMI analyses have been published as clinical analyses, however no evidence was found of this approach being used previously to inform economic analyses (55, 126). The approach requires a number of assumptions to implement into a cost-effectiveness analysis, therefore results should be considered exploratory.

Methodology

The analysis compares the average patient's progression-free survival (PFS) when treated with larotrectinib versus the average patient's time-to-treatment progression (TTP) on their prior therapy. This results in a ratio 'the GMI' between 'Period A' (prior therapy) and 'Period B' (larotrectinib) used to assess the comparative effectiveness of larotrectinib versus the prior therapy in delaying disease progression (Figure 36).

Figure 36. GMI assessment



Two scenarios were conducted. These reflect the primary GMI and analysis and a sensitivity analysis:

- Assessment of GMI based on 52 patients (restricted to those whose previous treatment was in the metastatic disease setting). This additional criteria attempts to control for stage of disease, allows for a more comparable assessment.

- Assessment based on all 73 patients who had received at least 1 prior therapy.

The GMI (ratio of TTP/PFS) was applied as a multiplier to the modelled larotrectinib health outcomes. The mean value was selected to inform the GMI analysis as data was skewed by follow-up time. Using a median value excludes information from patients being treated for larotrectinib long enough to improve vs their previous therapy. (I.e. progression-free patients in the larotrectinib trial whose follow-up time is less than previous TTP, cannot show benefit through this analysis).

This scenario calculates the modelled health outcomes for a previous line of therapy comparator by applying a transformation to the larotrectinib trial data. Assuming that larotrectinib is GMI value times more effective than the previous line of therapy, for PFS and OS. This naïve comparison aims to provide insight into what is plausible when conceptualising performance of larotrectinib in comparison to the previous line of therapy for the analysed cohort. The exploratory analysis focuses on modifying the base case results, remaining a simple naïve analysis. As a result, there are strong limitations and assumptions.

- The GMI multiplier was applied to all health outcomes (OS/PFS life years and overall QALYs) and therefore assumes the same relationship between PFS and OS as larotrectinib.

- The GMI ratio can only be derived for TTP/PFS. The analysis assumes this ratio can also be applied to post-progression survival. This seems a fair assumption given the absence of data to inform this input.

- The analysis compares treatment with larotrectinib against a previous line of therapy where disease is less advanced. This is likely to underestimate the relative benefit of larotrectinib. To attempt to control for this difference an additional criteria was added, restricting to patients whose previous treatment was in the metastatic setting only, results are presented for this subgroup below.

- PFS on larotrectinib is heavily influenced by patients censored due to follow-up, whilst TTP on prior therapy is not. In this respect results from the scenarios represent a 'worst-case' with the GMI likely to improve with follow-up.

Comparator costs were based on the pooled comparator base case results, in order to isolate the analysis on the impact using an alternative data set as a control. The costs therefore do not account for any changes in survival or treatment time driven by varied progression-free and overall survival. Results from the scenario with each GMI value are presented in Table 59 below.

Table 59. Previous line of therapy comparison results

GMI source	GMI value	Larotrectinib			Pooled comparator			
		Costs	QALYs	Life years	Costs	QALYs	Life years	ICER
Base case		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
All patients who received a prior systemic therapy (mean GMI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
All patients receiving prior systemic therapy in the metastatic disease setting (mean GMI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

The naïve comparison shows greater QALYs gained per life year in the comparator arm compared to the base case. However, this result of higher QALYs for the comparator arm in the scenario vs. the base case could be seen as realistic based on patients on previous line of therapy potentially being less advanced than a direct comparison with their current standard of care. Although the exploratory analysis is naïve, the results provide further evidence of plausible and consistent cost-effectiveness of larotrectinib versus the pooled comparator.

NTRK Adjustment Scenario

A SLR was conducted to assess the prognostic nature on NTRK fusion status. An assessment was conducted to explore how this information could be incorporated into an economic evaluation.

Six studies were identified in the SLR (See Appendix D). One study on CMN was excluded because of a lack of comparative data for PFS or OS outcomes (127). Four studies were identified for thyroid papillary. Two studies were excluded as they did not report outcomes of interest (128) (129) . Musholt 2000 did not report a quantified relationship between NTRK1 and no arrangements and median survival was not reached by NTRK1 patients (130). Musholt 2010 found no differences between NTRK1 and BRAF, RET/PTC or unknown mutation. Therefore this paper is reflective of the base case analysis and was not leveraged in this exploratory analysis (131).

A study in colorectal cancer by Pietrantonio et al (132) reported a hazard ratio (HR) for overall survival for a group of patients (n=27) with NTRK (n=13), ALK (n=11) and ROS (n=3) rearrangements versus those without rearrangement (N=319) and was identified as the most appropriate source to incorporate into a scenario analysis.

Methodology

The scenario assumes that the pooled NTRK/ALK/ROS population is representative of NTRK patients. This is a limitation, however NTRK was the most common genetic alteration in the group.

The unadjusted HR of 2.17 was applied as a relative adjustment to the PFS and OS curves of the model. The HR applied to OS was also used for PFS, as no data was presented in the publication for this outcome. The unadjusted OS HR of 2.17 from the univariate analysis was used as opposed to the multivariate model HR of 2.33 to be conservative. All other inputs from the base case remained constant.

By adjusting the comparator arm of the model (to account for NTRK) we assume that these patients do not have NTRK. The scenarios conducted considered the tumour

site within the publication (colorectal cancer) and tumour sites with NTRK prevalence less than 25% in order to minimize bias (in adjusting patients who are NTRK+).

Two scenarios were conducted:

1. Applying the HR only to the colorectal tumour site engine only (reflecting this publication considered patients with colorectal cancer)
2. Applying the HR to all comparators where NTRK prevalence is >25%

Results

Naively applying the HR to comparator arm survival led to shorter cumulative PFS and OS for the included tumour sites, resulting in lower overall QALYs and costs. Given the less favourable results for the comparator arm, the ICER for larotrectinib improved under both scenarios. This change was more prominent with the scenario where the HR was applied across multiple tumour sites (ICER: £[redacted]) compared to when it was applied to colorectal only (ICER: £[redacted]). Results are presented in Table 60 and Table 61.

Table 60. Scenario 1 - survival adjustment for NTRK+ only applied to CRC

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Larotrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Comparator	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 61. Scenario 2 -survival adjustment for NTRK+ applied to all tumour sites where NTRK incidence (<25%)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Larotrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Comparator	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

B.3.9 Subgroup analysis

Not applicable given histology independent nature of the intervention and no identifiable subgroups.

B.3.10 Validation

Comparison of outcomes – model and clinical trial

As part of the validation process, results from the model were compared with outcomes from the larotrectinib clinical trial programme. A summary of this comparison in terms of median OS and PFS is presented in Table 62. The results show close alignment between model and outcomes, with a slight underestimation of time-to-event outcomes for larotrectinib at later time points in the model.

Table 62. Comparison of base case model and trial outcomes

Larotrectinib						
Outcome	Source	3 months (%)	6 months (%)	12 months (%)	18 months (%)	24 months (%)
OS (Weibull)	Trial	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Model	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS (Weibull)	Trial	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Model	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scoping of the cost-effectiveness analysis

A number of steps were taken to validate the approach taken for the economic evaluation. In order to ensure the scientific rigor of this appraisal Bayer partnered with a number of Health Economic advisors.

Scoping of economic model

- An independent health economic and outcomes research consultancy were commissioned to review previous NICE Technology Appraisals to understand how challenges, related to histology independent treatments have previously been addressed in NICE technology appraisals.

- An independent health economic and outcomes research consultancy were commissioned to provide economic analysis and insight into best modelling

practices and advised on the modelling structure and methodology. [redacted]

- A formal advisory board was held in London on 19th November 2018, 8 academic health economists and statisticians provided input on the modelling methodologies that were used to inform the analysis.

Validation of the economic model

Clinical validation

Bayer conducted interviews with a number of UK clinical experts, targeting the broad range of tumour locations included within the larotrectinib clinical programme, in order to validate approaches, data sources and assumptions.

A medical communications agency was commissioned to recruit experts and set up interviews. This followed a stakeholder mapping exercise where experts in the UK had been previously identified according to specialism. All participants completed a declaration of potential conflict of interest.

A discussion guide for the interviews, was created to cover key approaches, assumptions and data sources for this complex submission. The telephone interviews were led by a Bayer health economist and a Bayer clinician and facilitated by the medical communications agency. As a result of the interviews which supported our methodology, there were two data sources that were questioned:

- For adult STS, an alternative source was proposed for the efficacy data for the comparator arm and we have tested this within scenario analysis (section B.3.6.3).

- For salivary gland cancer, advice was that treatment is often based on anecdotal evidence, and whilst our chosen comparator was not 'wrong', as an alternative to the base case, we could explore:

    - Platinum drug + 5-FU

    - Carboplatin + docetaxel
    - Carboplatin + paclitaxel
- We re-reviewed the data identified in the tumour specific SLRs and did not find any data using these regimens that we could test in sensitivity analysis

When validating the approach for equalising resource use, all clinicians commented that whilst it was a fair assumption given the published evidence, it would likely be conservative (given that larotrectinib is a targeted therapy), and that the assumption could result in over-estimation of the resource use with larotrectinib.

Economic validation

Two validation exercises were conducted upon completion of the economic model.

An initial validation was conducted by health economists that had not been involved in the development process. The validation involved checks on the selection and results of different modelling options, calculation spot checks, cross checks against source data and extreme value scenarios to check if the model behaved logically.

The quality check explored the following general aspects of the model:

- Top down tests. This involved systematic variation of the model input parameters to establish whether changes in inputs results in predictable changes in the model outputs. These tests were designed to identify failures in model logic or material computation errors
- Model internal functionality (e.g. testing of all key model parameters, extreme value testing).
- Internal consistency. Accuracy of input data. This was checked by comparing the model inputs in Excel against the data sources referenced

Overall, the validation identified no major issues with the computational accuracy of the model. A number of small inaccuracies were identified and rectified.

A final model validation ran in parallel and was conducted by an independent health economic and outcomes research consultancy. The validation provided a strategic

review of the analytical approach (checking the overall approach is fit for purpose) Quality-control checks covered, but were not limited to, a checklist of basic validity checks (e.g. setting all costs to zero and ensuring the model outputs zero costs), sheet by sheet check of model logic (e.g. checking patient flow sheet calculations), module by module check of VBA logic, validity assessment of outcomes (e.g. comparing available trial data with the outcomes of the model), and editorial checks (e.g. performing a spell check of model content).

B.3.11 Interpretation and conclusions of economic evidence

In order to reduce uncertainty in decision making, Bayer proposes that larotrectinib is made available in a timely manner via the cancer drugs fund

Are the results from this economic evaluation consistent with the published economic literature? If not, why do the results from this evaluation differ, and why should the results in the submission be given more credence than those in the published literature?

This is the first economic evaluation of a histology independent therapy to be conducted in England. Systematic literature reviews found no evidence of published economic evaluations for patients with TRK-Fusion cancer.

Comparator data derived from the literature was benchmarked against the source data (e.g. aligning QALY and LY projections to previous HTAs) from previous submissions, where available, to ensure consistency in model implementation. Minor differences in estimates were noted (especially when complex survival models were adopted and information to reproduce the analysis was limited). Estimates from the de-novo model were generally in line with past models and submissions. The impact of varying inputs such as survival curves for each comparator tumour site is considered in scenario analyses and leads to small changes in the ICER (£[redacted]- £[redacted]/QALY). As the comparator arm is stratified by tumour site, an individual source or assumption has a relatively minimal impact on the ICER.

In the absence of published economic evaluations, or guidance on best-practice for modelling histology independent treatments, alternative modelling methods were explored, including a comparison versus the non-responder population and versus prior therapy. These approaches have limitations and require additional assumptions. Using these methods saw the ICER range between [redacted]- £[redacted]/QALY) suggesting the ICER is robust to alternative modelling methods.

Is the economic evaluation relevant to all groups of patients who could potentially use the technology as identified in the decision problem?

The economic evaluation includes patients with NTRK gene fusions enrolled in the larotrectinib clinical trial programme. This is a highly targeted population, with all patients testing positive for the primary oncogenic driver of the disease.

When considering a conventional treatment paradigm it would be understandable to look at tumour site and consider this a heterogeneous population. Conversely the prognosis of any potential tumour site not captured in the economic evaluation is likely to be reflected in the broad range of tumour sites enrolled in the study.

Due to the rarity of TRK Fusion cancer, inclusion criteria was broad and specific ‘groups’ of patients were not excluded from the study. On this basis there is no reason to suggest this sample of all TRK Fusion patients is not representative of the overall population of TRK Fusion patients.

How relevant (generalisable) is the analysis to clinical practice in England?

No published evidence was found relating specifically to patients with TRK-Fusion cancer in England. However there is no reason to suggest that the tumour site - and age agnostic efficacy (and safety) demonstrated in the larotrectinib studies would not be generalisable to the population found in clinical practice in England. Further information on clinical generalisability is presented in section B.2.13.2.

The comparator arm of the analysis is populated, where possible, with data from previous NICE technology appraisals, reflecting the Committee’s preferred assumptions and data sources. Given the process and scrutiny undertaken to inform these, the data and assumptions of the comparator arm are highly reflective of clinical practice in England.

What are the main strengths and weaknesses of the evaluation? How might these affect the interpretation of the results?

- There are currently no published economic evaluations (or guidance for conducting evaluations) for histology independent treatments. The evaluation presented adheres as closely as possible to the stipulated NICE reference case, uses previously accepted methods and in doing this maintains transparency for decision makers. Alternative modelling methodologies and scenarios are presented allowing for assessment of uncertainty.

- The evaluation independently models standard of care on a tumour site level. This reflects the conventional treatment of cancer in clinical practice. Modelling each tumour site has facilitated:

    - Use of past NICE technology appraisals to inform inputs and assumptions
    - External clinical validation
    - Scenario analyses considering the sensitivity of modelling assumptions for each tumour site.

- Survival data from the larotrectinib clinical trial programme is immature, this is driven, especially for overall survival by the low event numbers. Data from the clinical trial programme is still being collected, once available it can incorporated into the economic model.

- Due to the rarity of TRK Fusion cancer, the number of patients enrolled per tumour site does not currently allow for further matching of patients on their baseline characteristics through conventional methods such as propensity score matching.

- Some tumour sites were very rare and comparator data reflecting the proposed license and positioning could not be identified, or these publications did not have the survival data needed to inform the model. In these cases tumour sites were grouped and later, where possible, this was validated by a clinical expert. Sensitivity of these groupings is considered in a scenario analysis through testing alternative survival models. These consider different projections and may be considered representative

of alternative sources. Results show that further stratification on this basis, even when using extreme values, is unlikely to affect the model results.

- It was not possible to derive comparator data from an NTRK positive population. Instead the comparator population reflects current standard of care where treatment is not targeted towards a genetic alteration and NTRK status is unknown. Scenario analyses considering the potential prognostic effect of NTRK based on the results of a systematic literature review, is presented. Results show the ICER to range between £[redacted]/QALY and £[redacted]/QALY depending on the assumptions used.

What further analyses could be carried out to enhance the robustness or completeness of the results?

- Updated survival analyses (OS and PFS) will allow more accurate estimates of long-term outcomes for patients with TRK Fusion cancer. Later data cuts will also include additional patients increasing the robustness in extrapolated outcomes.
- It is expected that ongoing and future research considering the natural history of TRK-Fusion may be incorporated into future economic evaluations.

B.4 References

1. Tao JJ, Schram AM, Hyman DM. Basket Studies: Redefining Clinical Trials in the Era of Genome-Driven Oncology. Annu Rev Med. 2018;69:319-31.

2. Love-Koh J, Peel A, Rejon-Parrilla JC, Ennis K, Lovett R, Manca A, et al. The Future of Precision Medicine: Potential Impacts for Health Technology Assessment. Pharmacoeconomics. 2018;36(12):1439-51.

3. Brose MS, Bauer TM, Burris HA, Doebele RC, Farago AF, Shaw AT, et al. Loxo-101, a selective pan-TRK inhibitor for patients with TRK-alterations. Thyroid. 2015;25:A156.

4. Hong DS, Bauer TM, Lee JJ, Dowlati A, Brose MS, Farago AF, et al. Larotrectinib in adult patients with solid tumours: a multi-centre, open-label, phase I dose-escalation study. Ann Oncol. 2019;30(2):325-31.

5. Laetsch TW, DuBois SG, Mascarenhas L, Turpin B, Federman N, Albert CM, et al. Larotrectinib for paediatric solid tumours harbouring NTRK gene fusions: phase 1 results from a multicentre, open-label, phase 1/2 study. Lancet Oncol. 2018;19(5):705-14.

6. Hsiao SJ, Zehir A, Sireci AN, Aisner DL. Detection of tumor NTRK gene fusions to identify patients who may benefit from TRK inhibitor therapy. J Mol Diagn. 2019.

7. Vaishnavi A, Le AT, Doebele RC. TRKing down an old oncogene in a new era of targeted therapy. Cancer Discov. 2015;5(1):25-34.

8. Penault-Llorca F, Rudzinski ER, Sepulveda AR. Testing algorithm for identification of patients with TRK fusion cancer. J Clin Pathol. 2019.

9. Meldolesi J. Neurotrophin Trk Receptors: New Targets for Cancer Therapy. Reviews of physiology, biochemistry and pharmacology. 2018;174:67-79.

10. Stransky N, Cerami E, Schalm S, Kim JL, Lengauer C. The landscape of kinase fusions in cancer. Nature communications. 2014;5:4846.

11. Prasad ML, Vyas M, Horne MJ, Virk RK, Morotti R, Liu Z, et al. NTRK fusion oncogenes in pediatric papillary thyroid carcinoma in northeast United States. Cancer. 2016;122(7):1097-107.

12. Bongarzone I, Vigneri P, Mariani L, Collini P, Pilotti S, Pierotti MA. RET/NTRK1 rearrangements in thyroid gland tumors of the papillary carcinoma family: correlation with clinicopathological features. Clin Cancer Res. 1998;4(1):223-8.

13. Vokuhl C, Nourkami-Tutdibi N, Furtwängler R, Gessler M, Graf N, Leuschner I. ETV6-NTRK3 in congenital mesoblastic nephroma: A report of the SIOP/GPOH nephroblastoma study. Pediatric blood & cancer. 2018;65(4).

14. Lassen UN, Doz F, Berlin JD, Oh D-Y, Bielack SS, McDermott R, et al. 409O Larotrectinib efficacy and safety in TRK fusion cancer: An expanded clinical dataset showing consistency in an age and tumor agnostic approach. Annals of Oncology. 2018;29(suppl_8).

15. Drilon A, Laetsch TW, Kummar S, DuBois SG, Lassen UN, Demetri GD, et al. Efficacy of Larotrectinib in TRK Fusion-Positive Cancers in Adults and Children. The New England journal of medicine. 2018;378(8):731-9.

16. Tognon C, Knezevich SR, Huntsman D, Roskelley CD, Melnyk N, Mathers JA, et al. Expression of the ETV6-NTRK3 gene fusion as a primary event in human secretory breast carcinoma. Cancer Cell. 2002;2(5):367-76.

17. Amatu A, Sartore-Bianchi A, Siena S. NTRK gene fusions as novel targets of cancer therapy across multiple tumour types. ESMO Open. 2016;1(2):e000023.

18. Chao MV. Neurotrophins and their receptors: a convergence point for many signalling pathways. Nat Rev Neurosci. 2003;4(4):299-309.

19. Nakagawara A. Trk receptor tyrosine kinases: a bridge between cancer and neural development. Cancer Lett. 2001;169(2):107-14.

20. Barr FG. Fusion genes in solid tumors: the possibilities and the pitfalls. Expert Rev Mol Diagn. 2016;16(9):921-3.

21. Kummar S, Lassen UN. TRK Inhibition: A New Tumor-Agnostic Treatment Strategy. Target Oncol. 2018;13(5):545-56.

22. Lange AM, Lo HW. Inhibiting TRK Proteins in Clinical Cancer Therapy. Cancers (Basel). 2018;10(4).

23. Massard C, Michiels S, Ferte C, Le Deley MC, Lacroix L, Hollebecque A, et al. High-Throughput Genomics and Clinical Outcome in Hard-to-Treat Advanced Cancers: Results of the MOSCATO 01 Trial. Cancer Discov. 2017;7(6):586-95.

24. Haslem DS, Van Norman SB, Fulde G, Knighton AJ, Belnap T, Butler AM, et al. A Retrospective Analysis of Precision Medicine Outcomes in Patients With Advanced Cancer Reveals Improved Progression-Free Survival Without Increased Health Care Costs. J Oncol Pract. 2017;13(2):e108-e19.

25. Schwaederle M, Zhao M, Lee JJ, Eggermont AM, Schilsky RL, Mendelsohn J, et al. Impact of Precision Medicine in Diverse Cancers: A Meta-Analysis of Phase II Clinical Trials. J Clin Oncol. 2015;33(32):3817-25.

26. Schwaederle M, Zhao M, Lee JJ, Lazar V, Leyland-Jones B, Schilsky RL, et al., editors. Impact of precision medicine in refractory malignancies: A meta-analysis of 13,203 patients in phase I clinical trials2016.

27. White Al-Habeeb N, Kulasingam V, Diamandis EP, Yousef GM, Tsongalis GJ, Vermeulen L, et al. The Use of Targeted Therapies for Precision Medicine in Oncology. Clinical chemistry. 2016;62(12):1556-64.

28. Haslem DS, Chakravarty I, Fulde G, Gilbert H, Tudor BP, Lin K, et al. Precision oncology in advanced cancer patients improves overall survival with lower weekly healthcare costs. Oncotarget. 2018;9(15):12316-22.

29. Bell PSJ. Life Sciences Industrial Strategy – A report to the Government from the life sciences sector. 2017.

30. NHS. The NHS Long Term Plan. 2019.

31. England N. Improving Outcomes Through Personalised Medicine. 2016.

32. England N. 2018/19 final draft National Genomics Test Directory FAQ 2018.

33. England N. National Genomic Test Directory for Cancer. 2019.

34. Mateo J, Chakravarty D, Dienstmann R, Jezdic S, Gonzalez-Perez A, Lopez-Bigas N, et al. A framework to rank genomic alterations as targets for cancer precision medicine: the ESMO Scale for Clinical Actionability of molecular Targets (ESCAT). Ann Oncol. 2018;29(9):1895-902.

35. Brose MS, Albert CM, Waguespack SG, al. e. Activity of larotrectinib in patients with advanced TRK fusion thyroid cancer. Thyroid. 2018;28(S1):A-4.

36. Farago A, Kummar S, Ibabekci S, Corsi-Travali S, Cruickshank S, Cox M, et al. P1.13-40 Rapid, Robust and Durable Responses to Larotrectinib in Patients with TRK Fusion Non-Small Cell Lung Cancer. Journal of Thoracic Oncology. 2018;13(10):S597-S8.

37. Nathenson M, Demetri G, Lassen U, Hong D, Boni V, Deeken J, et al. O-020Activity of larotrectinib in patients with TRK fusion GI malignancies. Annals of Oncology. 2018;29(suppl_5).

38. Wirth L, Drilon A, Albert C, Farago A, Wel D, Ma P, et al. Larotrectinib Is Highly Active in Patients With Advanced Recurrent TRK Fusion Thyroid (TC) and Salivary Gland Cancers (SGC). International Journal of Radiation Oncology • Biology • Physics. 2018;100(5):1318.

39. DuBois SG, Laetsch TW, Federman N, Turpin BK, Albert CM, Nagasubramanian R, et al. The use of neoadjuvant larotrectinib in the management of children with locally advanced TRK fusion sarcomas. Cancer. 2018;124(21):4241-7.

40. Lemery S, Keegan P, Pazdur R. First FDA Approval Agnostic of Cancer Site - When a Biomarker Defines the Indication. N Engl J Med. 2017;377(15):1409-12.

41. Bayer. Interim Statistical Analysis Plan (SAP). Study LOXO-TRK-15002. 2017 Contract No.: v1.0 4th August 2017.

42. Bayer. Statistical Analysis Plan (SAP). Clinical study LOXO-TRK-14001. 2017 Contract No.: v1.0 3rd August 2017.

43. Bayer. Statistical Analysis Plan (SAP). Clinical Study LOXO-TRK-15003. 2017 Contract No.: v1.0 3rd August 2017.

44. Bayer. Data on file. EMA submission. 2.7.3 Summary of Clinical Efiicacy (2731953). 2018.

45. Bayer. Data on file. EMA submission. CSE_tables_figures_listings_cut-off_30JUL2018. 2019.

46. Bayer. CSS_tables_figures_listings_cut-off_30JUL2018. 2019.

47. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228-47.

48. Wen PY, Macdonald DR, Reardon DA, Cloughesy TF, Sorensen AG, Galanis E, et al. Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group. J Clin Oncol. 2010;28(11):1963-72.

49. Cocks K, King MT, Velikova G, de Castro G, Jr., Martyn St-James M, Fayers PM, et al. Evidence-based guidelines for interpreting change scores for the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30. Eur J Cancer. 2012;48(11):1713-21.

50. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. J Clin Oncol. 1998;16(1):139-44.

51. Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. Health Qual Life Outcomes. 2007;5:70.

52. Cote CJ, Lerman J, Todres D. A practice of anesthesia for infants and children. 2009.

53. Davies J, Patel M, Gridelli C, de Marinis F, Waterkamp D, McCusker ME. Real-world treatment patterns for patients receiving second-line and third-line treatment for advanced non-small cell lung cancer: A systematic review of recently published studies. PloS one. 2017;12(4):e0175679.

54. Park IH, Lee KS, Ro J. Effects of second and subsequent lines of chemotherapy for metastatic breast cancer. Clinical breast cancer. 2015;15(1):e55-62.

55. Von Hoff DD, Stephenson JJ, Jr., Rosen P, Loesch DM, Borad MJ, Anthony S, et al. Pilot study using molecular profiling of patients' tumors to find potential targets and select treatments for their refractory cancers. J Clin Oncol. 2010;28(33):4877-83.

56. Von Hoff DD. There are no bad anticancer agents, only bad clinical trial designs--twenty-first Richard and Hinda Rosenthal Foundation Award Lecture. Clinical cancer research : an official journal of the American Association for Cancer Research. 1998;4(5):1079-86.

57. Bayer. Data on file. GMI analysis. 2019.

58. Quigley JM, Thompson JC, Halfpenny NJ, Scott DA. Critical appraisal of nonrandomized studies-A review of recommended and commonly used tools. J Eval Clin Pract. 2019;25(1):44-52.

59. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health. 1998;52(6):377-84.

60. Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovitch C, Song F, et al. Evaluating non-randomised intervention studies. Health Technol Assess. 2003;7(27):iii-x, 1-173.

61. Vaishnavi A, Capelletti M, Le AT, Kako S, Butaney M, Ercan D, et al. Oncogenic and drug-sensitive NTRK1 rearrangements in lung cancer. Nature medicine. 2013;19(11):1469-72.

62. Skalova A, Vanecek T, Sima R, Laco J, Weinreb I, Perez-Ordonez B, et al. Mammary analogue secretory carcinoma of salivary glands, containing the ETV6-NTRK3 fusion gene: a hitherto undescribed salivary gland tumor entity. The American journal of surgical pathology. 2010;34(5):599-608.

63. Loh ML, Ahn P, Perez-Atayde AR, Gebhardt MC, Shamberger RC, Grier HE. Treatment of infantile fibrosarcoma with chemotherapy and surgery: results from the Dana-Farber Cancer Institute and Children's Hospital, Boston. J Pediatr Hematol Oncol. 2002;24(9):722-6.

64. Orbach D, Brennan B, De Paoli A, Gallego S, Mudry P, Francotte N, et al. Conservative strategy in infantile fibrosarcoma is possible: The European paediatric Soft tissue sarcoma Study Group experience. European journal of cancer (Oxford, England : 1990). 2016;57:1-9.

65. Orbach D, Rey A, Cecchetto G, Oberlin O, Casanova M, Thebaud E, et al. Infantile fibrosarcoma: management based on the European experience. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2010;28(2):318-23.

66. Nagasubramanian R, Wei J, Gordon P, Rastatter JC, Cox MC, Pappo A. Infantile Fibrosarcoma With NTRK3-ETV6 Fusion Successfully Treated With the Tropomyosin-Related Kinase Inhibitor LOXO-101. Pediatric blood & cancer. 2016;63(8):1468-70.

67. Suehara Y, Trahair T, Kirby M, Saito S, Capra M, Ghosh T, et al. Expanded access of larotrectinib to manage pediatric TRK fusion sarcomas. SIOP 2018, 50th Congress of the International Society of Paediatric Oncology, 16 – 19 November 2018, Kyoto, Japan2018.

68. Solutions BRH. Data on File. Exploratory Analyses of Quality of Life Based on Patient-Reported Outcome Assessments for Larotrectinib Studies 15002 and 15003. 2019.

69. Doebele RC, Davis LE, Vaishnavi A, Le AT, Estrada-Bernal A, Keysar S, et al. An Oncogenic NTRK Fusion in a Patient with Soft-Tissue Sarcoma with Response to the Tropomyosin-Related Kinase Inhibitor LOXO-101. Cancer Discov. 2015;5(10):1049-57.

70. Clark MA, Fisher C, Judson I, Thomas JM. Soft-tissue sarcomas in adults. N Engl J Med. 2005;353(7):701-11.

71. Shi E, Chmielecki J, Tang C-M, Wang K, Heinrich MC, Kang G, et al. FGFR1 and NTRK3 actionable alterations in "Wild-Type" gastrointestinal stromal tumors. Journal of translational medicine. 2016;14(1):339.

72. Bayer. Data on file. Case studies. 2019.

73. Ronsley R, Rassekh SR, Shen Y, Lee AF, Jantzen C, Halparin J, et al. Application of genomics to identify therapeutic targets in recurrent pediatric papillary thyroid carcinoma. Cold Spring Harb Mol Case Stud. 2018;4(2).

74. Parisi MT, Eslamy H, Mankoff D. Management of Differentiated Thyroid Cancer in Children: Focus on the American Thyroid Association Pediatric Guidelines. Semin Nucl Med. 2016;46(2):147-64.

75. Fard-Esfahani A, Emami-Ardekani A, Fallahi B, Fard-Esfahani P, Beiki D, Hassanzadeh-Rad A, et al. Adverse effects of radioactive iodine-131 treatment for differentiated thyroid carcinoma. Nucl Med Commun. 2014;35(8):808-17.

76. Shukla N, Roberts SS, Baki MO, Mushtaq Q, Goss PE, Park BH, et al. Successful Targeted Therapy of Refractory Pediatric ETV6-NTRK3 Fusion-Positive Secretory Breast Carcinoma. JCO precision oncology. 2017;2017.

77. Xcenda B. Data on file. Systematic Literature Review to Support Larotrectinib in Solid Tumors Harboring Neurotrophic Tyrosine Receptor Kinases (NTRK) Gene Fusions. 2019.

78. Bayer. Draft Summary of Product Characteristics. 2019.

79. National Cancer Intelligence Network and Cancer Research UK. Cancer Incidence and Survival by Major Ethnic Group, England, 2002-2006. 2009.

80. Broekman F, Giovannetti E, Peters GJ. Tyrosine kinase inhibitors: Multi-targeted or single-targeted? World J Clin Oncol. 2011;2(2):80-93.

81. Phillippo D, Ades T, Dias S, Palmer S, Abrams KR, Welton N. NICE DSU technical support document 18: methods for population-adjusted indirect comparisons in submissions to NICE. 2016.

82. Briggs A, Sculpher M, Claxton K. Decision modelling for health economic evaluation: OUP Oxford; 2006.

83. Stahl JE. Modelling methods for pharmacoeconomics and health technology assessment. Pharmacoeconomics. 2008;26(2):131-48.

84. National Insitute of Health and Care Excellence. Guide to the methods of technology appraisal,. 2013.

85. Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, et al. State-transition modeling: a report of the ISPOR-SMDM modeling good research practices task force-3. Value in Health. 2012;15(6):812-20.

86. Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Med Res Methodol. 2012;12:9.

87. Nicholas Latimer. NICE DSU Technical support document 14: survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data 2011.

88. National Health Service (NHS). NHS reference costs collection guidance for 2017 to 2018. 2018. Available from: https://improvement.nhs.uk/resources/reference-costs/.

89. Ishak KJ, Kreif N, Benedict A, Muszbek N. Overview of parametric survival analysis for health-economic applications. Pharmacoeconomics. 2013;31(8):663-75.

90. Latimer NR. Survival analysis for economic evaluations alongside clinical trials—extrapolation with patient-level data: inconsistencies, limitations, and a practical guide. Medical Decision Making. 2013;33(6):743-54.

91. Collett D. Modelling survival data in medical research. 2003.

92. Office of National Statistics (ONS). National life tables, UK: 2014 to 2016. Available from: https://www.ons.gov.uk/releases/nationallifetablesuk2014to2016. 2017.

93. National Institute for Health and Care Excellence. Pegylated liposomal irinotecan for treating pancreatic cancer after gemcitabine. Technology appraisal guidance [TA440]. 2017. Available from: https://www.nice.org.uk/guidance/TA440.

94. National Institute for Health and Care Excellence. Lenvatinib and sorafenib for treating differentiated thyroid cancer after radioactive iodine [TA535]. 2018. Available from: https://www.nice.org.uk/guidance/ta535

95. National Institute for Health and Care Excellence. Nivolumab for treating squamous cell carcinoma of the head and neck after platinum-based chemotherapy. Technology appraisal guidance [TA490]. 2019. Available from: https://www.nice.org.uk/guidance/ta490

96. Shepherd FA, Rodrigues Pereira J, Ciuleanu T, Tan EH, Hirsh V, Thongprasert S, et al. Erlotinib in Previously Treated Non–Small-Cell Lung Cancer. New England Journal of Medicine. 2005;353(2):123-32.

97. Batchelor TT, Mulholland P, Neyns B, Nabors LB, Campone M, Wick A, et al. Phase III randomized trial comparing the efficacy of cediranib as monotherapy, and in combination with lomustine, versus lomustine alone in patients with recurrent glioblastoma. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2013;31(26):3212-8.

98. Valle J, Wasan H, Palmer DH, Cunningham D, Anthoney A, Maraveyas A, et al. Cisplatin plus gemcitabine versus gemcitabine for biliary tract cancer. N Engl J Med. 2010;362(14):1273-81.

99. National Institute for Health and Care Excellence. Position statement on use of the EQ-5D-5L valuation set for England (updated November 2018) 2018 [09 April 2019]. Available from: https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/technology-appraisal-guidance/eq-5d-5l.

100. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: An EQ-5D-5L value set for England. Health Econ. 2018;27(1):7-22.

101. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value Health. 2012;15(5):708-15.

102. Khan KA, Petrou S, Rivero-Arias O, Walters SJ, Boyle SE. Mapping EQ-5D utility scores from the PedsQL generic core scales. Pharmacoeconomics. 2014;32(7):693-706.

103. Longworth L, Rowen D. NICE Decision Support Unit Technical Support Documents. NICE DSU Technical Support Document 10: The Use of Mapping Methods to Estimate Health State Utility Values. London: National Institute for Health and Care Excellence (NICE)
unless otherwise stated. All rights reserved.; 2011.

104. National Institute for Health and Care Excellence. Erlotinib and gefitinib for treating non-small-cell lung cancer that has progressed after prior chemotherapy 2015 [26 February 2019]. Available from: https://www.nice.org.uk/guidance/ta374.

105. National Institute for Health and Care Excellence. Pembrolizumab for treating unresectable, metastatic melanoma after progression with ipilimumab [ID760]. Technology appraisal guidance [TA357]. 2015. Available from: https://www.nice.org.uk/guidance/ta357

106. National Institute for Health and Care Excellence. Trifluridine–tipiracil for previously treated metastatic colorectal cancer. Technology appraisal guidance [TA405]. 2016. Available from: https://www.nice.org.uk/guidance/ta405

107. National Institute for Health and Care Excellence. Regorafenib for previously treated unresectable or metastatic gastrointestinal stromal tumours. Technology appraisal guidance [TA488]. 2017. Available from: https://www.nice.org.uk/guidance/ta488.

108. National Institute for Health and Care Excellence. Trabectedin for the treatment of advanced soft tissue sarcoma. Technology appraisal guidance [TA185]. 2010. Available from: https://www.nice.org.uk/guidance/ta185

109. National Institute for Health and Care Excellence. Eribulin for treating locally advanced or metastatic breast cancer after 2 or more chemotherapy regimens. Technology appraisal TA423 2016. Available from: https://www.nice.org.uk/guidance/ta423/resources/eribulin-for-treating-locally-advanced-or-metastatic-breast-cancer-after-2-or-more-chemotherapy-regimens-pdf-82604662012357.

110. National Institute for Health and Care Excellence. Guidance on the use of temozolomide for the treatment of recurrent malignant glioma (brain cancer) [TA23]. 2001. Available from: https://www.nice.org.uk/guidance/ta23.

111. Liberato NL, Rognoni C, Rubrichi S, Quaglini S, Marchetti M, Gorlia T, et al. Adding docetaxel to cisplatin and fluorouracil in patients with unresectable head and neck cancer: a cost-utility analysis. Ann Oncol. 2012;23(7):1825-32.

112. Delea T, Amdahl J, Nakhaipour H, Manson S, Wang A, Fedor N, et al. Cost-effectiveness of pazopanib in advanced soft-tissue sarcoma in Canada. Curr Oncol. 2014;21(6):e748-59.

113. Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. Health Qual Life Outcomes. 2008;6:84.

114. Amdahl J, Manson SC, Isbell R, Chit A, Diaz J, Lewis L, et al. Cost-effectiveness of pazopanib in advanced soft tissue sarcoma in the United kingdom. Sarcoma. 2014;2014:481071.

115. National Health Service. Drugs and pharmaceutical electronic market information tool (eMIT) 2018. Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit.

116. The Royal Pharmaceutical society. British National Formulary. 2018. Available from: https://about.medicinescomplete.com/.

117. Airoldi M, Pedani F, Succo G, Gabriele AM, Ragona R, Marchionatti S, et al. Phase II randomized trial comparing vinorelbine versus vinorelbine plus cisplatin in patients with recurrent salivary gland malignancies. Cancer. 2001;91(3):541-7.

118. Mascarenhas L, Lyden ER, Breitfeld PP, Walterhouse DO, Donaldson SS, Paidas CN, et al. Randomized phase II window trial of two schedules of irinotecan with vincristine in patients with first relapse or progression of rhabdomyosarcoma: a report from the Children's Oncology Group. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2010;28(30):4658-63.

119. Round J, Jones L, Morris S. Estimating the cost of caring for people with cancer at the end of life: A modelling study. Palliat Med. 2015;29(10):899-907.

120. Personal Social Services Research Unit. Unit Costs of Health & Social Care 2018 2018. Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2018/.

121. National Institute for Health and Care Excellence. Blood transfusion. Available from: https://www.nice.org.uk/guidance/ng24. 2015.

122. Monthly Index of Medical Specialities. Database of prescription and generic drugs, clinical guidelines 2018 2018. Available from: https://www.mims.co.uk/.

123. Roth JA, Carlson JJ. Cost-effectiveness of gemcitabine + cisplatin vs. gemcitabine monotherapy in advanced biliary tract cancer. J Gastrointest Cancer. 2012;43(2):215-23.

124. Claxton K, Sculpher M, McCabe C, Briggs A, Akehurst R, Buxton M, et al. Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. Health economics. 2005;14(4):339-47.

125. Hatswell AJ, Thompson GJ, Maroudas PA, Sofrygin O, Delea TE. Estimating outcomes and cost effectiveness using a single-arm clinical trial: ofatumumab for double-refractory chronic lymphocytic leukemia. Cost effectiveness and resource allocation : C/E. 2017;15:8.

126. Dean A, Byrne A, Marinova M, Hayden I. Clinical Outcomes of Patients with Rare and Heavily Pretreated Solid Tumors Treated according to the Results of Tumor Molecular Profiling. BioMed Research International. 2016;2016:9.

127. Vokuhl C, Nourkami-Tutdibi N, Furtwängler R, Gessler M, Graf N, Leuschner I. ETV6–NTRK3 in congenital mesoblastic nephroma: A report of the SIOP/GPOH nephroblastoma study. Pediatric blood & cancer. 2018;65(4):e26925.

128. Prasad ML, Vyas M, Horne MJ, Virk RK, Morotti R, Liu Z, et al. NTRK fusion oncogenes in pediatric papillary thyroid carcinoma in northeast United States. Cancer. 2016;122(7):1097-107.

129. Bongarzone I, Vigneri P, Mariani L, Collini P, Pilotti S, Pierotti MA. RET/NTRK1 rearrangements in thyroid gland tumors of the papillary carcinoma family: correlation with clinicopathological features. Clinical Cancer Research. 1998;4(1):223-8.

130. Musholt TJ, Musholt PB, Khaladj N, Schulz D, Scheumann GF, Klempnauer J. Prognostic significance of RET and NTRK1 rearrangements in sporadic papillary thyroid carcinoma. Surgery. 2000;128(6):984-93.

131. Musholt TJ, Schönefeld S, Schwarz CH, Watzka FM, Musholt PB, Fottner C, et al. Impact of pathognomonic genetic alterations on the prognosis of papillary thyroid carcinoma. Langenbeck's archives of surgery. 2010;395(7):877-83.

132. Pietrantonio F, Di Nicolantonio F, Schrock AB, Lee J, Tejpar S, Sartore-Bianchi A, et al. ALK, ROS1, and NTRK rearrangements in metastatic colorectal cancer. JNCI: Journal of the National Cancer Institute. 2017;109(12).

133. Drilon A, Sankhala KK, Liu SV, Cho BC, Blakely C, Chee CE, et al. Abstract CT060: STARTRK-2: A global phase 2, open-label, basket study of entrectinib in patients with locally advanced or metastatic solid tumors harboring TRK, ROS1, or ALK gene fusions. Cancer Research. 2017;77(13 Supplement):CT060-CT.

134. Demetri GD, Bazhenova L, John T, Besse B, Wolf J, Seto T, et al. LBA4 Efficacy and safety of entrectinib in patients with NTRK fusion-positive tumours: Pooled analysis of STARTRK-2, STARTRK-1, and ALKA-372-001. Annals of Oncology. 2018;29(suppl_9).

135. Desai AV, Brodeur GM, Foster J, Berg SL, Basu EM, Shusterman S, et al. Phase 1 study of entrectinib (RXDX-101), a TRK, ROS1, and ALK inhibitor, in children, adolescents, and young adults with recurrent or refractory solid tumors. Journal of Clinical Oncology. 2018;36(15_suppl):10536-.

136. Drilon AE, Ou S-HI, Cho BC, Kim D-W, Lee J, Lin JJ, et al. A phase 1 study of the next-generation ALK/ROS1/TRK inhibitor ropotrectinib (TPX-0005) in patients with advanced ALK/ROS1/NTRK+ cancers (TRIDENT-1). Journal of Clinical Oncology. 2018;36(15_suppl):2513-.

137. Drilon A, Siena S, Ou SI, Patel M, Ahn MJ, Lee J, et al. Safety and Antitumor Activity of the Multitargeted Pan-TRK, ROS1, and ALK Inhibitor Entrectinib: Combined Results from Two Phase I Trials (ALKA-372-001 and STARTRK-1). Cancer Discov. 2017;7(4):400-9.

138. Pietrantonio F, Di Nicolantonio F, Schrock AB, Bardelli A, Amatu A, Falcone A, et al. ALK, ROS1, and NTRK Rearrangements in Metastatic Colorectal Cancer. JNCI: Journal of the National Cancer Institute. 2017;109(12).

139. Musholt TJ, Schönefeld S, Schwarz CH, Watzka FM, Musholt PB, Fottner C, et al. Impact of pathognomonic genetic alterations on the prognosis of papillary thyroid carcinoma. ESES vienna presentation. Langenbeck's archives of surgery. 2010;395(7):877-83.

140. Musholt TJ, Musholt PB, Khaladj N, Schulz D, Scheumann GF, Klempnauer J. Prognostic significance of RET and NTRK1 rearrangements in sporadic papillary thyroid carcinoma. Surgery. 2000;128(6):984-93.

141. Farago AF, Taylor MS, Doebele RC, Zhu VW, Kummar S, Spira AI, et al. Clinicopathologic Features of Non-Small-Cell Lung Cancer Harboring an NTRK Gene Fusion. JCO precision oncology. 2018;2018.

142. Boon E, Valstar MH, van der Graaf WTA, Bloemena E, Willems SM, Meeuwis CA, et al. Clinicopathological characteristics and outcome of 31 patients with ETV6-NTRK3 fusion gene confirmed (mammary analogue) secretory carcinoma of salivary glands. Oral Oncol. 2018;82:29-33.

143. Gooskens SL, Houwing ME, Vujanic GM, Dome JS, Diertens T, Coulomb-l'Hermine A, et al. Congenital mesoblastic nephroma 50 years after its recognition: A narrative review. Pediatric blood & cancer. 2017;64(7).

144. Jehangir S, Kurian JJ, Selvarajah D, Thomas RJ, Holland AJA. Recurrent and metastatic congenital mesoblastic nephroma: where does the evidence stand? Pediatr Surg Int. 2017;33(11):1183-8.

145. Wang ZP, Li K, Dong KR, Xiao XM, Zheng S. Congenital mesoblastic nephroma: Clinical analysis of eight cases and a review of the literature. Oncol Lett. 2014;8(5):2007-11.

146. Furtwaengler R, Reinhard H, Leuschner I, Schenk JP, Goebel U, Claviez A, et al. Mesoblastic nephroma--a report from the Gesellschaft fur Padiatrische Onkologie und Hamatologie (GPOH). Cancer. 2006;106(10):2275-83.

147. van den Heuvel-Eibrink MM, Grundy P, Graf N, Pritchard-Jones K, Bergeron C, Patte C, et al. Characteristics and survival of 750 children diagnosed with a renal tumor in the first seven months of life: A collaborative study by the SIOP/GPOH/SFOP, NWTSG, and UKCCSG Wilms tumor study groups. Pediatric blood & cancer. 2008;50(6):1130-4.

148. Baghai F, Yazdani F, Etebarian A, Garajei A, Skalova A. Clinicopathologic and molecular characterization of mammary analogue secretory carcinoma of salivary gland origin. Pathology, research and practice. 2017;213(9):1112-8.

149. Chiosea SI, Griffith C, Assaad A, Seethala RR. Clinicopathological characterization of mammary analogue secretory carcinoma of salivary glands. Histopathology. 2012;61(3):387-94.

150. Skálová A, Vanecek T, Sima R, Laco J, Weinreb I, Perez-Ordonez B, et al. Mammary analogue secretory carcinoma of salivary glands, containing the ETV6-NTRK3 fusion gene: a hitherto undescribed salivary gland tumor entity. The American journal of surgical pathology. 2010;34(5):599-608.

151. Skálová A, Vanecek T, Majewska H, Laco J, Grossmann P, Simpson RHW, et al. Mammary analogue secretory carcinoma of salivary glands with high-grade transformation: report of 3 cases with the ETV6-NTRK3 gene fusion and analysis of TP53, β-catenin, EGFR, and CCND1 genes. The American journal of surgical pathology. 2014;38(1):23-33.

152. Ito Y, Ishibashi K, Masaki A, Fujii K, Fujiyoshi Y, Hattori H, et al. Mammary analogue secretory carcinoma of salivary glands: a clinicopathologic and molecular study including 2 cases harboring ETV6-X fusion. Am J Surg Pathol. 2015;39(5):602-10.

153. Bongarzone I, Fugazzola L, Vigneri P, Mariani L, Mondellini P, Pacini F, et al. Age-related activation of the tyrosine kinase receptor protooncogenes RET and NTRK1 in papillary thyroid carcinoma. The Journal of clinical endocrinology and metabolism. 1996;81(5):2006-9.

154. Motheral B, Brooks J, Clark MA, Crown WH, Davey P, Hutchins D, et al. A Checklist for Retrospective Database Studies—Report of the ISPOR Task Force on Retrospective Databases. Value in Health. 2003;6(2):90-7.

155. Meric-Bernstam F, Shukla N, Peled N, Landman Y, Onitilo A, Montez S, et al. Abstract P6-20-02: Activity of larotrectinib, a highly selective inhibitor of tropomyosin receptor kinase, in TRK fusion breast cancers. Cancer Research. 2019;79(4 Supplement):P6-20-02-P6-20-02.

156. Landman Y, Ilouze M, Wein S, Neiman V, Yerushalmi R, Yakimov M, et al. Rapid Response to Larotrectinib (LOXO-101) in an Adult Chemotherapy-Naive Patients With Advanced Triple-Negative Secretory Breast Cancer Expressing ETV6-NTRK3 Fusion. Clinical breast cancer. 2018;18(3):e267-e70.

157. Gibson E, Koblbauer I, Begum N, Dranitsaris G, Liew D, McEwan P, et al. Modelling the Survival Outcomes of Immuno-Oncology Drugs in Economic Evaluations: A Systematic Approach to Data Analysis and Extrapolation. Pharmacoeconomics. 2017;35(12):1257-70.

158. Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. Stat Med. 2002;21(15):2175-97.

159. Jr. FEH. Regression Modeling Strategies. Statistics SSi, editor2001.

160. Cleveland WS. Robust Locally Weighted Regression and Smoothing Scatterplots. Journal of the American Statistical Association. 1979;74(368):829-36.

161. Robert E. Kass & Adrian E. Raftery. Bayes Factors. Journal of the American Statistical Association 1995;90(430).

162. Mdzinarishvili T, Sherman S. Weibull-like model of cancer development in aging. Cancer informatics. 2010;9:CIN. S5460.

163. George B, Seals S, Aban I. Survival analysis and regression models. Journal of Nuclear Cardiology. 2014;21(4):686-94.

164. Zhang J, Hanson T, Zhou H. Bayes factors for choosing among six common survival models. Lifetime data analysis. 2019;25(2):361-79.

165. Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. Health and quality of life outcomes. 2008;6:84-.

166. Review IfCaE. Treatment Options for Advanced Non-Small Cell Lung Cancer: Effectiveness, Value and Value-Based Price Benchmarks 2016. Available from: https://icer-review.org/wp-content/uploads/2016/10/MWCEPAC_NSCLC_Final_Evidence_Report_Meeting_Summary_110116.pdf.

167. Hamid O, Puzanov I, Dummer R, Schachter J, Daud A, Schadendorf D, et al. Final analysis of a randomised trial comparing pembrolizumab versus investigator-choice chemotherapy for ipilimumab-refractory advanced melanoma. European journal of cancer (Oxford, England : 1990). 2017;86:37-45.

168. Ribas A, Puzanov I, Dummer R, Schadendorf D, Hamid O, Robert C, et al. Pembrolizumab versus investigator-choice chemotherapy for ipilimumab-refractory melanoma (KEYNOTE-002): a randomised, controlled, phase 2 trial. Lancet Oncol. 2015;16(8):908-18.

169. National Institute for Health and Care Excellence. Ipilimumab for previously treated advanced (unresectable or metastatic) melanoma - Single Technology Appraisal [TA268]. Available from: https://www.nice.org.uk/guidance/ta268. 2012.

170. Tejani MA, ter Veer A, Milne D, Ottesen R, Bekaii-Saab T, Benson AB, 3rd, et al. Systemic therapy for advanced appendiceal adenocarcinoma: an analysis from the NCCN Oncology Outcomes Database for colorectal cancer. Journal of the National Comprehensive Cancer Network : JNCCN. 2014;12(8):1123-30.

171. Mayer RJ, Van Cutsem E, Falcone A, Yoshino T, Garcia-Carbonero R, Mizunuma N, et al. Randomized trial of TAS-102 for refractory metastatic colorectal cancer. N Engl J Med. 2015;372(20):1909-19.

172. Yoshino T, Mizunuma N, Yamazaki K, Nishina T, Komatsu Y, Baba H, et al. TAS-102 monotherapy for pretreated metastatic colorectal cancer: a double-blind, randomised, placebo-controlled phase 2 trial. Lancet Oncol. 2012;13(10):993-1001.

173. Grothey A, Van Cutsem E, Sobrero A, Siena S, Falcone A, Ychou M, et al. Regorafenib monotherapy for previously treated metastatic colorectal cancer (CORRECT): an international, multicentre, randomised, placebo-controlled, phase 3 trial. Lancet. 2013;381(9863):303-12.

174. Demetri GD, Reichardt P, Kang YK, Blay JY, Rutkowski P, Gelderblom H, et al. Efficacy and safety of regorafenib for advanced gastrointestinal stromal tumours after failure of imatinib and sunitinib (GRID): an international, multicentre, randomised, placebo-controlled, phase 3 trial. Lancet (London, England). 2013;381(9863):295-302.

175. Abel J, Pring A, Rich A, Malik T, Verne J. The impact of advance care planning of place of death, a hospice retrospective cohort study. BMJ Supportive & Palliative Care. 2013;3(2):168-73.

176. van der Graaf WT, Blay JY, Chawla SP, Kim DW, Bui-Nguyen B, Casali PG, et al. Pazopanib for metastatic soft-tissue sarcoma (PALETTE): a randomised, double-blind, placebo-controlled phase 3 trial. Lancet (London, England). 2012;379(9829):1879-86.

177. Demetri GD, von Mehren M, Jones RL, Hensley ML, Schuetze SM, Staddon A, et al. Efficacy and Safety of Trabectedin or Dacarbazine for Metastatic Liposarcoma or Leiomyosarcoma After Failure of Conventional Chemotherapy: Results of a Phase III Randomized Multicenter Clinical Trial. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2016;34(8):786-93.

178. National Institute for Health and Care Excellence. Pemetrexed for the maintenance treatment of non-small-cell lung cancer 2017. Available from: https://www.nice.org.uk/guidance/ta190.

179. Lloyd A, Nafees B, Narewska J, Dewilde S, Watkins J. Health state utilities for metastatic breast cancer. Br J Cancer. 2006;95(6):683-90.

180. Tsukiyama I, Ejiri M, Yamamoto Y, Nakao H, Yoneda M, Matsuura K, et al. A Cost-Effectiveness Analysis of Gemcitabine plus Cisplatin Versus Gemcitabine Alone for Treatment of Advanced Biliary Tract Cancer in Japan. J Gastrointest Cancer. 2017;48(4):326-32.

181. Swinburn P, Lloyd A, Nathan P, Choueiri TK, Cella D, Neary MP. Elicitation of health state utilities in metastatic renal cell carcinoma. Curr Med Res Opin. 2010;26(5):1091-6.

182. Peasgood T, Ward SE, Brazier J. Health-state utility values in breast cancer. Expert Rev Pharmacoecon Outcomes Res. 2010;10(5):553-66.

183. Mansel R, Locker G, Fallowfield L, Benedict A, Jones D. Cost-effectiveness analysis of anastrozole vs tamoxifen in adjuvant therapy for early stage breast cancer in the United Kingdom: the 5-year completed treatment analysis of the ATAC ('Arimidex', Tamoxifen alone or in combination) trial. Br J Cancer. 2007;97(2):152-61.

184. Centre for Bayesian Statistics in Health Economics (CHEBS). Excel macros for sampling from multivariate Normal and Dirichlet distributions [April 2019]. Available from: https://www.sheffield.ac.uk/chebs/software.

185. Longworth L. RD. NICE DSU Technical Support Document 10: the use of mapping methods to estimate health state utility values. April 2011.

186. National Institute for Health and Care Excellence (NICE). EQ5D-5L NICE position statement. November 2018. Available from: https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/technology-appraisal-guidance/eq-5d-5l.

187. Dolan P. Modeling valuations for EuroQol health states. Medical care. 1997;35(11):1095-108.

A1. Priority question: Have analyses for a more recent data cut than July 2018 been performed? If they have, please provide an update for the clinical analyses; in particular, updating Tables 12-15 (company submission, CS) and all Kaplan-Meier curves (Figures 14-17, CS). If no update has been performed, please explain why.

Limited data analysis has been performed for the purpose of disclosure at the upcoming congress - ESMO 2019. As such, we cannot directly update the tables and figures as requested at this time.

Updated data in an expanded cohort of [redacted] total TRK fusion patients treated with larotrectinib, with [redacted] (55 primary + [redacted] supplemental) evaluable for efficacy will be presented. Data cut off was 19 February 2019 and disease status was assessed by investigators using RECIST 1.1. Independent review committee (IRC) assessed data were not available at the time of the analysis. The updated analysis confirmed the marked tissue-agnostic efficacy and long durability of response in patients with TRK fusion cancer treated with larotrectinib. Larotrectinib continued to demonstrate a favourable long-term safety profile. A more detailed analysis is planned for a later time (projected data cut late summer 2019).

We have attached the ESMO abstract, which should remain academic in confidence until after publication, as appendix 1.

A2. Priority question: Please supply fully anonymised individual participant data (subject to prior creation and approval of a suitable Data Sharing Agreement) for the latest data cut for all three included trials, including distinguishing the following patient groups:

i. **Patients with primary central nervous system (CNS) tumours**
ii. **Adult patients**
iii. **Paediatric patients**

The following variables are requested:

i. **Tumour type (specific site and NTRK fusion type)**
ii. **Line of therapy (including previous therapies received)**
iii. **Response (complete response, partial response etc.)**
iv. **Duration of response**
v. **Time of progression**
vi. **Time of death**
vii. **Censoring time**

Bayer does not have permission to share the patient level data with the ERG.

A3. Priority question: The ERG considers it essential that potential heterogeneity in efficacy is investigated. For the ePAS2 and SAS3 data (30th July 2018 data cut off, n=102), please provide subgroup data for the outcomes listed in Table 1 at the end of this document. Where feasible, please also provide Kaplan-Meier curves for progression-free survival and overall survival. Results data should be presented for the following subgroups:

i. **By tumour site**
ii. **By age: Adults (≥18 years) vs children/adolescents (<18 years)**
iii. **By overall response rate status: responders vs non-responders**
iv. **By response category (separately for complete response and partial response)**
v. **By fusion type: NTRK 1,2,3**
vi. **By isoform: ETV6-NTRK3, TPM3-NTRK1, LMNA-NTRK1**

As discussed during the clarification teleconference on 26th June. We do not believe that providing subgroup data is justified or helpful in terms of decision-making. There are two main reasons for this 1) based on the totality of the trial data there is no evidence of heterogeneity in treatment effect according to the subgroups listed 2) patient numbers are already small and further post-hoc 'slicing and dicing' of the data will only serve to increase uncertainty. We believe that provision of subgroup data only serves as a distraction and introduces the potential for decision-making to be based on chance findings.

The totality of the clinical and nonclinical body of evidence supports a tissue-agnostic/histology-independent indication since larotrectinib has demonstrated a large magnitude of effect irrespective of tumour site. We do not believe the uncertainty inherent to small datasets is improved by cutting the data further.

1) There is no evidence of heterogeneity in treatment effect according to the subgroups listed

Tumour site is not relevant for tumour-agnostic therapies

Tropomyosin receptor kinases (TRK) fusion cancer are among the first truly genetically defined cancer (Drilon et al. 2018), where tumour site of origin (i.e. histology) is a minor variable in the pathologic description of the disease. In respect of larotrectinib the site of the tumour is not relevant as the mechanism of action is **independent** of tumour site and is

entirely dependent on the presence/absence of NTRK fusion proteins i.e. if NTRK fusion proteins are present larotrectinib is effective and if they are not it is of no benefit. Treatment of TRK fusion cancer patients with larotrectinib exhibited rapid, substantial antitumor activity with durable disease control that appears to be independent of site (See Appendix E associated with main submission).

By fusion type and isoform

Table 1 presents the ORR for subgroups requested and shows the widely overlapping 95% confidence intervals. The anti-tumour activity appears to be independent of NTRK fusion type and isoform.

The NICE methods guide states that subgroup effects should be statistically robust if they are to be considered in a CE model, as well as having some *a priori* justification. "In practice it would be difficult to sustain and argument that a treatment should be accepted or rejected based on a statistically weak interaction".

Table 1. ORR for larotrectinib according to NTRK gene fusion or major NTRK gene isoforms (IRC, ePAS2) [reproduced from table 83 Appendix E]

	N	ORR, % (95% CI)
Fusion		
NTRK3	[redacted]	[redacted]
NTRK1	[redacted]	[redacted]
NTRK2	[redacted]	[redacted]
Isoform		
ETV6-NTRK3	[redacted]	[redacted]
TPM3-NTRK1	[redacted]	[redacted]
LMNA-NTRK1	[redacted]	[redacted]

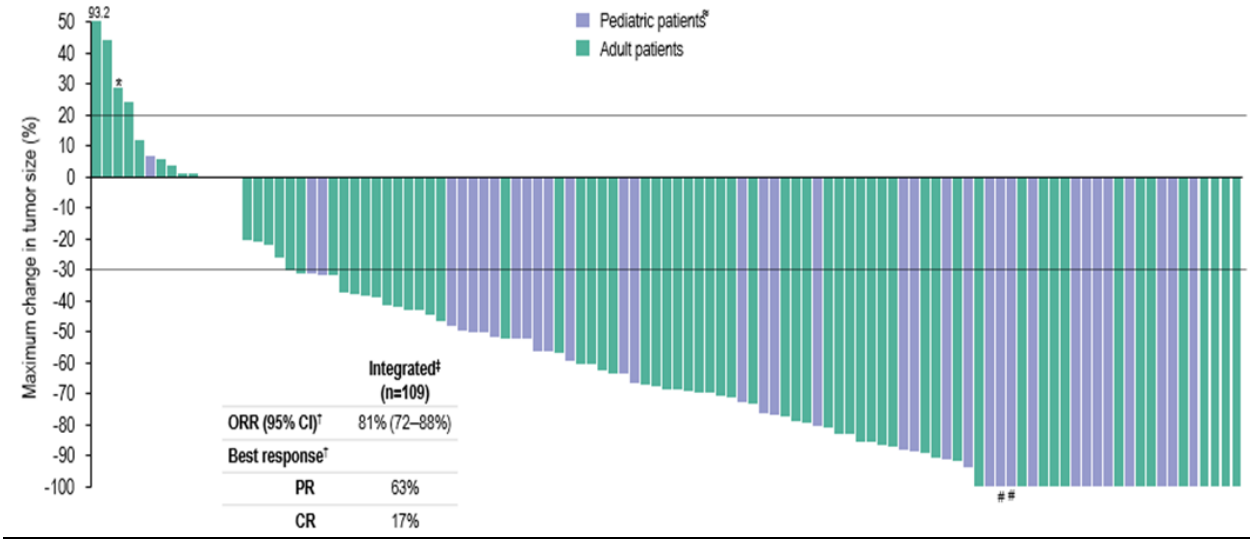
CR=complete response; ePAS=extended primary analysis set; PR=partial response; IRC=independent review committee; ORR=overall response rate.

By age

The efficacy of larotrectinib is independent of age:

- Figure 1 shows efficacy according to maximum change in tumour size and indicates no difference according to age
- Table 2 shows overall response rate with widely overlapping confidence intervals which do not support any difference in efficacy according to age.

Figure 1. Efficacy Results With Larotrectinib in the Integrated Analysis by Patient Age (Investigator Assessment)



- surgical CR

Table 2. ORR for larotrectinib according to patient age (reproduced from Appendix E)

Baseline characteristic		ePAS2	
		N	ORR, % (95% CI)
Overall		**93**	[redacted]
Age	**Paediatrics (<18 years)**	[redacted]	[redacted]
	1 month to <2 years	[redacted]	[redacted]
	2 to <6 years	[redacted]	[redacted]
	6 to <12 years	[redacted]	[redacted]
	12 to <18 years	[redacted]	[redacted]
	Adults (≥18 years)	[redacted]	[redacted]
	18 to <65 years	[redacted]	[redacted]
	≥65 years	[redacted]	[redacted]

A random effects logistic regression analysis for adults vs children/adolescents has not been performed. A statistical random effects model only makes sense if the variable for which a random effect is estimated is a variable which can be assumed to be randomly chosen from a population. It can be assumed to be true for patients study sites. But for adults vs children/adolescents, this assumption is not true: the result of combining adults with children/adolescents, is a full population already, and cannot be considered as a random drawing from a population.

By overall response or response category

We apologise if we are misinterpreting the question but we do not understand the definition of subgroups categorised by response to treatment as opposed to patient characteristics.

Random Effects model by tumour type

An assessment of heterogeneity by tumour type has been undertaken. The primary endpoint is overall response rate. Thus, assessment on heterogeneity was conducted using this endpoint.

Bayer has assessed heterogeneity in the primary endpoint, by tumour type for the ePAS2 population. A separate analysis for SAS3 has not been considered.

A random effects logistic regression model, with tumour type included as a normal-distributed random effect has been performed. Utilising the estimates of this model, a prediction for the distribution of ORR for ''not yet'' studied tumour types was generated. These results are displayed in Figure 1. This distribution indicates that an estimated [redacted]% of newly ''to be'' studied tumour types will express an ORR of 40% or higher (see Table 3 for further quantiles of the distribution).

A table displaying ORR by tumour histology has been provided in the Appendix E of the main submission.

Figure 2. Estimated distribution of probability of response in new tumour types (ePAS2)

Table 3. Distribution of various thresholds (t) for response probability (as determined by the random effects model) (ePAS2)

T	Approximative Probability for Response > t (*)
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

(*) calculated with trapezoidal rule

Bayer: /by-sasp/patdb/ia/2731953/stat/2019/test_emasag01/pgms/table_3005_est_distr_epas2.sas eopsf 07MAR2019 12:05

2) Patient numbers are too small for meaningful results in subgroups

Table 2 shows the small patient numbers in the trial overall and the even smaller numbers in the subgroups of interest – the largest subgroup being [redacted] patients and the smallest [redacted] patient. We believe that consideration of subgroup results would be perilous and should be avoided. We do not believe the uncertainty inherent to small datasets is improved by cutting the data further.

Table 2. Patient numbers by subgroup

Tumor Type	**N**
Overall	93
Soft tissue sarcoma	[redacted]
Salivary gland	[redacted]
Infantile fibrosarcoma	[redacted]
Thyroid	[redacted]
Lung	[redacted]
Melanoma	[redacted]
Colon	[redacted]
GIST	[redacted]
Bone sarcoma	[redacted]
Cholangiocarcinoma #	[redacted]
Appendix #	[redacted]
Breast #	[redacted]
Congenital mesoblastic nephroma	[redacted]
Pancreas #	[redacted]
Fusion	
NTRK3	[redacted]
NTRK1	[redacted]
NTRK2	[redacted]
Isoform	
ETV6-NTRK3	[redacted]
TPM3-NTRK1	[redacted]
LMNA-NTRK1	[redacted]
Age	
Paediatrics (<18 years)	[redacted]
Adults (≥18 years)	[redacted]

A4. Priority question: The total number of NTRK patients in the trials, as reported in Table 8 (CS), does not match the numbers in the ePAS2 and SAS3 sets in Table 9 (CS). This appears to be due to some safety analysis set patients being excluded from the efficacy analysis set (20 in NAVIGATE, 13 in SCOUT). Was this because disease could not be measured at baseline in these 33 patients?

i. **Page 62 (CS) states that analyses were performed according to the intention-to-treat principle. With this in mind please describe the rationale for excluding these 33 patients from the progression-free survival and overall survival analyses.**

ii. **To further clarify this, please provide CONSORT flow diagrams - separately for each of the three trials - illustrating the flow of participants from screening to inclusion in the analyses. Please provide: number screened for eligibility, number ineligible/excluded (with reasons), number who declined participation, number recruited into study, number who received at least 1 dose of larotrectinib, number who discontinued treatment (with reasons), and the numbers included and excluded in the overall response rate, progression-free survival and overall survival analyses (with reasons).**

iii. **Where data are available, please provide progression-free survival and overall survival results for the safety analysis set (at the latest available cut-off date).**

Exclusion of patients from ePAS2

The primary analysis set includes, as per the SAP for integrated efficacy analysis, the first 55 consecutively enrolled patients harbouring a solid tumour with NTRK fusion that were treated, had measurable lesion at baseline (as assessed by investigator) and had no primary CNS tumour These patients formed the PAS population, used for the primary evaluation performed on the July 2017 cut-off data, which ensured a follow-up of at least 6 months for these patients.

Separate from the PAS, three supplementary sets were also defined for patients treated with larotrectinib:

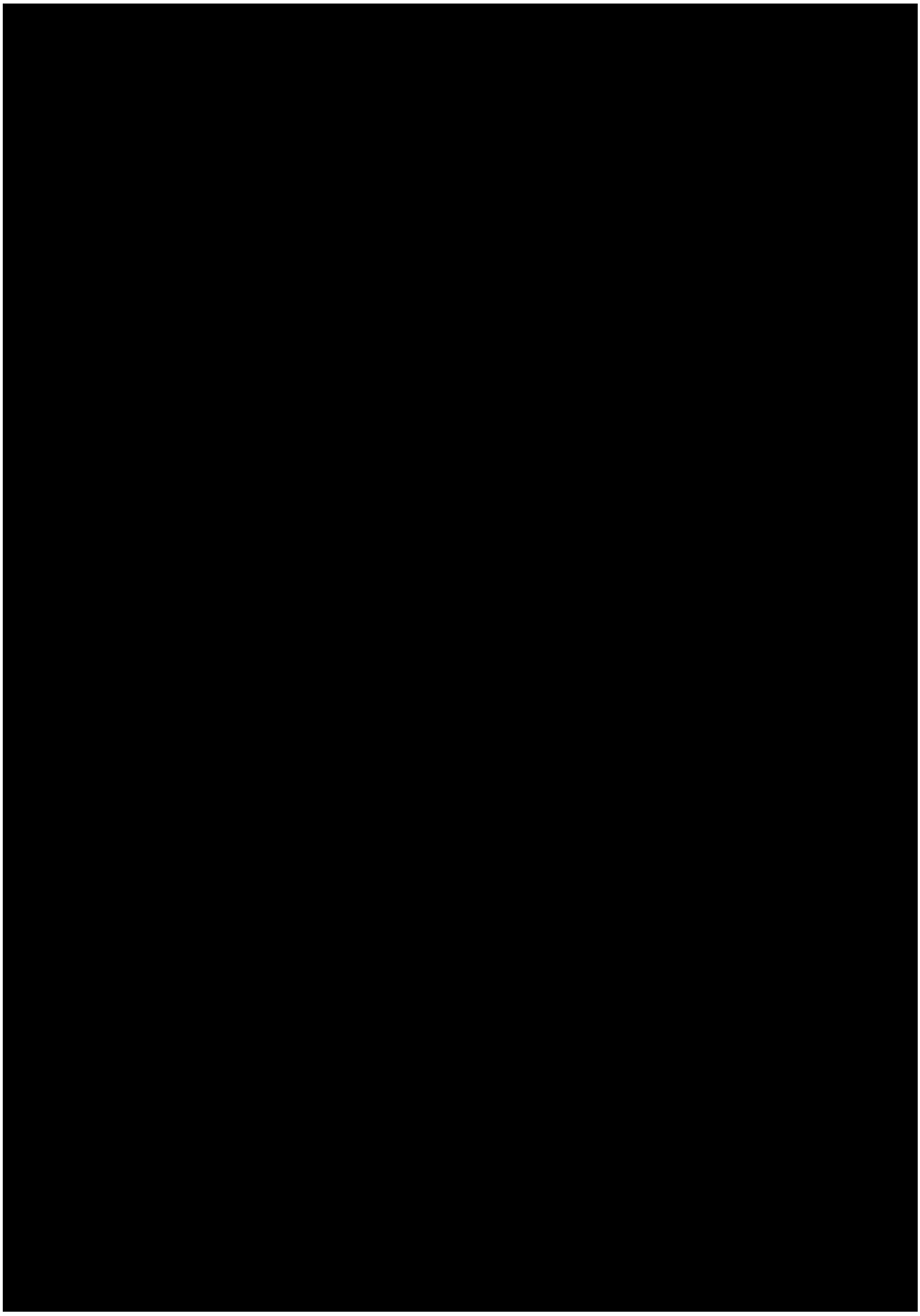
1. SAS1: Patients that fulfilled PAS criteria, but were enrolled after the initial 55 patients constituting the PAS population.
2. SAS2: Patients with solid tumour with NTRK fusion that had no measurable lesion
3. SAS3: Patients with primary CNS tumours.

The ePAS2 includes the initial 55 consecutively recruited patients of the primary analysis set (PAS) as well as 38 further patients from the SAS1 dataset, that were consecutively recruited before the February 2018 cut-off date. This analysis set was first analysed at the cut-off in July 2018, with the potential of 6 or more months of follow-up for the patients included.

An additional 9 patients were studied within the SAS3 dataset.

Patients in the SAS1 dataset that were recruited later than February 2018 (N=28) were not included into ePAS2. Together with patients in SAS2 (N=7), these were 35 patients. An updated Table 8 (and Table 79) have been provided below. Apologies for the mistake in the original submission.

Thus, altogether [redacted] NTRK positive patients have been studied, as is also outlined within Figure 19 below.



The primary endpoint of this study was overall response rate (ORR), analysed using the PAS. Related to ORR, the secondary efficacy variables time to response and duration of response were also analysed using the PAS.

As is common in clinical studies, and to allow for comparable results, with the primary efficacy analysis, the PFS and OS analysis were analysed using the identical population with no separate populations being defined.

The number of patients excluded from PAS/ePAS2 due to non-measurable lesions at baseline was low (SAS2 included N=7 patients at 30th July 2018 cut-off).

In order to ensure consistent minimum follow-up for the ORR endpoint comparable to what was defined in the SAP for the primary analysis based on PAS, the ePAS2 population had only included patients recruited until 19th February 2018.

No patients were excluded from SAS3.

Upon request by NICE, the PFS and OS analyses for the N=137 NTRK positive patients were performed and are described in response iii below.

ii.

In line with the teleconference on 26th June 2019, as the studies were originally in the hands of LOXO, Bayer do not have the full CONSORT flow diagrams at this time but can provide these at a later date.

iii.

The median PFS in the N=137 NTRK positive patients was [redacted] months (95% CI: [redacted]). See also Figure 1 and Table 101. The median follow-up for PFS is [redacted] months.

The overall survival is not yet mature, with an estimated [redacted] of patients surviving one year or longer (see Figure 2 and Table 102). The median follow-up in Overall survival is [redacted] months.

Figure 1: PFS evaluation in N=137 NTRK positive patients (cut-off 30 JUL 2018)

Figure 2: Overall survival for N=137 NTRK positive patients

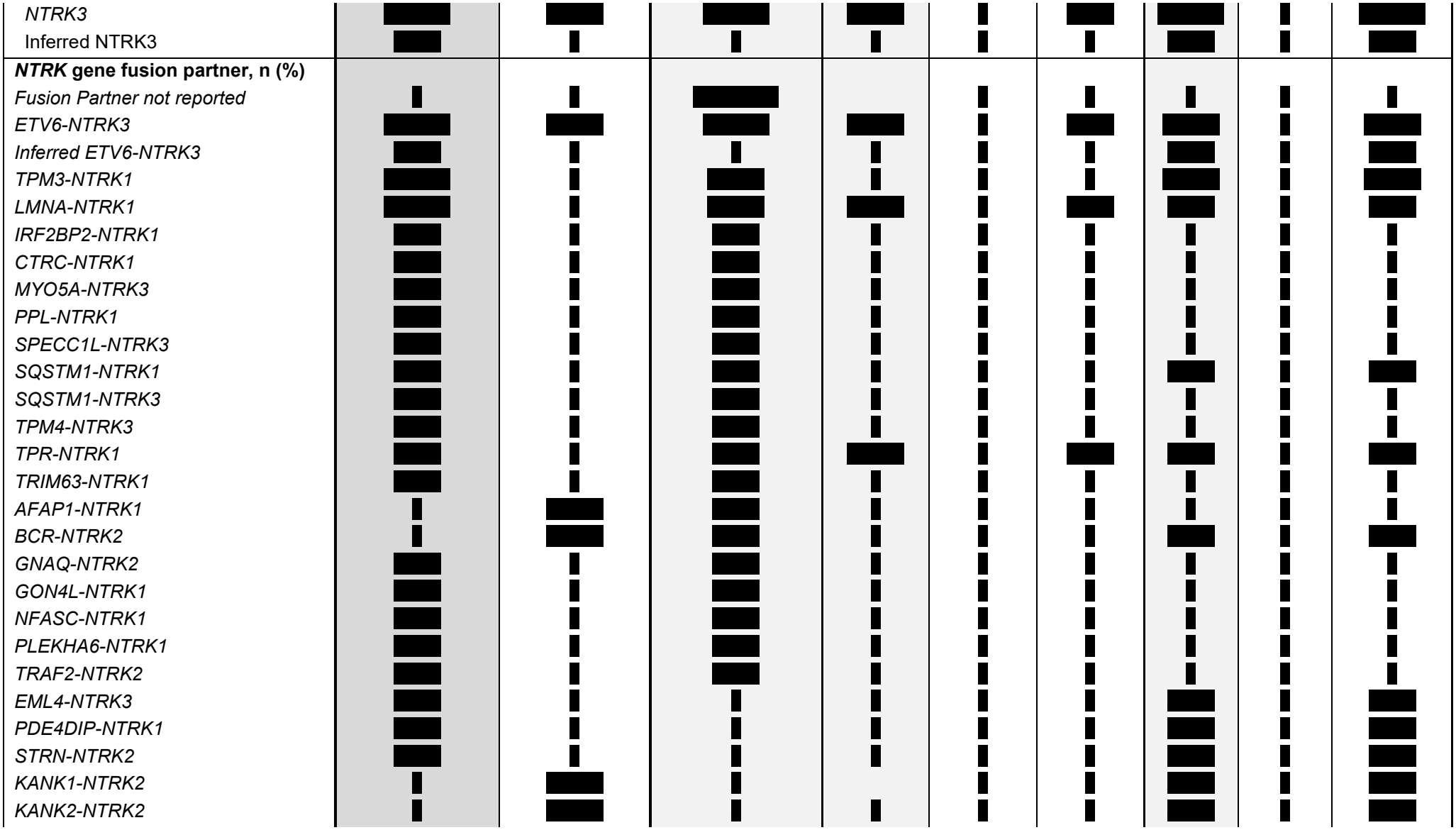
Table 8. Baseline demographic and disease characteristics for pooled analysis and individual larotrectinib study populations (efficacy evaluable patients)(data cut-off 30th July 2018) (45)

Trial number (acronym) Baseline Characteristics	Pooled Analysis of patients from LOXO-TRK-14001, SCOUT and NAVIGATE trials N=102		NAVIGATE LOXO-TRK-15002 NCT02576431 N=62	Adult Phase 1 LOXO-TRK-14001 NCT02122913			SCOUT LOXO-TRK-15003 NCT02637687		
	ePAS2 n=93	SAS3 n=9		NTRK N=8	Non-NTRK N=62	Total N=70	NTRK N=32	Non-NTRK N=9	N=41
Age, n (%)									
Median age, years (range)									
Mean									
< 2 yr									
2-<6 yr									
6-<12 yr									
12-<16 yr									
16-<18 yr									
18-<45 yr									
45-<65 yr									
65<75 yr									
≥ 75 yr									
Sex, n (%)									
Male									
Female									
Race, n (%)									
White									
Black or African American									
Asian									
American Indian or Alaska Native									

Trial number (acronym) Baseline Characteristics	Pooled Analysis of patients from LOXO-TRK-14001, SCOUT and NAVIGATE trials N=102		NAVIGATE LOXO-TRK-15002 NCT02576431 N=62	Adult Phase 1 LOXO-TRK-14001 NCT02122913			SCOUT LOXO-TRK-15003 NCT02637687		
	ePAS2 n=93	SAS3 n=9		NTRK N=8	Non-NTRK N=62	Total N=70	NTRK N=32	Non-NTRK N=9	N=41
Native Hawaiian or Pacific Islander	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Multiple / Other									
Declined to state/Not reported									
ECOG PS, n (%)									
0									
1									
2									
Not reported/unknown									
Primary tumour type, n (%)									
NSCLC									
IFS									
STS									
Colon									
Salivary gland									
Breast									
Pancreas									
Thymus									
Thyroid									
Bone sarcoma									
Cholangiocarcinoma									
Gastric									
GIST									
Hepatic									
Melanoma									
Anal									
Appendix									
Cancer of unknown primary									
Endometrial									
Larynx									
Neuroblastoma									
Oral									

Ovarian									
Primary CNS									
Renal									
Congenital mesoblastic nephroma									
Ewing sarcoma									
Other									
Stage at initial diagnosis, n (%)									
I									
II									
III									
IV									
Not reported/Unknown									
Disease extent at enrollment n (%)									
Locally advanced									
Metastatic									
Other / not reported									
Prior cancer therapy - Yes, n (%)									
Surgery									
Radiotherapy									
Systemic therapy									
0 prior systemic									
1-2									
≥3									
Mean no. prior systemic									
Median no. prior systemic									
***NTRK* gene fusion status, n (%)**									
None / not known									
NTRK1									
NTRK2									


NTRK3
Inferred NTRK3
NTRK gene fusion partner, n (%)
Fusion Partner not reported
ETV6-NTRK3
Inferred ETV6-NTRK3
TPM3-NTRK1
LMNA-NTRK1
IRF2BP2-NTRK1
CTRC-NTRK1
MYO5A-NTRK3
PPL-NTRK1
SPECC1L-NTRK3
SQSTM1-NTRK1
SQSTM1-NTRK3
TPM4-NTRK3
TPR-NTRK1
TRIM63-NTRK1
AFAP1-NTRK1
BCR-NTRK2
GNAQ-NTRK2
GON4L-NTRK1
NFASC-NTRK1
PLEKHA6-NTRK1
TRAF2-NTRK2
EML4-NTRK3
PDE4DIP-NTRK1
STRN-NTRK2
KANK1-NTRK2
KANK2-NTRK2

SPECC1L-NTRK2	
AGTPBP1-NTRK2	

CNS=central nervous system; ECOG=Eastern Cooperative Oncology Group; GIST=gastrointestinal stromal tumour; IFS=infantile fibrosarcoma; n=number; no.=number; NSCLC=non-small cell lung cancer; NTRK=neurotrophin receptor tyrosine kinase; PS=performance status; STS=soft tissue sarcoma; yr=year;
Note: due to rounding, percentages may not total to exactly 100%

[redacted]

Table 79. Patient disposition by study and pooled analysis sets (data cut-off 30 July 2018) (45)

Study	Pooled analysis		NAVIGATE [a]	LOXO-TRK-14001[a]			SCOUT [a]		
Disposition	**ePAS2 n (%)**	**SAS3 n (%)**	**NTRK n (%)**	**NTRK n (%)**	**Non-NTRK n (%)**	**Total n (%)**	**NTRK n (%)**	**Non-NTRK n (%)**	**Total n (%)**
Screened	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Enrolled and treated	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Disease progression	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Disease-free	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment ongoing, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment continued post-progression	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Discontinuation of treatment, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Disease progression	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Clinical progression	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Physician decision	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse event [b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patient decision	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-compliance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Protocol deviation	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Death	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Other [c]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] Safety analysis set

[b] Discontinuations due to AEs:

LOXO-TRK-14001 - [redacted]

NAVIGATE - [redacted]

SCOUT – [redacted]

[c] Included:

SCOUT – [redacted]

A5. Priority question: Please provide details of any treatments received after the development of larotrectinib resistance or after disease progression, including the number of patients who:

i. **Developed larotrectinib resistance**
ii. **Continued to receive larotrectinib**
iii. **Received LOXO-195**
iv. **Received other interventions not currently available or not recommended in the NHS.**

Please also provide the median (and interquartile range) or mean (and 95% confidence intervals) duration of larotrectinib treatment beyond the point of progression. Please also provide details of how patients who received these post-progression treatments were handled in overall survival estimation.

Of the 93 patients in the dataset, [redacted] had progressed at the time of data cut-off (30th July 2018). Mutations were identified as a mechanism for resistance in [redacted] patients. Overall, among the 93 patients included in the ePAS2 dataset, [redacted] continued to receive larotrectinib post-progression.

Available data show that the duration of treatment post-progression ranged from [redacted] to >[redacted] days (2 patients continuing to receive treatment). The median duration of post-progression treatment was [redacted] months. The mean in ePAS2 was [redacted] and in NTRK fusion cancers was [redacted] months (Table 1).

Table 1. Post-progression treatment

Bayer Protocols 20288, 20289, 20290
(Loxo Oncology Protocols 14001, 15002, 15003)
Integrated Summary of Efficacy (Visit Cutoff 30-JUL-2018)

Table 103: Summary of Treatment Duration after Progression

		ePAS2 (N=93)	NTRK Fusion Cancers (N=137)
Component of Treatment after Progression (months)	n	[redacted]	[redacted]
	Mean		
	SD		
	Min		
	Median		
	Max		

Bayer: /by-sasp/patdb/ia/2731953/stat/2019/prod_gbr01/pgms/table_103_trt_after_pd.sas ggmka 26JUN2019 15:43
End of table

Handling of patients who received post-progression treatment were handled in the survival analysis

Treatments received were not a consideration in overall survival analysis. Each patient enrolled in the studies contributed to OS analysis.

Patients receiving LOXO-195

Nineteen patients enrolled in the larotrectinib studies have gone on to receive LOXO-195.

Patients receiving other treatments not currently available or recommended in the NHS

These data is not available at this time. However, if patients go on to receive other interventions not currently available or recommended for use in the NHS after larotrectinib in the trials, Bayer would not expect to adjust for this in any cost-effectiveness analysis (cross reference to question B5). No previous oncology appraisals have been identified were analysis on such data has been deemed appropriate by NICE. With such rare tumours, with poor prognosis, despite treatments not being routinely available on the NHS, it is not unreasonable to assume that some patients would go on to receive further innovations as part of a clinical trial, compassionate access to medicines not yet licensed or drugs approved via a system of individual funding requests. This would be equally applicable to patients in the comparator arms. As such, the dataset would be reflective of expected clinical practice in the UK and in line with an ITT analysis.

Effect of post-progression treatment on cost-effectiveness

As patients would not be treated beyond progression in clinical practice, making an adjustment for it would not have an impact on OS. Hence such an adjustment was not part of the base case in the model.

However, in addition to the 'treat-to-progression' strategy presented in the base case, we have reviewed the potential impact of patients remaining on larotrectinib after progression. To inform this scenario, we took the average treatment duration from the two trial populations. Applying the cost of larotrectinib treatment for the average durations mentioned in the post-progression increased the ICER per QALY to £[redacted] when using the ePAS2 dataset estimate and £[redacted] when using the NTRK Fusion Cancers population estimate (range £[redacted] - £[redacted]). Note that this scenario assumed a fixed proportion of paediatric patients amongst all patients using the baseline split. In addition, this scenario did not consider the proportion of patients receiving larotrectinib after disease progression but rather applied costs to all patients, and therefore the real increase in the ICER is likely to be smaller than seen here.

A6. The results illustrated in Figure 1 (p22, CS) do not match those in Table 12 (p73, CS). In particular, Figure 1 has 2 surgical complete responses rather than the 1 in Table 12, and Figure 1 appears to show that around 25 patients had a CR (100% decrease in tumour size) compared to 15 patients in Table 12. Please explain the inconsistency of these results.

Apologies for the inconsistency. Figure 1 (the waterfall plot) was generated for conference purposes and is based on investigator assessment of all efficacy-evaluable subjects at the time of data cut-off (30th July 2018). Excluded from the plot are those more recently-enrolled subjects without post-baseline assessments.

For Table 12, the ePAS2 set for regulatory submission purposes was determined by pre-specified criteria of which subjects will have IRC assessments performed, i.e., those efficacy-evaluable subjects who have had the opportunity to receive at least approximately 6 months of treatment. That is, subjects who started treatment within approximately 6 months of the data cut-off (30th July 2018) are excluded from the primary integrated analysis at that time. These patients will be included in future planned analyses where subjects have had the opportunity to receive at least 6 months of treatment.

Regardless of the known investigator assessment at data cut-off (confirmed CR or PR response, or PD), such subjects starting treatment less than 6 months prior to data cut-off are not assessed by IRC until a future analysis.

It should also be noted that we cannot derive response rate from a waterfall plot as only the target lesions at baseline are taken into account.

In addition, the ePAS2 set for regulatory submission excluded a subject whose NTRK fusion status was not confirmed until after start of treatment.

A7. Please provide a version of Figure 1 (p22, CS) which presents results derived from independent review committee data.

Figure 1 was created for conference purposes and is based on investigator assessment of all efficacy-evaluable subjects at the time of data cutoff. Excluded from the plot are those more recently-enrolled subjects who do not yet have post-baseline assessments. A similar figure was not repeated for IRC assessments. However, we present below the waterfall plot based on IRC assessment for the ePAS2 dataset colour coded by tumour type.

Note that for 1 patient the histology was later updated from GIST to soft tissue sarcoma and is reflected as such in the plot below.

A8. Please provide Figure 19 (CS) categorised by tumour type and by adult/paediatric patients (or provide individual-level GMI data as per question A2.)

For the 53 patients included in the analysis, the waterfall plot for GMI across tumour types and age groups (adult/paediatric) is shown below.

Of the 53 patients, there were [redacted] adults and [redacted] paediatric patients. The majority of patients had a GMI ≥1 with [redacted] adult patients ([redacted]) and [redacted] paediatric patients ([redacted]) with GMI ≥1.33.

The waterfall plots for GMI by primary diagnosis indicates that results are consistent across the majority of tumour types and across adult and paediatric patients. Note that certain tumour types only have 1 patient each (appendix, bone sarcoma, and breast, and pancreas), so for those tumour types there is limited data from which to derive conclusions.

A9. If available, please provide more detail on the exact tumour types included in the analysis set. For example, is the breast cancer case secretory breast cancer? This could be included with data provided as per question A2.

Patient distribution by tissue histology and subtypes in the extended efficacy patient pool (ePAS2+ SAS3) plus primary CNS is summarised in Table 1.

Table 1: Patient distribution by tissue histology and subtypes (Extended Primary Analysis Set 2 + Primary CNS)

Tumour type	No. patients (N=102)
Soft tissue sarcoma[a]	[redacted]
Infantile Myofibromatosis	[redacted]
Inflammatory Myofibroblastic Tumor	[redacted]
Inflammatory Myofibroblastic Tumor Of Kidney	[redacted]
Lipofibromatosis	[redacted]
Myopericytoma	[redacted]
Nos	[redacted]
Peripheral Nerve Sheath	[redacted]
Spindle Cell	[redacted]
Spindle, Epitheliod	[redacted]
Infantile Myofibromatosis	[redacted]
Salivary gland[a]	[redacted]
Adenocarcinoma	[redacted]
Masc	[redacted]
Parotid	[redacted]
Parotid; Adenocarcinoma	[redacted]
Parotid; Adenoid Cystic	[redacted]
Parotid; Glandular, Sarcomatoid	[redacted]
Parotid; Mucoepidermoid	[redacted]
Infantile fibrosarcoma[a]	[redacted]
Thyroid[a]	[redacted]
Differentiated	[redacted]
Non-Differentiated	[redacted]
Primary CNS[b]	[redacted]
Astrocytoma	[redacted]
Glioblastoma	[redacted]
Glioma	[redacted]
Nos	[redacted]
Lung[a]	[redacted]
Non-Small Cell	[redacted]
Small Cell	[redacted]
Melanoma[a]	[redacted]
Colon[a]	[redacted]
Gastrointestinal stromal tumour[a]	[redacted]
Bone sarcoma[a]	[redacted]
Chondrosarcoma	[redacted]
Nos	[redacted]
Cholangiocarcinoma[a]	[redacted]
Congenital mesoblastic nephroma[a]	[redacted]
Cellular	[redacted]
Appendix[a]	[redacted]
Breast (non-secretory adenocarcinoma)[a]	[redacted]
Pancreas	[redacted]

a: Independent review committee analysis by RECIST 1.1
b: Patients with a primary CNS tumour were evaluated per investigator assessment using either RANO or RECIST v1.1 criteria

A10. In Appendix O, results by trial are presented for the February 2018 cut-off. Please provide the results for each respective trial for the July 30th 2018 cut-off (or a more recent cut-off) if available.

Limited analysis was performed for individual studies based on 30 July 2018 data cut and included ePAS2 and SAS3 datasets (table below). Analysis that also includes SAS2 dataset has not been conducted and therefore is not available at this point.

	NAVIGATE LOXO-TRK-15002 NCT02576431 (INV) N=62	**Adult Phase 1 LOXO-TRK-14001 NCT02122913 (INV) N=70**		**SCOUT LOXO-TRK-15003 NCT02637687 (INV) N=43**	
		NTRK N=8	**Non-NTRK N=62**	**NTRK N=32**	**Non-NTRK N=9**
Best overall response					
CR, confirmed	[redacted]	[redacted]		[redacted]	
CR, pending confirmation					
Surgical complete response				[redacted]	
PR, confirmed	[redacted]	[redacted]	[redacted]	[redacted]	
PR, unconfirmed					
Stable disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressive disease	[redacted]		[redacted]		[redacted]
Not evaluable	[redacted]				
Not determined	[redacted]		[redacted]		[redacted]
	[redacted]				
Number of patients evaluable (INV)	[redacted]	[redacted]	[redacted]	[redacted]	
ORR (CR+PR), n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI for ORR	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Duration of follow-up, months					
Median	[redacted]	[redacted]	[redacted]	[redacted]	
25th, 75th percentiles	[redacted]	[redacted]		[redacted]	
Duration of response, months, n (%)					

Median (Min, Max)	[redacted]	[redacted]	[redacted]	[redacted]	
≤ 6 months	[redacted]		[redacted]	[redacted]	
>6 months	[redacted]			[redacted]	
>12 months	[redacted]	[redacted]		[redacted]	
>18 months	[redacted]	[redacted]		[redacted]	
>24 months	[redacted]	[redacted]		[redacted]	

A11. Permitted treatments include palliative radiotherapy at specific disease sites (Table 5, CS). Please provide proportions of patients who received this treatment and the proportion by cancer site if available.

Given the small number of patients as well as trial designs, it was not practical to present the results by proportions of patients. Please find instead the results presented in the descriptive table below.

Please note that, while NAVIGATE (LOXO-TRK-15002) collected the information on palliative radiotherapy systematically with a dedicated CRF, in SCOUT (LOXO-TRK-15003) this information has been captured only in concomitant procedures. The LOXO-TRK-14001 protocol did not contain provisions to collect data on concurrent palliative radiotherapy.

Subject ID	Primary Diagnosis	Category	Dose	Dose Units	Location	Start Date	Start Study Day	End Date	End Study Day
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Subject ID	Primary Diagnosis	Description of Procedure	Reason	Outcome	Date	Study Day
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

A12. Figure 12 identifies [redacted] patients who have had surgery (during the studies); please provide data on the tumour types involved.

Please see below the breakdown of histology for the [redacted] patients who had surgery in Figure 12:

Histology	Number of Patients
[redacted]	[redacted]
[redacted]	[redacted]

A13. Please provide the individual clinical study reports for all three studies (NAVIGATE, SCOUT and LOXO-TRK-14001).

Bayer have previously provided the pooled data that was the basis of the EMA submission which was based on the July 2018 data cut off. CSRs for the individual studies for the July 2018 were not produced as the focus was on providing the pooled analysis for EMA. However, we have provided the interim CSRs for the three individual studies.

A14. Priority question: The review of comparators (Appendix D) describes a large number of inclusions. Only limited summaries of these results are given in the submission. Please provide further detail on how relevant comparator data were identified, specifically:

i. **The search terms used to identify appropriate guidance.**
ii. **Documentation of selection decisions and reasons for excluding potentially relevant guidance.**
iii. **Documentation on decisions regarding which summary data or survival curves to extract, where multiple choices were available.**
iv. **The complete systematic review report (Xcenda: reference 77, CS)**

In answer to points i and iv, in separate files, appendix 2, we have included:

- the full systematic review report, including search terms and strategies,
- the data extraction files
- study exclusions

The information in these files also answers question B14.

Comprehensive SLRs were conducted and updates run. The systematic reviews were commissioned and specified prior to knowledge of the likely summary of product characteristics wording and were conducted for the global organisation. They therefore contain data wider than needed for the NICE appraisal. The evidence generated from these reviews was too comprehensive to present in its entirety in the CS.

With the multiple tumour sites and multiple potential comparators and therefore the complexity of the submission, we adhered as far as possible to the NICE reference case applying a systematic algorithm, considering hierarchy of evidence, for all comparators and source data.

We apologise if our process was not described clearly and provide additional detail below.

A note on terminology used throughout the document: Proxy BSC refers to active treatments that are not deemed satisfactory (eg: not approved by NICE and/or not in guidelines or where clinicians have advised may be used in clinical practice but would be considered unsatisfactory) and that are used once all other lines of active treatments have been exhausted. Larotrectinib is expected to displace those as they could be considered not to be satisfactory. BSC refers to placebo arms with no active treatments.

Regarding points ii and iii

General approach to search and select NICE submission evidence

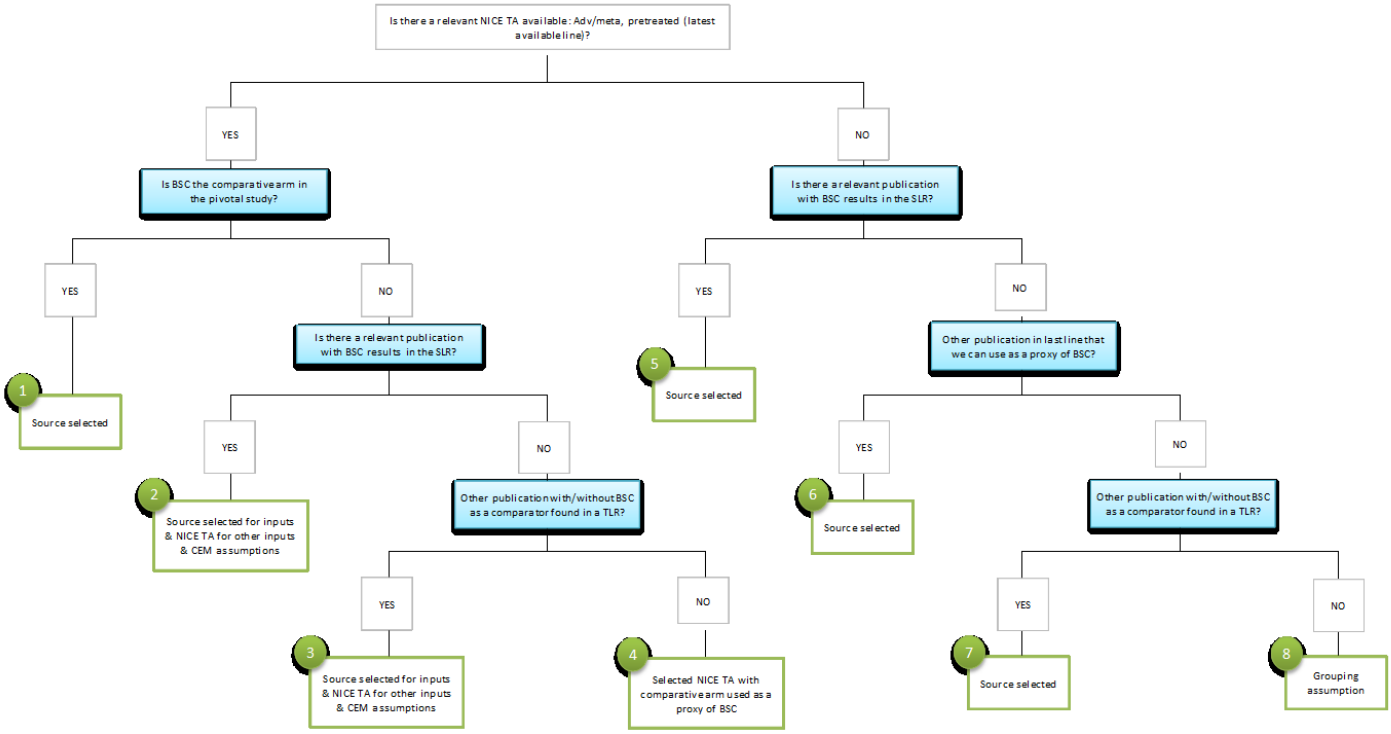
SLRs of clinical, economic and health-related quality of life evidence have been conducted for each tumour type represented in the larotrectinib trial programme. Due to the vast amount of evidence generated, across many tumour types, where possible we make use of past committee decision making. In some cases, the selection of the primary source to inform NICE submission was motivated by the following reasons:

- The SLR protocols were designed some time ago when the labelling assumption was tumour and line agnostic and when the initial scope was to search literature in the advanced/metastatic population regardless of the line of therapy to inform HTA submissions across countries (global project).
- Given the high number of tumour types, the list of comparators and the search terms for each tumour type was limited to the specific histologies in the studies, to make the SLRs manageable, resulting in a limited number of the publications eventually used in the model not being captured by the SLR.
- NTRK gene fusion is a novel target and there are no existing treatments for patients with NTRK gene fusion cancer hence no other additional studies for NTRK gene fusion cancer to inform the submission.

For all these reasons and due to the unusually high number of tumour types to consider for the cost-effectiveness model, and in order to ensure the model includes robust inputs and assumptions, Bayer prioritised the most recent NICE TA sources, where multiple TAs were available.

The overall search and selection strategy is illustrated in the figure below.

Summary overall search and selection approach



Given the process and scrutiny undertaken in each technology appraisal to select the Committee's preferred inputs and assumptions these sources were determined to be most suitable for decision making in England, and allowed the data and assumptions used in the model to reflect the Committee's preferred assumptions. This minimises uncertainty, and allows incorporation of input from the wide range of stakeholders who contributed to previous appraisals. When several sources met the selection criteria, further elements were taken into consideration to select the most appropriate source:

1. If NICE TA is considered, the appropriateness of the identified NICE TA were judged based on :
    - ✓ the trial comparator arm used as a proxy of BSC accepted by ERG/NICE
    - ✓ the extent to which the publication matches the treatment criteria as applied in the larotrectinib trial protocol,

- ✓ source with the most advanced patients (e.g. last line of systemic therapy)

2. Date of publication prioritising more recent publications. If a publication is considered, the appropriateness of the identified publications in the SLR and targeted searched were judged based on:
   - ✓ the extent to which the publication matches the treatment criteria as applied in the larotrectinib trial protocol,
   - ✓ source with the most advanced patients (e.g. last line of systemic therapy)
   - ✓ date of publication prioritising more recent publications, and reports outcomes that inform the health economic model inputs.

Comparators and main sources considered for each tumour type

Main source	Tumour types	Comparator
NICE TA	Thyroid	BSC
	GIST	BSC
	CRC	BSC
	NSCLC*	BSC
	Pancreas*	5FU + LV Proxy
	Melanoma	Investigator choice of chemo
	Breast	Treatment of physician's choice
	STS non GIST	Historical control data (TA185)
Publication (SLR)	Cholangiocarcinoma*	Gemcitabine + Cisplatin
	Salivary gland	Cisplatin + vinorelbine
Publication (TLR)	STS paediatrics	Irinotecan + vincristine
	Gliomas/CNS*	Lomustine
Other tumour used as a proxy	IFS & CMN	Irinotecan + vincristine
	Bone sarcoma	Historical control data (TA185)

(tumour groupings)	Appendix	BSC

**Mixed sources. Some tumour types have input coming from different sources. Tumour types are classified under the main data source*

Justification NICE TA/publication choice for tumour types

Tumour types	BSC source	Justification
Thyroid	NICE TA 535 + NICE Excel model	1) Latest line of active therapy with NICE TA is 2L 2) Only one MTA in this population (Lenvatinib and sorafenib) 3) Both product with BSC as comparator 4) Final choice: TA535 (sorafenib), Bayer had access to patient-level data which allowed replication of the committees preferred assumptions within CE model.
GIST	NICE TA 488 + NICE Excel model	1) Latest line of active therapy with NICE TA is 3L 2)Only one STA in this population is TA 488 (Regorafenib) 3) BSC as comparator 4) Final choice : TA 488 (regorafenib)
CRC & Appendix (for grouping explanation see B10 and appendix M)	NICE TA 405	1) Latest line of active therapy with NICE TA is 3L 2)Only one STA in this population is TA 405 (Trifluridine–tipiracil) 3) BSC as comparator 4) Final choice : TA 405 (Trifluridine–tipiracil) 5) In the absence of an identified clinical source for patients with appendix cancer and the low number of patients enrolled in the larotrectinib clinical trial programme with

		an appendix tumour site (n=1), the decision was made to use the colorectal cancer cohort as a proxy. Please see Appendix M for further details.
NSCLC*	NICE TA 374 + Sheppard 2005	1) TA374 included patients who had progressed following prior chemotherapy. The population considered was EGFR status unknown. These were presumed to be most representative of patients that would be eligible for larotrectinib (as to opposed to another biomarker source EGFR+, ROS1 etc) 2) The placebo arm of the Sheppard 2005 (cited in TA374, no active treatment) was used to represent current patient management. The study was also identified and selected in a recent NICE technology appraisal (TA483) to represent standard of care for patients who had exhausted available chemotherapy options

Justification NICE TA/publication choice for tumour types where satisfactory treatment options have been exhausted

For some tumour types, no treatment could be identified where patients had exhausted all satisfactory treatment options. In those cases, a search for sources of BSC (placebo arms in trials) was conducted. In other cases, it appeared that there were active treatments options that were not considered satisfactory (e.g. not approved by NICE and/or not in guidelines). When no available BSC source (placebo arms in trials) met the search criteria mentioned in in the previous section, the alternative approach was to expand the search to an active treatment that could be used as a proxy of BSC (active treatments that are not deemed satisfactory (not approved by NICE and/or not in guidelines)) (e.g. last line active treatment) following the same process.

When several sources met the selection criteria, further elements were taken into consideration to select the most appropriate source.

Tumour types	Treatment	Justification
Pancreas	5FU + LV (TA440)	1) Pegylated liposomal irinotecan 2nd line: Latest line of therapy with NICE TA. No NICE TA with more advanced population. 2) In the absence of BSC sources found in previous NICE TA or in the literature Bayer considered the comparative arm in Pegylated liposomal irinotecan clinical trial (5FU + LV) as a clinically relevant proxy of BSC
Melanoma	Investigator choice (IC) of chemo including dacarbazine, temozolomide,	1) NICE TA using the comparative arm as a proxy of BSC (TA357: IC chemotherapy and TA268: gp100 vaccine) Final choice : IC chemotherapy

	carboplatin, paclitaxel, or carboplatin+paclitaxel (from KEYNOTE-002 in NICE TA 357)	2) more representative of current clinical practice compared to gp100 vaccine, more likely to be used 3) IC chemotherapy as a proxy of BSC accepted by NICE 4) TA357 is the most recent
Breast	Treatment of physician's choice (TA423)	1) Eribulin 3rd line: latest line of therapy for breast cancer, adv/meta, with a NICE TA (TA423). No NICE TA with a more advanced population. 2) In the absence of BSC sources found in previous NICE TA or in the literature Bayer considered the comparative arm in Eribulin clinical trial (cisplatine + vinorelbine) as a clinically relevant proxy of BSC
STS non GIST	Historical control data (TA185)	1) Trabectedin 3rd line: latest line of therapy for STS, adv/meta, with a NICE TA (TA185). There is no NICE TA with a more advanced population 2) Historical control data considered to be equivalent to best supportive care (BSC; see sections 3.7 to 3.9 Company submission). These data were derived from studies in the database of the European Organisation for Research and Treatment of Cancer Soft Tissue and Bone Sarcoma Group (EORTC STBSG) 3) Approach accepted by NICE despite limitations (See section 4.4 NICE TA 185) 4) Please note an alternative has been considered after being raised as potentially also valid in a clinician interview, and a scenario analysis has been conducted.

Cholangiocarcinoma (or biliary tract cancer)*	Gemcitabine + Cisplatin	1) No specific NICE guidance was available for Cholangiocarcinoma and clinical evidence for pre-treated advanced/metastatic patients is limited and there is no established SoC (Valle 2010). 2) Based on ESMO, EASL and UK guidelines and identified within the SLR, gemcitabine + cisplatin represents a common treatment for adv/metastatic pre-treated patients based on a comparative phase II study (Valle 2010) 3) In the absence of BSC sources found in previous NICE TA and in the literature as well as the lack of established SoC in patients with adv/meta pre-treated disease, Bayer has considered the regimen Gemcitabine + Cisplatin as a clinically relevant to represent a proxy of BSC (Valle 2010)
Salivary gland	Cisplatin + vinorelbine	1) No specific NICE guidance was available for salivary gland tumours. Further, a review of the ASCO and NCCN guidelines along with manual searches confirm the lack of established SoC. 2) Cancer research UK recommends cisplatin, carboplatin, cyclophosphamide, doxorubicin, methotrexate and or paclitaxel alone or in combination. 3) Cisplatin + vinorelbine is a common treatment for recurrent salivary gland cancer based on a 2001 phase II study (Airoldi 2001). 4) In the absence of BSC sources found in previous NICE TA and in the literature as well as the lack of established SoC in patients with adv/meta pre-treated disease, Bayer has considered the regimen Cisplatin + vinorelbine as a clinically relevant to represent a proxy of BSC (Airoldi 2001)

STS paediatrics	Irinotecan + vincristine	1) No specific NICE guidance was available for STS paediatrics and clinical evidence for pre-treated advanced /metastatic paediatric patients is limited 2) Based on NCCN guidelines and targeted literature search of commonly used treatments Irinotecan + vincristine represents a treatment commonly used for pre-treated advanced/metastatic paediatric patients based on a 2010 phase II study (Mascarenhas 2010). In clinical validation the irinotecan + vincristine arm was considered to reflect patients that had failed previous therapies, in line with the anticipated marketing authorisation, and reflective of the patients enrolled into the larotrectinib clinical trial programme. This was accepted as a valid approach in the clinical expert validation. Please see Appendix M for further details.
Gliomas/CNS*	Lomustine	1) One NICE TA assessing temolozomide for patients with recurrent gliomas is available. However the comparative arm was not considered appropriate to extract survival and adverse event data for the CEM as procarbazine alone is rarely used in the UK for adv/meta recurrent disease, PCV or lomustine being considered as more appropriate (NICE TA 23 See section 4.1.3) 2) The SLR did not identify a suitable source data. 3) In the absence of BSC sources found in previous NICE TA and in the literature, Bayer has considered the single agent lomustine as a clinically relevant to represent a proxy of BSC as opposed to a multiple agents regimen usually used in earlier lines. 4) Two publications were found with outcomes in a similar population to larotrectinib patients profile (advanced/metastatic) recurrent gliomas already pretreated with systemic therapy): Batchelor 2013 (inclusion criteria: recurrent Glioblastoma, previous

		RT, previous treatment with temolozomide) and Wick et al 2010 (inclusion criteria: recurrent glioblastoma, previous RT and chemo, <2 prior systemic treatments). 5) Batchelor 2013 has been chosen over Wick et al 2010 because this study includes patients pre-treated with temolozomide like in the larotrectinib trials and doesn't restrict the number of previous line of systemic therapies like Wick et al 2010.
IFS	Irinotecan + vincristine	Assumption: IFS pooled with STS paediatrics. STS paediatrics comparator used as a proxy of BSC. It was determined through clinical validation that patients in the studies identified in the SLR and their outcomes were not representative of those enrolled in the clinical trial programme. Thus, following clinical advice, as a type of soft tissue sarcoma, IFS has been grouped with the paediatric STS patients to reflect a treatment relapsed population. This was accepted as a valid approach in the clinical expert validation. Please see Appendix M and B10 for further details.
CMN	Irinotecan + vincristine	Assumption: CMN pooled with IFS paediatrics, hence pooled with STS. STS paediatrics comparator used as a proxy of BSC
Bone sarcoma	Historical control data (TA185)	Assumption: Bone sarcoma pooled with STS adults. STS adults comparator used as a proxy of BSC

Manual search: Justification for source publications not being picked up in the SLR

Tumour type	Publication (input)	Justification
Salivary gland	Liberato 2012 (utility)	No relevant papers were identified in the SLR that considered specifically Salivary Gland patients. The patient population search terms were restricted to salivary gland (and very specific synonyms for subtypes/locations of salivary gland tumours), and this publication reports on the much broader disease of “head and neck cancer”. Please note that a lot of “head and neck cancer” publications are caught in the searches because salivary gland cancer is often grouped as such. However, this is also a function of how the articles are indexed. This one must not be indexed in the same way (eg mapped to salivary gland cancer in any way, either through keyword, MeSH, or Emtree).
STS paediatric	Mascarenhas 2010 (dosing, response status, AE, survival)	No relevant studies were identified in the SLR. This trial is specific to rhabdomyosarcoma, which is not one of the 7 STS subtypes included in the reviews. Searches were specific to the histologic subtype in each review. This was also explored in clinical validation interviews.
	Zuluga-Sanchez 2018 (utility)	No relevant studies were identified in the SLR. This is an STS model that is not specific to one of the 7 STS sub types investigated. Searches were specific to the histologic subtype in each review.

	Delea 2014 (utility)	No relevant studies were identified in the SLR. This is an STS model that is not specific to one of the 7 STS sub types investigated. Searches were specific to the histologic subtype in each review.
	Amdahl 2014 (Health state cost)	No relevant studies were identified in the SLR. This is an STS model that is not specific to one of the 7 STS sub types investigated. Searches were specific to the histologic subtype in each review.
Pancreas	Swinburn 2010 (disutility)	No relevant studies were identified in the SLR. The pancreas cancer review was restricted to studies of pancreas tumours only. Since this is a study of metastatic renal cell carcinoma, it was not caught in the searches. This study was used in the pancreas NICE TA so was included in the analysis

Rationale for tumour types groupings assumptions

A summary of the rationale is below. Further information on groupings is presented Appendix 3 and in response to question B10.

Tumour type	Grouping	Justification
IFS	With STS paediatrics	1) No relevant NICE TA in the relevant population (IFS, adv/meta/ last line) 2) No publication in the relevant population (IFS, adv/meta, last line) found to inform the NICE submission. Infantile fibrosarcoma is a type of soft tissue sarcoma. 3) In the absence of relevant source to inform the model, the low number of patients and the fact that IFS is a subtype of soft tissue sarcoma, IFS has been pooled with STS paediatrics in the CEM. 4) This grouping was confirmed as a valid approach by a clinical expert.

CMN	With IFS	1) No relevant NICE TA in the relevant population (CMN, adv/meta/ last line) 2) The SLR did not uncover suitable source data. 3) No publication in the relevant population (CMN, adv/meta, last line) found to inform the NICE submission. 4) However, cellular CMN are histologically similar to IFS and they also share cytogenetic abnormalities (Whittle et al 2010). 5) In the absence of relevant source to inform the model, the low number of patients and the similarity with IFS, CMN has been grouped with IFS in the CEM. 6) This grouping was confirmed as a valid approach by a clinical expert.
Bone sarcoma	STS non GIST	1) No relevant NICE TA in the relevant population 2) The SLR did not uncover suitable source data. 3) In larotrectinib trial, two patients with available results. 1 comes from NAVIGATE (adult) and the 2nd comes from SCOUT but she is a young adult (>20yrs) and has received several previous lines of systemic therapy including trabectedin (recommended by NICE for STS patients) 4) Decision to pool with STS patients has been taken in the absence of relevant source to inform the CEM as one of the bone sarcoma patients has received trabectedin in the course of her treatment

		5) In the absence of relevant source to inform the model, the low number of patients and similar outcomes (confirmed by clinical expert opinion), bone sarcoma has been grouped with STS in the CEM
Appendix	Colorectal	1) No relevant NICE TA in the relevant population (Appendix, adv/meta, last line) 2) The SLR did not uncover suitable source data. 3) No publication in the relevant population (appendix, adv/meta, last line) found via manual search to inform the NICE submission. 4) Appendix cancers are rare tumours that represents 1% of all diagnosed CRC (Tejani et al 2014). Appendix treatments and outcomes are similar to those seen in CRC (Tejani et al 2014) 5) In the absence of relevant source to inform the model, the low number of patients and the similarity with CRC, Appendix has been grouped with CRC in the CEM.

Documentation on decisions regarding which summary data or survival curves to extract, where multiple choices were available.

Detailed description of data sources and rationale is presented in Appendix M by tumour location. We are unsure what additional information is needed. We would be happy to set up a TC to discuss further if it would be helpful.

A15. Priority question: When comparators are considered for the economic model (Appendix M) there is only a single STA for each tumour type. This does not accord with the range of evidence in Appendix D, and that most tumour types have numerous potentially relevant STAs. Please justify why those particular STAs were selected, providing detail as in question A14.

Please refer to response to A14.

A16. Priority question: For some tumour types placebo or best supportive care is used for the survival curves, for others it's an active therapy. Please justify the comparator treatment selected for each tumour type.

Please refer to response to A14.

A17. Please provide more statistical detail on how the average comparator survival curves (Figure 30) were constructed from the curves for each tumour type.

The comparative survival results were based on the independently modelled PFS and OS survival curves for each of the tumour locations, which were then weighted by the trial-based tumour distribution for the purpose of presenting the pooled comparator results for LYs and QALYs.

To model survival on the tumour level, each tumour location was assigned its own Markov engine in the model. For each tumour location, the probability of survival was tracked using a partitioned survival approach using 7-day cycles as described in the CS. The average comparator survival curve for the pooled comparator (as presented in Figure 30 of the CS) was estimated by weighting survival across all tumour locations for each cycle according to the trial-based tumour distribution (Table 32 of the CS).

Note that the pooled comparator outcomes for LYs and QALYs as reported in the model results sheet are based on the individual engine results weighted by the trial-based tumour distribution. Figure 30 of the CS presented in order to show a representation of the pooled comparator survival data.

Bayer trust the above adequately answers the ERG's question. However, should more statistical detail be required, Bayer would suggest a TC where we can provide further explanations as needed.

A18. Priority question: The anticipated marketing authorisation for larotrectinib states that patients will be eligible for treatment where there are "[redacted]." Please expand on what constitutes [redacted], where larotrectinib will be placed as a line of therapy, and how this might vary by tumour type.

The final wording of the anticipated marketing authorization is subject to change at this time, however, Bayer can provide information regarding the ongoing discussions with EMA.

[redacted].

A19. According to the anticipated marketing authorisation for larotrectinib, patients are eligible for treatment depending on the presence of a NTRK fusion regardless of tumour site. However, the larotrectinib studies did not collect data on the full set of tumour sites that can present NTRK fusions. Please comment on the generalisability of these data, detailing which tumour sites were not covered by the trials, and referring to any other evidence available for these tumour sites (e.g. in terms of prognosis). Please comment also on the potential implications for the clinical and cost-effectiveness of larotrectinib.

The rarity of TRK fusion cancer and that *NTRK* gene fusions are found across many different tumour types are well recognized. The *NTRK* gene fusions seen in different types of cancers are similar, involving the carboxy-terminal kinase domain of TRK and various upstream amino-terminal partners leading to overexpression of the chimeric protein, resulting in constitutively active, ligand-independent downstream signaling (Vaishnavi et al. 2015), (Drilon et al. 2018). Larotrectinib is a precision medicine which specifically binds the protein product of the *NTRK* gene fusions, agnostic to the histology of the tumour.

A wide range of tumour types were represented in the clinical program. The response rate observed across the different histologies supports the use of larotrectinib in a histology agnostic population. Treatment of TRK fusion cancer patients with larotrectinib exhibited rapid, substantial antitumor activity with durable disease control that appears to be independent of *NTRK* isoform, tumor type and patient age. There was no effect in patients without TRK fusion cancer, irrespective of tumor type. This is not surprising given the mechanism of action of larotrectinib as a potent and selective inhibitor of TRKA, TRKB, and TRKC.

In terms of ongoing data collection, it is worth noting that the NAVIGATE and SCOUT clinical studies are still recruiting, so data for more patients and potentially more tumour histologies will be available over the next couple of years.

Research into precision medicines and specifically epidemiology data for NTRK-fusion positive patients are still not widely reported. In order to identify the frequency of NTRK-fusion patients in the general population Bayer conducted a SLR on the epidemiology of NTRK gene fusion in solid

tumours (see appendix 3). A meta-analysis was performed to provide pooled NTRK fusions rates. No tumour types were identified which were not represented in the larotrectinib clinical studies.

Sensitivity analyses were conducted to investigate the impact of varying the frequency of tumour types on cost effectiveness. The results seem to suggest that this does not have a significant impact on the overall cost-effectiveness. See question B10.

A20. Please provide details of the testing used to identify NTRK fusions in the included trials (e.g. immunohistochemistry (IHC) tests and next-generation screening), and how this varied across tumour types. Please comment on whether this screening approach (or any other approach) would be plausible in the UK, were larotrectinib to be approved.

NTRK gene fusion testing in studies 14001, 15002, or 15003 was conducted using a variety of analytical assays such as NGS, FISH or RT-PCR. Although no central assay was used, to be acceptable for use as an inclusion criteria, the assay had to be run in a CAP/CLIA (or equivalent) laboratory to ensure high technical quality, and every pathology report was reviewed at the time of enrolment to assess the methodology and the type of fusion for each patient. It was up to the local clinical practice to decide which test should be used for each type of tumour.

In 102 patients (93 ePAS2 + 9 SAS3) enrolled with NTRK gene fusion, non-CNS primary tumour and received at least 1 dose of larotrectinib: NTRK gene fusion was confirmed by NGS (N=[redacted]); FISH (N=[redacted]); PCR (N=[redacted]).

The range of testing methodologies utilised in the trials is compatible with the general approach set out in the national test directory. Most pertinently, the shift to NGS panel testing establishes a testing regime that will detect a wide range of genetic biomarkers, including NTRK 1-3 fusions:

> "In cancer, it is anticipated that wherever possible genetic testing will be delivered using panel testing rather than individual gene / variant tests with a move towards larger panels and WGS as the number of actionable targets increases for any specific tumour type"

We know from discussions with senior stakeholders at Genomic Laboratory Hubs that the hubs have been tasked with adopting and validating panels for solid and haematological cancers, and that when developing these panels they will cover genes that are of interest in research as well as those used in current clinical practice (i.e. a reimbursed actionable target and/or relevant to a full diagnosis).

From the publically available information and our interactions with NHS stakeholders it is our understanding that NTRK testing is available through a comprehensive range of testing

methodologies and whilst NGS panel testing will be the most widely-used method, in line with the testing methodologies in the trials, it is clear that as long as the results are obtained through a validated test, the screening approach is testing methodology agnostic.

i) https://www.england.nhs.uk/publication/national-genomic-test-directories/ (accessed 1st July 2019)

A21. Please justify assumptions about wider healthcare benefits of whole genome sequencing and the contribution of NTRK gene fusion treatments to these benefits (Table 1, CS). Please comment on the current availability of whole genome sequencing and other potential methods of NTRK gene testing in NHS clinical practice.

The 100 000 Genomes Project, completed in 2018, focused on cancer and rare disease, harnessed advances in whole-genome sequencing to improve diagnoses, inform targeted treatments and drive clinical research. (i) When reflecting on the success of the project, the Chief Scientific Officer, Dame Sue Hill, reaffirmed the wider healthcare benefits of WGS for diagnosis and treatment of patients in the clinical setting and how the scale of the project generated a vast evidence base to drive forward the discovery of new treatments and care approaches(ii). In announcing its extension, with the goal to sequence 1 million whole genomes by 2023, the Secretary of State for Health and Social Care reiterated the wider benefits of comprehensive genomic testing and stated that:

> "I'm incredibly excited about the potential for this type of technology to improve the diagnosis and treatment for patients to help people live longer, healthier lives – a vital part of our long-term plan for the NHS." (ii)

Larotrectinib is just one of a growing group of personalised medicines that can improve outcomes for certain patients identified as having the right genetic profile for the corresponding intervention. We have engaged with NHS Genomic Lead Hubs (GLHs) that are implementing the use of WGS (and panel testing) in a broad range of cancers in line with the national test directory. NHS England has been very clear about its strategic objectives and implementation of the National Genomics Medicine Service is progressing. NHS England has procured the GLHs, created the national directory and revised it at least once and it is currently working on the detailed contracting arrangements for the Hubs and sub-contracted partners. We have heard from stakeholders that this process is not without its challenges but it is definitely progressing.

Currently, WGS is listed in the National Genomic Test Directory for Cancer against 109 specific paediatric and adult solid tumour cancers and NGS panel testing is listed against 100 specific paediatric and adult solid tumours (there are also listings for WGS and NGS for haematological cancers but we have not included specific numbers as this sits outside the licence for Larotrectinib). The availability of WGS and panel testing for the majority of cancers is made clear in the 'notes' to the National Genomic Test Directory as they state that:

> "In cancer, it is anticipated that wherever possible genetic testing will be delivered using panel testing rather than individual gene / variant tests with a move towards larger panels and WGS as the number of actionable targets increases for any specific tumour type"

As part of the National Genomics Services, the Genomic Laboratory Hubs (GLHs) have been tasked by NHS England to develop / adopt a 'solid tumour panel'. Bayer has been notified by leading pathologists in several NHS GLHs that the panels being developed will contain target genes associated with those therapies currently reimbursed, and a degree of future proofing. i.e. the panels will test a wider range of genes than those which are currently actionable so that the panel will not need continual updating. On this basis, we believe that NTRK 1-3 is already being tested within existing gene panels or will be included in any panels under development and subject to validation. For example, ThermoFisher's Oncomine panel tests cover 161 genes in their 'comprehensive panel' and 52 in their 'focus panel' including NTRK 1-3 as fusion drivers in both panels. Similarly, FoundationOne CDx and multiple Ilumina cancer panels cover NTRK1-3 [iii-v]

It is clear that whether panels covering a wide range of genes or WGS testing is implemented, NHSE is not developing a national service solely for treatment with larotrectinib or any other precision medicine in development.

The National Genomic Test Directory also lists individual gene / variant testing for NTRK3 for congenital paediatric mesoblastic nephroma (NTRK3-ETV FISH/RT-PCR and NTRK3 rearrangement FISH). The notes in the test directory recognise that there may be circumstances where it is necessary to utilise multiple tests:

> "In cancer, those clinical indications listed as being eligible for whole genome sequencing can have this performed in parallel to the current standard of care testing in the directory."

From the publically available information and our interactions with NHS stakeholders it is our understanding that NTRK testing is available through a comprehensive range of testing methodologies and whilst panel testing (including NTRK amongst a wide range of genes) will be the most widely-used method, there will be a longer-term move towards widespread use of WGS for many cancer types. Therefore, the NHS is implementing a testing system that will identify NTRK fusion positive patients but this is only one part of a more complex and wider picture that is being driven by NHS England's strategic cancer and life sciences objectives to improve cancer diagnosis, treatment and research.

i) https://onlinelibrary.wiley.com/doi/full/10.1002/bjs.10786. Accessed 25th June 2019

ii)https://www.gov.uk/government/news/matt-hancock-announces-ambition-to-map-5-million-genomes. Accessed 25th June 2019

iii) https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FCSD%2FFlyers%2Foncomine-ffpe-gene-list-flyer.pdf&title=Rmx5ZXI6IE9uY29taW5lIEFzc2F5IGdlbmUgbGlzdHMgZm9yIEZGUEUgdGlzc3VlIHByb2ZpbGluZw==. Accessed 25th June 2019

iv)https://assets.ctfassets.net/vhribv12lmne/4ZHUEfEiI8iOCk2Q6saGcU/c3361163e2c9bfeb33e934f2d00b0612/F1CDx_Tech_Specs_April_2019_1_.pdf. Accessed 25th June 2019

v)https://www.illumina.com/products/selection-tools/gene-panel-finder.html#/targeted-panels/results. Accessed 25th June 2019

Section B

B1. Priority question: The submitted version of the electronic model does not allow the probabilistic sensitivity analysis (PSA) or the one way sensitivity analysis (OWSA) to run.

The PSA does not appear to generate numeric values for the majority of the simulated comparator results and crashes after approximately 10% of the iterations. The issue may be related with the fact that age in the model is currently set probabilistically, and is sampled from normal distributions for both adults and children. This effectively allows for negative starting age values, which will then result in errors in the calculation of background mortality. Furthermore, there are errors in the Live variables sheet for a number of cells where random draws of utility estimates are calculated (e.g. cells W184:W187 and W579:W582). Please clarify why the age parameters were set stochastically in the model.

The OWSA does not generate ICER values (either lower bound, upper bound or both) for the following parameters:

i. **OS Weibull shape (p) - Larotrectinib adults**

ii. **Model adult start age (years)**

iii. **Model mixed cohort start age (years)**

Please submit a corrected version of the electronic model that is fully functional.

The submitted model with AIC and CIC markings mistakenly altered the standard upper/lower variation input from 0.2 to 2. As a result, variables for which upper and lower limits were unavailable were varied by 200%, rather than the correct 20%. This caused issues with running OWSA and calculations that fed into the parameter variation for PSA.

This has been reversed within the updated model with the OWSA (also addressing parts i, ii and iii) and PSA running correctly, replicating the results within the submitted dossier with runtimes of around 1 hour each.

Age was initially included stochastically within the model in order to investigate any uncertainty around the average age of the NTRK patient cohort.

We apologies for the error and the inconvenience caused.

B2. Priority question: The ERG was unable to replicate the results of exploratory scenarios reported in Table 57 to 61 (CS), as the submitted version of the electronic model does not appear to have the functionality to run these analyses implemented. Please submit a version of the model that runs the following scenarios probabilistically:

1. Previous line of therapy comparison (page 215 to 217, CS)

i) All patients who received a prior systemic therapy (mean GMI)

ii) All patients receiving prior systemic therapy in the metastatic disease setting (mean GMI)

2. NTRK adjustment scenarios (page 215 to 217, CS)

i) Applying the HR of [redacted] only to progression-free survival and overall survival of the colorectal tumour site engine (reflecting that the publication only considered patients with colorectal cancer)

ii) Applying the HR of [redacted] to progression-free survival and overall survival of all comparators where NTRK prevalence is >25%

Results for these alternative comparator methodology scenarios are based on simple calculations based on the results of the model. Whilst they were already within the model, they were not displayed for the user to avoid any potential confusion when adjusting results away from the base case. Bayer have now included links to the alternative scenarios calculations within the 'Results' sheet of the cost-effectiveness model, in order to allow the ERG to transparently view calculations.

B3. Priority question: The ERG was unable to replicate the full set of scenario analyses reported in table 57 (page 206-212, CS). For example, for scenario 18 ('Larotrectinib PFS – Gompertz') The CS reports an ICER of [redacted] per QALY. When the ERG attempted to replicate this result, the calculated ICER was [redacted] per QALY. Furthermore, Table 57 does not report the Utility scenario listed on Table 52, page 196 ('Replace larotrectinib utilities with alternative utility model: Revised patient pool excluding paediatrics under age 11). Please correct and update table 57 so that no discrepancies between model results and the ones reported in the CS remain, and include the utility scenario results.

We apologise for the error in presented results outlined above. It was due to a change in the order of the automation of the scenarios, leading to the 'Larotrectinib PFS – Gompertz' scenario to be incorrect. The table below reflects the updated results from the scenario analyses, in line with those produced by the cost-effectiveness model from the original submission and including the additional utility scenarios. The utility scenarios were not previously presented within the dossier due to the expectation that the base case utility values derived from the trial patient reported outcomes were the most representative for the modelled population, with scenarios based on naïve calculations and pooled literature containing larger uncertainty than the base case values.

Table 1. Scenario analyses results

Scenario number	Scenario category	Description	Larotrectinib			Pooled comparator			ICER
			Costs	QALYs	Life years	Costs	QALYs	Life years	
0		Base case results	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
1	Discount rate	Replace 3.5% discount rates for cost and outcomes with 1.5% rate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2	Utility	Replace larotrectinib utilities with weighted comparator utilities for progression-free health state	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
3		Replace larotrectinib utilities with alternative utility model: Revised patient pool excluding paediatrics under age 11	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
4		Replace larotrectinib utility for progressed disease state only with literature based relative reduction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
5	Drug costs	Full daily dose for larotrectinib adults (200mg)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
6		Larotrectinib time-to-discontinuation curve for time on treatment (Weibull)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
7		Larotrectinib time-to-discontinuation curve for time on treatment (Exponential)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
8		10 year time horizon	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario number	Scenario category	Description	Larotrectinib			Pooled comparator			ICER
			Costs	QALYs	Life years	Costs	QALYs	Life years	
9	Time horizon	20 year time horizon	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
10	Health state costs	Replace tumour location specific health state costs with consistent costs for every tumour location; weighted average of all tumour location sources	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
11		Remove health state costs if not reported in the source documents for each tumour location	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12	Adverse events	Alternative AE inclusion criteria; all AE with individual 5% rates reported in source publication	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
13	Non-GIST survival source	Use survival data from alternative source (pazopanib)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
14	Survival;	Larotrectinib OS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
15		Larotrectinib OS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
16		Larotrectinib OS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario number	Scenario category	Description	Larotrectinib			Pooled comparator			ICER
			Costs	QALYs	Life years	Costs	QALYs	Life years	
17	alternative fits	Larotrectinib OS - Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
18		Larotrectinib OS - Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
19		Larotrectinib PFS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
20		Larotrectinib PFS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
21		Larotrectinib PFS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
22		Larotrectinib PFS - Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
23		Larotrectinib PFS - Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
24		Salivary OS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
25		Salivary OS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
26		Salivary OS - Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
27		Salivary OS - Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
28		Melanoma OS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
29		Melanoma OS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
30		Melanoma OS - Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario number	Scenario category	Description	Larotrectinib			Pooled comparator			ICER
			Costs	QALYs	Life years	Costs	QALYs	Life years	
31		Melanoma OS - Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
32		Colorectal OS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
33		Colorectal OS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
34		Colorectal OS - Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
35		Colorectal OS - Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
36		Colorectal PFS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
37		Colorectal PFS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
38		Colorectal PFS - Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
39		Colorectal PFS - Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
40		STS GIST OS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
41		STS GIST OS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
42		STS GIST OS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
43		STS GIST OS - Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
44		STS GIST PFS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario number	Scenario category	Description	Larotrectinib			Pooled comparator			ICER
			Costs	QALYs	Life years	Costs	QALYs	Life years	
45		STS GIST PFS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
46		STS GIST PFS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
47		STS GIST PFS - Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
48		STS non-GIST OS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
49		STS non-GIST OS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
50		STS paediatrics OS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
51		STS paediatrics OS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
52		STS paediatrics OS - Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
53		STS paediatrics PFS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
54		STS paediatrics PFS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
55		STS paediatrics PFS - Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
56		Cholangiocarcinoma OS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
57		Cholangiocarcinoma OS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
58		Cholangiocarcinoma OS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario number	Scenario category	Description	Larotrectinib			Pooled comparator			ICER
			Costs	QALYs	Life years	Costs	QALYs	Life years	
59		Cholangiocarcinoma OS - Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
60		Cholangiocarcinoma PFS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
61		Cholangiocarcinoma PFS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
62		Cholangiocarcinoma PFS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
63		Cholangiocarcinoma PFS - Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
64		CNS/Glioma OS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
65		CNS/Glioma OS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
66		CNS/Glioma OS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
67		CNS/Glioma OS - Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
68		CNS/Glioma PFS - Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
69		CNS/Glioma PFS - Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
70		CNS/Glioma PFS - Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
71		CNS/Glioma PFS - Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

B4. Priority question: The ERG notes that the base-case analysis assumes treatment specific post-progression utilities, without presenting a clinical rationale as to why the utility weight for post-progression would differ between larotrectinib and comparator treatments. Please justify this assumption.

To best reflect the quality of life among patients receiving larotrectinib, the base case assumed the post-progression utility from the larotrectinib clinical trial programme.

The overarching aim during the development of larotrectinib was to create a precision medicine specifically targeting patients with an NTRK gene fusion while minimizing off-target toxicity. Patients who progress on larotrectinib would have higher quality of life because the treatment is very well tolerated, with extremely rare cases of grade 3 and 4 side effects. The incidence of grade 1 and 2 side effects is low as well, compared to treatments such as chemotherapy. Additionally, the type of side effects potentially occurring while taking larotrectinib – fatigue and weight gain – have a less negative/serious impact on patient's quality of life than with chemotherapy (e.g. peripheral neuropathy or cardiotoxicity). Finally, many side effects commonly related to standard chemotherapy have a potential of long term or irreversible damage (e.g. severe cardiac conduction abnormalities with paclitaxel). Clinical experts recognize the possibility of those long term side effects of chemotherapy due to organ damage/nerve damage.
Although larotrectinib's long term data is not yet mature, the extremely rare occurrence of grade 3 and 4 side effects suggest that assuming different utility weight between larotrectinib and comparator treatments for post-progression is a highly plausible approach.

Assumptions of different utilities post progression are not unusual in oncology appraisals in cases where newer highly targeted treatments with consequently less adverse events are compared against older chemotherapy regimens.

There are scenarios to explore the influence of variance in utility assumptions. One scenario tested the results when post-progression utility was derived from the progression-free utility, according to the literature-based ratio between the utilities for these two health states. Another

scenario assigned larotrectinib the weighted utilities from other comparator treatments. See responses to B3, B7 and B8 for results of utility based scenarios.

B5. Priority question: Please present a scenario analysis adjusting the treatment effect to reflect current NHS clinical practice, i.e. excluding LOXO-195 and other therapies not currently available and/or recommended in the NHS (as described in question A5).

Bayer do not believe such an analysis is appropriate. Please see below and our response to question A5.

If patients go on to receive other interventions not currently available or recommended in the NHS after larotrectinib in the trials, Bayer would not expect to adjust for this in any cost-effectiveness analysis as it could potentially compromise the validity of the ITT approach adopted in the larotrectinib trials. Such an analysis could be considered as equivalent to analysing data associated to treatment switching, which as per NICE TSD16, would not be an appropriate method when conducting an “intention to treat” analysis.

Additionally, the results of such a scenario analysis, to the best of our knowledge have not previously been requested in any NICE TA, would be out of line with clinical practice and therefore irrelevant and out of scope of this appraisal. With such rare tumours, with poor prognosis, despite treatments not being routinely available on the NHS, it is not unreasonable to assume that some patients would go on to receive further innovations as part of a clinical trial, compassionate access to medicines not yet licensed or drugs approved via a system of individual funding requests. As such, the dataset would be reflective of expected clinical practice in the UK.

While some patients who receive larotrectinib may have later received other exploratory therapies or treatments this would also be the case for patients at the same line of therapy for the comparator arm of the economic model, this is therefore expected to have a minimal impact on the results of the cost effectiveness analysis.

B6. Please present the base case cost-effectiveness results (as per table 53 and 55, page 198 and 203 in the CS respectively) and cost-effectiveness acceptability curve (as per figure 34, page 203 in the CS), using the discounted price as agreed with the PAS liaison unit.

NHS England has approved a simple discount patient access scheme for larotrectinib, however the level of discount has not yet been confirmed. At the current time it would only be possible to submit an indicative price which would not be helpful. As soon as the level of discount is confirmed we will submit the PAS template with the confirmed PAS discount.

B7. The electronic model is set up to run a scenario labelled 'Replace larotrectinib utility for progressed disease state only with literature based relative reduction', which is not mentioned in the CS. Please describe this scenario including the data sources used to inform it and present results for it.

In order to build the cost-effectiveness model for this appraisal, a vast body of evidence had to be explored and accounted for. This means the model underwent many iterations and that numerous sensitivity analysis have been conducted in the process. Unfortunately this increases the probability of a legacy code to be inadvertently left in the final version of the model. The scenario labelled 'Replace larotrectinib utility for progressed disease state only with literature based relative reduction' was considered in an early version which was not meant to be part of the final model as it wasn't considered sufficiently relevant. Bayer apologises for the confusion.

This utility scenario takes the relative risk reduction between 'progression-free' and 'progressed' health state utilities for each tumour location comparator and weights them using the clinical trial programme cohort split in order to create a weighted relative risk between the two health state utilities. This relative risk is then applied to the 'progression-free' health state utility for the larotrectinib arm to calculate a 'progressed' health state utility. The results of selecting this utility source with all other settings in line with the base case are presented in the table below.

Table 1. Additional health state utility scenario results

Scenario description	Larotrectinib			Pooled comparator			ICER
	Costs	QALYs	Life years	Costs	QALYs	Life years	
Base case results	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Replace larotrectinib utility for progressed disease state only with literature based relative reduction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

B8. The ERG notes that the electronic model allows running another utility related scenario (labelled 'Weighted average of comparators') that is also not mentioned in the CS. Please describe this scenario including the data sources used to inform it and present results for it.

In order to build the cost-effectiveness model for this appraisal, a vast body of evidence had to be explored and accounted for. This means the model underwent many iterations and that numerous sensitivity analysis have been conducted in the process. Unfortunately this increases the probability of a legacy code to be inadvertently left in the final version of the model. The scenario labelled 'Weighted average of comparators' was considered in an early version and was not meant to be part of the final model as was not considered sufficiently relevant. Bayer apologises for the confusion.

The dropdown option within the settings sheet switched the larotrectinib arm health state utility values from the base case clinical trial HRQoL outputs to using a weighted average of the comparator arm tumour location health state utility values; weighting in with the same method as the pooling of results using the clinical trial programme cohort. The results of selecting this utility source with all other settings in line with the base case are presented in the table below.

Table 1. Additional health state utility scenario results

Scenario description	Larotrectinib			Pooled comparator			ICER
	Costs	QALYs	Life years	Costs	QALYs	Life years	
Base case results	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Use of weighted comparator health state utilities for the larotrectinib arm	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

B9. The drop down button that allows selection of the parametric distribution for the larotrectinib time to treatment discontinuation curve (Settings sheet) is not working. Please correct this feature.

The treatment discontinuation curve is not used within the base case. As a result, the functionality to select the parametric distribution for larotrectinib time to discontinuation was moved from the 'Settings' sheet to the 'Tx duration' sheet. Within the updated cost-effectiveness model, the functionality has been reinstated within the 'Settings' sheet to allow the user to transparently control this within the single location.

B10. Priority question: The weighting of comparator data by tumour site is based on data reported in table 32 (page 144, CS). The CS states that due to the absence of data for certain tumour sites, tumours such as colorectal and appendix cancer were grouped together. However, it does not present the clinical rationale for each specific grouping and it is unclear how some tumour sites (e.g. congenital mesoblastic nephroma) were grouped. Please provide:

i) **A comparison between the distribution of tumour sites as applied in the model (as presented in table 32) and the distribution of NTRK fusion per tumour site in the general population**

ii) **Please describe the rationale for grouping across tumour types, where comparator data were not available for particular tumour types. Comment on the possible consequences for cost-effectiveness results**

Please discuss the generalisability of the weighting used in the model to the population of NRTK-fusion positive patients that would be seen in clinical practice and comment on the consequences to cost effectiveness results.

For comment (i) above, we have compared the distribution of tumour sites in the table below. The data come from an SLR and meta-analysis which can be found in appendix 3. NTRK prevalence is not always available for each of the included tumour locations. Using the NTRK prevalence in the overall oncology population or amongst patients with certain tumour types is likely to lead to bias because (1) NTRK fusion may not be evenly presented in the included tumour sites and (2) data from the overall oncology population may not be fully representative of patients who are treated with larotrectinib. Thus, the distribution from the larotrectinib clinical trial programme provides the best proxy to the weights among larotrectinib-treated patients in the real-world.

NTRK incidence per tumour was calculated using the tumour specific frequency in combination with tumour incidence (see table below). Distributions from the trial and that from the general population differ, with relatively bigger variation for STS paediatrics, non-GIST, colorectal cancer and breast cancer. However, the STS paediatrics data was proxied by data for infantile/congenital fibrosarcoma, which may not capture the true prevalence among the relevant population. Non-GIST data was based on a single study for Ewing sarcoma, which could also be an underestimate. Overall, the population distribution of larotrectinib clinical trials accounts for [redacted]% of the population distribution of all known NTRK patients.

Tumour locations	Tumor Incidence (per 100,000)	NTRK Incidence per Tumour (per 100,000)	Trial-based distribution	Population-based distribution (modelled tumour types)	Population-based distribution (total NTRK population)
STS paediatrics/IFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Salivary	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cholangiocarcinoma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
STS adults (GIST)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
STS adults (non-GIST)/Bone sarcoma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Thyroid	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Colorectal/Appendix	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

NSCLC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Melanoma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pancreas	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CNS/Glioma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Breast	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total			[redacted]	[redacted]	[redacted]

Source: Please see Appendix 3 for the full SLR report.

The population-based distribution reflecting the modelled population has been tested in the model and led to an improved (lower) ICER of [redacted] per QALY.

ii Please describe the rationale for grouping across tumour types, where comparator data were not available for particular tumour types. Comment on the possible consequences for cost-effectiveness results

IFS, bone sarcoma, and appendix are grouped with others based on clinical expert review given lack of data to support modelling them separately.

Infantile Febrile Sarcoma (IFS)

IFS (n=13) is grouped with STS paediatric patients given that no relevant publications for comparator data were identified in the SLR in the advanced or metastatic setting, where patients had failed previous therapies.

Publications were found in paediatric IFS patients who had early stage disease. However, these patients have a very different prognosis to those in the advanced or metastatic setting which is where larotrectinib would be used. Prior to failing therapy patient's disease can be well controlled and prognosis on available treatments is good. Once the patient relapses on available therapies such as chemotherapy, the prognosis of their disease worsens dramatically. It was determined through clinical validation that patients in the studies identified in the SLR and their outcomes were not representative of those enrolled in the clinical trial programme. Thus, following clinical advice, as a type of soft tissue sarcoma, IFS has been grouped with the paediatric STS patients to reflect a treatment relapsed population. This was accepted as a valid approach in the clinical expert validation. Without clinical data specific to IFS, data from STS paediatric patients are the best proxies. Excluding IFS patients would bring down the weight for STS paediatrics, which would lower the ICER to [redacted]/QALY.

Congenital Mesoblastic Nephroma (CMN)

For congenital mesoblastic nephroma (CMN) no studies were identified via the SLR that could inform the comparator arm of the economic model. There is only one patient with congenital mesoblastic nephroma who, following clinical advice, is grouped under STS paediatrics, given the age of 1.25 years, and are not built within the model. Removing the single patient from the STS paediatric results in an ICER of [redacted]/QALY.

Bone Sarcoma

Bone sarcoma (n=2) is grouped with STS adults non-GIST based on clinical review. No relevant NICE TA was found for this relevant population (bone sarcoma, advanced /metastatic) and there were no results available from the SLR. One of the patients enrolled in the trial received trabectedin as their prior therapy, suggesting a similar treatment pathway as patients with STS. In the clinical validation interviews, Bayer were advised that similarities in outcomes between STS and bone sarcoma patients that had failed previous therapies would be expected. Inclusion/exclusion of this tumour site has minimum impact on the overall results given that very few patients has this condition. Removing these patients from the STS non-GIST results in an ICER for [redacted]/QALY.

Appendix

In the absence of an identified clinical source for patients with appendix cancer and the low number of patients enrolled in the larotrectinib clinical trial programme with an appendix tumour site (n=1), following clinical advice the decision was made to use the colorectal cancer cohort as a proxy. No relevant NICE TA in was found for the population (appendix, advanced/metastatic), and there were no results available from the SLR for patients with appendix cancer (as this population was enrolled late in the study). No publication in the relevant population (appendix, advanced/ metastatic) was found via manual search to inform the NICE submission, however evidence was identified to suggest similarities between the cancer sites. Inclusion/exclusion of this tumour site is expected to have minimum impact on the overall results given that very few patients have this condition. An NCCN report confirmed that clinically 'appendix cancers are rare tumour that represents 1% of all diagnosed CRC cases' and that outcomes are similar 'appendix treatments and outcomes are similar to those seen in CRC' (Tejani et al 2014). Removing the single appendix patient from the colorectal weighting results in an ICER of [redacted]/QALY

Impact of grouping tumour types

Removing the proxy based weightings from the analysis results in an ICER of [redacted]/QALY based on 85 patients. However, the removal of proxy weighting impacts only the pooling calculation of the comparator arm, with the larotrectinib arm remaining pooled for all tumour

locations. We would therefore conclude that this would have no significant impact on larotrectinib's cost-effectiveness results.

It is unknown how generalisable the weighting used within the model, stemming from the trial, would be within clinical practice. Research into precision medicines and specifically epidemiology data for NTRK-fusion positive patients are still not widely reported. Whilst assumptions can be formulated around the relationship between the enrolment of patients onto clinical trials and real-world proportions, it uncertainty remains within the cost-effectiveness analysis. This uncertainty has been explored in a sensitivity analysis.

As previously indicated, effort has been made to review published studies and databases for NTRK frequencies by tumour location (see appendix 3). NTRK prevalence is not widely available for the included tumour locations. As a result, a DSA was conducted using weights from the general population (as presented in the table above). A PSA was also conducted only including the tumour weighting parameters to investigate the uncertainty and power of these inputs. The deterministic analysis resulted in an ICER of [redacted], suggesting that a difference in the distribution of NTRK fusion per tumour site in the general population versus the trial population would have no significant impact on larotrectinib's cost-effectiveness results. This conclusion seems to be further reinforced by the PSA, which resulted in an average ICER of [redacted] over the [redacted] iterations. The cost-effectiveness acceptability curve and cost-effectiveness plane show that the weighting uncertainty does not have a significant impact on the overall result.

Table 1. Tumour location weighting PSA results

Larotrectinib			**Comparators**			**Incremental**			
Life years	QALYs	Costs	Life years	QALYs	Costs	Life years	QALYs	Costs	ICER
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Figure 1. Cost-effectiveness plane for tumour location weighting PSA results

Figure 2. Cost-effectiveness acceptability curve for tumour location weighting PSA results

B11. Priority question: The pack size for the different larotrectinib presentations reported in Table 42 (page 179, CS) does not match the pack size values applied in the model (Unit cost sheet, cells G35:H38). Please clarify what is the correct pack size and total mg per pack, and correct either Table 42 or the model as appropriate, utilising the correct pack size.

Apologies for any confusion. Due to the early submission the pack sizes weren't finalised. The cost-effectiveness model uses a 'cost per 30 days of treatment' unit cost for larotrectinib. Therefore, the 'pack size' seen within the 'Unit costs' sheet reflect the 30-day treatment total dose, rather than the specified capsule pack and bottle of solution sizes.

With the updated approach of anchoring to a per-mg list price of £[redacted], we have updated the model 'Unit cost' sheet to use the per-mg price (Table 1).

Table 1. Larotrectinib presentations, pack size and list price

Formulation	Pack		Cost per mg
	Total dose	Total cost (list price)	
100mg capsules	5,600mg (56 capsules)	[redacted]	[redacted]
25mg capsules	1,400mg (56 capsules)	[redacted]	[redacted]
20mg/ml solution	2,000mg/100ml (1 bottle of solution)	[redacted]	[redacted]

B12. The company states in page 146 of the CS that treatment costs are assumed to occur at the beginning of the treatment cycle (without a half-cycle correction) to account for larotrectinib wastage due to patient discontinuation. However, the ERG notes that the cycle length is 7 days, and, therefore, the wastage assumption in the base-case considers that at most patients will be given a 7 days' supply of larotrectinib. While the CS states that this is reflective of clinical practice, the ERG considers that there may be NHS trust level variation in terms of how often patients will be supplied with medication, which is likely to impact on the estimates of cost effectiveness. Please present a scenario analysis exploring alternative wastage assumptions for larotrectinib (e.g. 2 and 4 weeks supply of larotrectinib). These scenario analyses should consider wastage in both children/adolescent and adult population.

Within the scenario analyses, Bayer presented a scenario to account for potential wastage using full treatment dosing over the observed received dosing from the clinical trial programme. In clinical practice, it has been observed that specialist oncology medicines would not be dispensed in large quantities and that unused products would be returned. However, the results of scenarios looking at prescribing two weeks and four weeks supply of larotrectinib vs the base case one-week supply has been provided below. Within these scenarios, the model accounts for a two or four week supply of larotrectinib being prescribed every second or fourth model cycle, rather than a weekly supply every cycle. This function has been added into the 'Settings' sheet of the updated cost-effectiveness model.

Table 1. Alternative prescribing patterns for larotrectinib scenarios

Scenario description	Larotrectinib			Pooled comparator			ICER
	Costs	QALYs	Life years	Costs	QALYs	Life years	
Base case results	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Two-weekly prescribing pattern	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Four-weekly prescribing pattern	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

B13. The health state costs per cycle reported in Table 48 (page 189, CS), do not match those applied in the electronic model (Health state cost sheet) for larotrectinib, and the comparators for all tumour sites with the exception of non-small cell lung cancer. Please clarify what is the correct set of values, and correct either Table 48 or the model as appropriate.

An updated version of Table 48 from the submitted dossier has been provided below with the correct calculated health state costs figures used within the cost-effectiveness analysis. Apologies for the error.

Table 1. Health state costs by tumour location

Tumour locations	Progression-free, start-up	Progression-free, per cycle	Progressed, start-up	Progressed, per cycle	Death/End-of-life
Larotrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Comparators with no active treatment					
NSCLC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Colorectal/Appendix	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GIST	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Thyroid anaplastic, follicular and papillary	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Active treatments accepted as a positioned last-line comparator					
Non-GIST/Bone sarcoma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Melanoma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Breast	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Glioma	[redacted]				[redacted]
Pancreas	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Salivary	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
STS paediatric/IFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cholangiocarcinoma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Health state costs are based on the source NICE TA or literature.

Start-up cost is the one-time cost of health resources required for assessment and/or treatment initiation when patients enter a health state. Start-up cost is assumed £0 if the source does not mention any HRU details or aggregate health state cost.

Glioma TA reported monitoring cost over the treatment period by a fixed schedule that did not fit a per-cycle calculation. Thus, the total costs were applied as a one-off cost to glioma health states.

Round 2015 was used to inform end-of-life cost for tumour locations that did not have this data in the TA or literature sources.

NSCLC, non-small-cell lung cancer; GIST, gastrointestinal stromal tumours; STS, soft tissue sarcoma; IFS, infantile fibrosarcoma.

B14. The search strategies for 3 of the 4 economic systematic reviews have not been included in the submission. The reviews for which the searches strategies are missing are:

i -Cost-effectiveness by tumour site (Appendix G, Section G.2, page 324)

ii- Cost and healthcare resource use (Appendix H, page 338)

iii - Health-related quality of life (Appendix I, page 350)

Please report the search strategies for each of these reviews.

The search strategies have been provided in response to question A14.

B15. Please present a scenario analysis that models the cost of testing for NTRK gene fusions consistent with the description of the decision problem (Table 1, CS).

Bayer feels that including the cost of testing for NTRK gene fusions in the CE analysis is not in line with the NICE process guide. The NICE processes guide stipulates that the costs associated with the companion diagnostic test should be incorporated into the assessments of clinical and cost effectiveness in instances where a diagnostic test for of a biomarker is carried out solely to support the treatment decision for the specific technology. This is clearly not the case for the testing for NTRK fusions.

The NHS Long Term Plan aims to offer whole genome sequencing (WGS) as part of routine cancer care. Given the comprehensive nature of WGS, one test can provide information on multiple targets. It is clear that whether panels covering a wide range of genes or WGS testing is implemented, NHSE is not developing a national service solely for treatment with larotrectinib or any other precision medicine in development. (For further details see question A21)

Furthermore there are at least five compounds in development for the treatment of NTRK fusion positive cancers so NTRK gene fusion testing would not be unique to larotrectinib. To apply costs to larotrectinib would be inequitable.

Hence, conducting a scenario analysis that models the cost of testing for NTRK gene fusions is not in line with the NICE process guide.

Section C

C1. Please confirm Figure 10 is complete and correct. We would expect one bar for each person in the SAS3 set, but only [redacted] bars are visible; Table 12 indicates that [redacted] patients had evaluable data available.

Bayer apologises for not explaining this inconsistency. Figure 10 only included patients with RECIST measurements, shown below is the updated figure that uses RANO measurements when available in addition to RECIST measurements. As shown in the figure, [redacted] of the [redacted] total SAS3 patients had post-baseline measurements recorded.

One patient out of [redacted] was non-evaluable and one had non-target lesions. Only measurable target lesions are presented in the waterfall plot.

NICE National Institute for Health and Care Excellence

Patient organisation submission

Larotrectinib for treating NTRK fusion-positive advanced solid tumours [ID1299]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	**Jayne Bressington**

2. Name of organisation	**GIST Support UK**
3. Job title or position	**Trustee & Vice Chair GIST Support UK** **Patient Director PAWS-GIST**
4a. Brief description of the organisation (including who funds it). How many members does it have?	**GIST Support UK is a registered charity (No. 1129219) formed in April 2009.** **We are a network of GIST cancer patients & carers working with top GIST specialists & National/International groups, to promote best practice. We exist to help GIST patients and their families come to terms with living with GIST cancer and we raise funds to:** • **Stimulate and fund GIST research.** • **Support Patients living with GIST cancer** • **Provide Information for GIST patients and their clinicians** • **Raise awareness of GIST cancer** **We receive no government funding and are run by a board of, currently ten, volunteer trustees who have a close association and experience of GIST cancer.** **GIST Support UK is not a membership organisation. Each year we engage with over a thousand GIST patients and carers, both newly diagnosed and longer-term survivors, via:** • **our telephone helpline,** • **regional patient carer meetings,** • **PAWS-GIST clinics,** • **our private online patient forum** • **social media Facebook & twitter platforms** **This amounts to many thousands of patient and carer experiences since we started in 2009.**

	We have been working hard to establish an infrastructure here in the UK that will help to stimulate research into GIST. In recent years we have helped to re-write and update the National GIST Guidelines, establish the National GIST Tissue Bank and the PAWS-GIST clinic at Addenbrookes hospital in Cambridge. All of these things help to stimulate research and our work continues... **We are the only UK based charity solely devoted to GIST cancer.**
4b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	**No**
5. How did you gather information about the experiences of patients and carers to include in your submission?	**GIST Support UK has gathered information about the experiences of patients since it became a charity in 2009.** **GIST Support UK engages with GIST patients, clinicians and researchers both in the UK and Internationally. Our work to find answers and treatments is extensive and has resulted in the implementation of infrastructure in the UK to support and stimulate GIST research e.g. the National GIST Guidelines, establishing the National GIST Tissue Bank and also the PAWS-GIST clinic at Addenbrookes hospital in Cambridge, for the rarer subsets of GIST such as those with NTRK fusions who currently do not have effective treatment options.** **Through our work to support GIST patients we gain valuable information about patient experiences. GIST Support UK also engages directly with patients in a variety of ways; our private listserve (email forum community) for patients and carers, patient and carer meetings (held 3 times per year), and via our telephone helpline.** **There are some pretty horrific things that can happen in a person's life. A cancer diagnosis is up there as one of the worst but when along with that diagnosis you are advised that there is no treatment and also no research, it is totally devastating.**

	This is what patients and their families experience. GIST is a heterogeneous cancer and there are many subtypes. **As a result of research that happened decades ago there are some types of GIST that respond well to treatments such as Imatinib, enabling patients to live long and productive lives. If the GIST that you are diagnosed with does not have mutations that can be targeted by drugs such as Imatinib the outlook is very frightening.** **And then there is hope.** **Hope inspired by the discovery that a new drug has been found and if you happen to be the subtype where the drug is effective then this is fantastic. Then there is the stress related to how you gain access to this new drug…** **In recent years we have been hearing that some GIST patients have a mutation called an *NTRK* fusion and that they should be given a drug called Larotrectinib, as the results for patients is quite remarkable. e.g. a patient with extensive metastatic GIST who had failed five prior therapies showing a massive tumour response when treated with Larotrectinib. We have seen the evidence presented at GIST conferences in mainland Europe, USA & UK recommending that Quadruple negative GIST patients should all be tested to see if they have an *NTRK* fusion.** **Quadruple negative GIST patients represent the second largest group of patients who have attended the PAWS-GIST clinic to date. Tests have commenced to see which ones carry an NTRK fusion.** ***NTRK* gene fusions can drive unregulated cell growth and proliferation in a range of cancer types. It is present in a range of cancers and the level of excitement surrounding the results it has shown in trials is re-enforced by things such as the ESMO Advanced Course on "NTRK Gene Fusion: A New Target in Treatment of Precision Cancer", which is being launched this year to train clinicians in this field. The discovery of this fusion and drugs such as Larotrectinib to treat it is state of the art ground breaking medicine.** **The advent of next-Generation Sequencing provides the most comprehensive view across a large number of genes and can identify *NTRK* gene fusions as well as other relevant alterations, with minimal sample tissue needed.**

	We understand that it will become standard practice in NHS England this year, for all cancer patients undergoing surgery to have their tumours sequenced. So, the incidence of finding patients for whom and *NTRK* fusion inhibitor Larotrectinib will be a suitable treatment will increase and be easier to find. We are very excited that this technology is becoming standard practice within the NHS.
Living with the condition	
6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	**Many GIST cancer patients manage, with effective treatment, to live relatively normal lives, continuing to work and play as best they can while manging the side effects of treatment. There are some who are fortunate that their GIST cancer is found early and before it has spread, they have it removed while still small and it does not return. This is as close to a cure as currently exists.** **Depending on the extent of disease, surgery can involve quite drastic interventions such as removal of the stomach. Often the disease has reached an advanced stage prior to diagnosis, limiting the potential for surgery to totally remove the cancer. Toxic side effects are also encountered from anticancer therapies, and tolerance of these side effects varies significantly. Side effects to the drug therapies currently available via NHS include hypertension, hypothyroidism, debilitating hand foot syndrome, diarrhoea, fatigue, nausea, skin rashes and so on. The list of side effects is quite extensive but with advice from oncologists, cancer nurse specialists and fellow patients we observe that these can be managed and tolerated to some degree by many patients, providing the chance to live longer and live a normal life. However, some patients are often forced to defer and put their lives on hold due to GIST cancer.**

	Living with GIST cancer as a patient and a carer is possible but every day that you wake up you hope that it was a bad dream and that it isn't real. This is a standard defence mechanism for cancer patients and their families. Learning to cope is something that you have to do and the last thing that you want to do as a carer is to give the impression that things will not be OK. You have to give your loved one hope. **The traumas and horrors of living with a type of GIST cancer that does not have a treatment that works can shatter family's lives. Carers take many forms, parents, partners, siblings, children and friends, all desperate to help and save the person that they love. A cancer diagnosis is the last thing that you think will happen to you or someone you love. It always happens to someone else, doesn't it?**
Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	**Patients and carers are very grateful for the treatments that are available via the NHS.** **Currently for GIST patients this consists of:** • **Surgery** • **Imatinib** • **Sutent** • **Regorafenib** **Unfortunately, not all GIST cancers are the same and there are many for whom the above treatments are not effective because either their primary mutation is not targeted by the above treatments or their disease metastasizes beyond the control of the above treatments.** **All GIST patients are currently given the above options. With the advent of knowledge such as the existence of NTRK fusion driven GIST and a specific targeted treatment such as Larotrectinib this may in future change the standard treatment pathway.**

	We are very grateful that there is some research happening in the world and that a treatment has been discovered for those GIST patients who have an *NTRK* fusion mutation. **Currently such a treatment is not available via the NHS but we hope that further to this appraisal that it will be available for patients with GIST caused by *NTRK* fusion mutations.**
8. Is there an unmet need for patients with this condition?	**Yes.** ***NTRK* mutated GIST patients do not currently have access to a targeted treatment for their type of GIST mutation via the NHS.** **For some patients with particular types of GIST, the anticancer drugs that are currently available are less effective. This includes PAWS-GIST patients (which includes those with NTRK fusion). A key reason for this is due to the lack of existing available therapies targeting specific mutations that drive these cancers, demonstrating a significant un-met need for targeted therapies such as Larotrectinib.**
Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	• **The advantages of this technology are that it is a precision medicine designed to target the *NTRK* fusion mutation which is the cause of the cancer. It blocks the signalling pathway so that the tumours cannot grow and is effective even when patients have been heavily pre-treated with other therapies.** • **Larotrectinib is administered orally in capsule or liquid form and studies to date show that it is well tolerated and has shown encouraging anti-tumour activity in all patients with TRK fusion positive tumours. In one trial 75% of patients remained on Larotrectinib and some were able to have surgery with curative intent as their tumours had shrunk sufficiently to make them operable.** • **Larotrectinib is suitable for both adults and children.** • **Drugs of this type are exactly what rare cancer patients are desperate to find and use to shrink**

	and stop their tumours and get their life back on track.
Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	**The only disadvantage that we can see is being an NTRK fusion cancer patient and not being able to access Larotrectinib.** **As with any drug there are side effects but those listed are tolerable and can be managed.**
Patient population	
11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	**Our understanding is that the patients currently classified as Quadruple negative GIST patients are the target group where NTRK fusion GIST's can be found.** **They have been named quadruple Wildtype GIST as they lack abnormalities in the four signalling pathways KIT, PDGFRA, SDH or RAS.** **When the whole genome sequencing starts as standard within the NHS this year it will speed up the discovery of patients whose cancer is caused by NTRK fusions for whom Larotrectinib is effective.**

	We are already screening all quadruple negative GIST patients who attend the PAWS-GIST clinic at Addenbrookes Hospital in Cambridge to find the ones that have NTRK fusions.
Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	**The only inequality we can see currently, is that the world is fast tracking NTRK fusion inhibitors to be available to patients with NTRK fusions and until this appraisal has concluded Larotrctinib is not available to the patients who will benefit in the UK.**
Other issues	
13. Are there any other issues that you would like the committee to consider?	**No**

NICE National Institute for Health and Care Excellence

Key messages
14. In up to 5 bullet points, please summarise the key messages of your submission: • **NTRK fusions are the root cause of some GIST cancers.** • **Larotrectinib is a precision medicine that targets NTRK fusion mutations.** • **Trials have resulted in dramatic results for GIST patients with NTRK fusions.** • **The NHS whole genome sequencing being launched this year will identify the patients with NTRK fusions.** • **Using Larotrectinib in GIST patients with an NTRK fusion will reduce unnecessary expenditure on other ineffective therapies that are very expensive for the NHS.**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

……

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

NICE National Institute for Health and Care Excellence

Patient organisation submission

Larotrectinib for treating NTRK fusion-positive advanced solid tumours [ID1299]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	[redacted]

2. Name of organisation	Sarcoma UK
3. Job title or position	Director of Research and Policy
4a. Brief description of the organisation (including who funds it). How many members does it have?	Sarcoma UK is the only cancer charity in the UK focusing on all types of sarcoma. Sarcoma UK raises awareness of sarcoma, funds peer reviewed Research and provides information and support to everyone affected by sarcoma. We are a registered as a charity in England and Wales (1139869) and in Scotland (SC044260) and a company limited by guarantee in England and Wales (7487432). The charity is funded by donations from supporters who predominantly have a personal connection with the cause. Sarcoma UK is not a membership organisation, we have a database of over 8000 active and engaged supporters. We receive no funding from government or other statutory sources.
4b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
5. How did you gather information about the experiences of patients and carers to include in your submission?	The Sarcoma UK Support line has answered over 4500 questions from individuals with, or carers of sarcoma, it gives us a unique understanding of living with the condition. Users of our Support Line are 50% patients and 50% carers; this gives us a balanced view of sarcoma which affects all ages and demographics. We also speak directly to patients at our support groups to gather their views about lack of treatment options when surgical resection is not possible. We were contacted by an Irish family whose child had been involved in the Larotrectinib trial and following treatment, the tumour had shrunk so that surgery to remove the tumour was possible.

Living with the condition	
6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	Sarcoma is a rare disease with approximately 3,800 people diagnosed with a soft tissue sarcoma a year, there are around 100 different sub-types. Sarcoma is one of the hardest cancers to diagnose, with patients visiting their GP more times than those with any other form of cancer before being diagnosed with sarcoma. Since setting up the Sarcoma UK Support line in February 2016, we have heard this confirmed from both patients and carers. This is also backed up by respondents to our National Sarcoma Survey on patient experience, published in 2016. https://sarcoma.org.uk/sites/default/files/resources/the_national_sarcoma_survey_feb_2017.pdf The uncertainty of sarcoma is described by our callers. We have patients who call in a cycle aligned to their follow up appointment. Recurrence of local disease is common and not unheard of after 5 or even 10 years. Commonly, it is the patient who picks up a local recurrence, whilst metastatic disease is usually picked up routinely on chest X-ray without any symptomatic suggestion to the patient that something is wrong. The constant fear of recurrence combined with the fear of the unknown is often described by callers, alongside their fears around prognosis and the limited treatment options available to sarcoma patients. They tell us that the rare nature of sarcoma means that they have to become experts and the source of further information around their disease We hear a lot from carers who reflect that lack of public awareness about sarcoma. They don't know anything about the condition and fail to understand what and why this happening to their loved one. Sarcoma affects all ages, from paediatric patients to the elderly and this is hard on family life, especially for carers who may not be involved in the early stages of diagnosis and treatment. Gough (2011) - https://www.ncbi.nlm.nih.gov/pubmed/22190862 reports that soft tissue sarcoma patients maintain a good quality of life with moderate symptoms until a rapid decline in the final weeks of their life. We believe this is unique to sarcoma, and is a contrast to other cancers like non-small cell lung cancer where there is a slow deterioration. This has implications for the patient and their families as home life and financial situations can change suddenly The end of treatment and the introduction of best supportive care

	is made on average only 3.4 weeks before the end of life, perhaps because of the good quality of life maintained until the end of life. Callers to our support line often report fatigue, pain, limitation to their mobility, impact of treatment to their quality of life and anxiety. The heterogeneity of the disease means that sarcoma patients can have a wide spread of symptoms dependent on the location of the primary tumour and or the metastatic disease.
Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	• Limited treatment options and all options are dated, with Doxorubicin having been prescribed for ~40 years. • New treatments for sarcoma are not emerging as fast as for other cancer groups. • Lack of clinical trials- National Sarcoma UK Survey 2015 (of 650 sarcoma patients in England and Scotland) found only a third of patients were offered a clinical trial and of these, only 20% took part. This clearly indicates that options are limited for access to new treatments and technologies once the small number of standard treatments have been exhausted. • No personalisation of treatments, no knowledge of whether a treatment will work for them.
8. Is there an unmet need for patients with this condition?	There is a high unmet need for sarcoma patients. Many patients are diagnosed late stage and curative surgery is not an option. Patients have: • Very limited treatment options. • No adjuvant treatment for most sub types, so predominately patient receive surgery with radiotherapy. Local recurrence is common but little option except further surgery. • Very few sub-type specific treatments • No curative treatment for metastatic disease

Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	• Larotrectinib is a drug which affects solid tumours which are confirmed NTRK-fusion positive and is a step towards personalised medicine for sarcoma (and other) patients. • It will only be given to patients who have confirmed NTRK-fusion positive tumours, patients who receive treatment will know their tumour will respond and can give informed consent. Uptake is likely to be very high in the eligible population. • Trial and anecdotal evidence indicates it may reduce soft tissue sarcoma to either give greater local control or allow for surgical removal or resection of the tumour. These tumours types have previously been untreatable and Larotrectinib has potential to give both increased survival and quality of life to patients. • Quality of life is essential for sarcoma patients. Late stage diagnosis means that many patients experience sarcoma in a palliative phase. Current treatment options to control metastatic disease are chemotherapy that requires intravenous or central access. Larotrectinib may allow for better quality of life within the palliative setting. • Oral delivery of treatment will have a huge benefit to patients, will mean less time in hospital and more “daily living”. Sarcoma affects all ages, for Paediatric and TYA patients this means less time away from school and social activities beneficial to life and well-being. Many patients have dependent children and good quality of life is essential to family life. • Oral delivery will have less economic impact on both the patients and NHS, requiring fewer visits, with less time away from work, travel to treatment centres, less planning life around appointments. Oral treatment will require less nursing and medical staff time, fewer clinic spaces and have less economic burden on Clinical services. • Sarcoma patients are listed on the NHS England directory to have Whole Genomic Sequencing as standard when the service is rolled out in ~July 2019 (current planned date given by Mark Caulfield). They will already have the confirmatory test as routine standard of care.

Disadvantages of the technology	
• 10. What do patients or carers think are the disadvantages of the technology?	• NONE
Patient population	
11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	• Patients who are eligible should be given this treatment as a first line as it may downstage tumours. Larotrectinib should be considered for patients with tumours which are either locally advanced, or technically difficult to operate and may contribute to curative treatment. • Sarcoma patients who have either locally advanced or metastatic solid tumours who have already used other therapies and have no other treatment options available except palliative care. • Reading the peer reviewed and published literature suggests that eligible Paediatric patients would especially benefit from treatment. Sarcoma is a less common cancer, but makes up ~15% of childhood cancers showing urgent unmet need for this group. Research has shown increased benefit in patients with Paediatric, Adolescent Wild Type GIST.
Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	Sarcoma is a rare cancer and unique to its make up is the heterogeneity. We know that in principle, sarcoma patients are younger and able to remain actively engaged in work and family life until very close to the end of their life. For many the time from primary diagnosis to local recurrence or metastatic disease can be years of productive life. It is important that these small numbers of people are not discriminated against because they are unfortunate enough to be diagnosed with a rare cancer. They should have equal access to treatments

Other issues	
13. Are there any other issues that you would like the committee to consider?	Whole Genomic Sequencing and reporting is part of the planned standard of care for all sarcoma patients. Testing for the NTRK-fusion positive gene must not be included in the economic assessment of Larotrectinib.

Key messages

14. In up to 5 bullet points, please summarise the key messages of your submission:

- Sarcoma is a less common cancer which has low public awareness
- Patients frequently experience difficult and late diagnosis leading to limited treatment options
- Larotrectinib may reduce tumour size to enable effective surgery
- The oral medication regimen is low burden on patients and NHS service
- We fully support the approval of Larotrectinib for NTRK-fusion positive patients

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

……

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☒ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

NICE National Institute for Health and Care Excellence

Professional organisation submission

Larotrectinib for treating NTRK fusion-positive advanced solid tumours [ID1299]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	[redacted]
2. Name of organisation	**Royal College of Pathologists**

3. Job title or position	[redacted]
4. Are you (please tick all that apply):	☒ an employee or representative of a healthcare professional organisation that represents clinicians? ☐ a specialist in the treatment of people with this condition? ☒ a specialist in the clinical evidence base for this condition or technology? ☒ other (please specify): A specialist in the companion diagnostics required to identify patients who may benefit from this therapy
5a. Brief description of the organisation (including who funds it).	**Professional body representing clinicians and scientists involved in diagnosis of disease.** Funded largely from registration fees.
5b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	**no**
The aim of treatment for this condition	
6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition,	To halt progression and reduce disease burden.

or prevent progression or disability.)	
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Availability of therapies targeted to particular genomic drivers is limited.
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	Several tumour types included so no single best practice. Given that eligible patients would lack other genomic drivers with approved therapy (eg EGFR, ALK etc) treatment would usually be with chemotherapy or, increasingly in some tumour types, immunotherapy if eligible.
• Are any clinical guidelines used in the treatment of the	Several tumour types included so no single best practice, but all tumour types have agreed standard-of-care therapy for patients with advanced disease. In some tumour types treatment is already based on the

condition, and if so, which?	results of companion diagnostic tests (IHC and/or genomics), whereas in others companion diagnostic testing may currently be less embedded.
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	Would expect broad agreement in most tumour types, notwithstanding iterative adjustment when a novel therapy such as this emerges.
• What impact would the technology have on the current pathway of care?	In tumour types in which companion diagnostic testing is already reflexed following a diagnosis of advanced or unresectable disease it would have little impact, particularly if this testing includes fusion detection as is the case in NSCLC. In tumour types for which this is not already the case, it would necessitate an additional testing stage and consequent small delay to therapeutic decision making.
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Yes
• How does healthcare resource use differ between the technology and current care?	By April 2020 it is likely that gene fusion screening in tumour types for which this is already necessary (eg NSCLC) would be performed using “panel-testing”, involving screening of a broad panel of genes known to be subject to fusion events in cancer, including NTRK genes, in a single test. In these tumour types, therefore, there would be little change in the diagnostic pathway. Tumour types for which fusion screening is not currently applied may require additional resource to fund this.

• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Clinical oncology
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	See above regarding healthcare resource. As mentioned, the test itself should already be available in most genomics laboratory hubs, but may need to be validated in appropriate additional tumour types.
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Trial data available to date suggests that this would be the case.
• Do you expect the technology to increase length of life more than current care?	As above.
• Do you expect the technology to increase health-related quality of life more than current care?	As above.

12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Effectiveness in NTRK-positive patients seems similar but frequency of NTRK rearrangements is greater in (albeit rare) tumour types seen in childhood.
The use of the technology	
13. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	See above for considerations regarding companion diagnostic testing.

14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Yes – additional testing requirements as mentioned above.
15. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	No
16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it	Yes. Enhances stratified approach to treatment.

improve the way that current need is met?	
• Is the technology a 'step-change' in the management of the condition?	For tumour types in which stratified approach to therapy is not currently standard of care, yes. Less so for tumour types for which treatment (for some patients) is already based on CDx testing, although even in these conditions, NTRK-positive patients are currently treated with standard chemotherapy, so this is a step change in that sense.
• Does the use of the technology address any particular unmet need of the patient population?	The need for more effective therapies.
17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	
Sources of evidence	

18. Do the clinical trials on the technology reflect current UK clinical practice?	Yes, broadly.
• If not, how could the results be extrapolated to the UK setting?	
• What, in your view, are the most important outcomes, and were they measured in the trials?	ORR, CRR, response duration/PFS Yes, these were measured.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	Unclear but sufficient to compare against current therapy.
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	
19. Are you aware of any relevant evidence that might	no

not be found by a systematic review of the trial evidence?	
20. How do data on real-world experience compare with the trial data?	comparable
Equality	
21a. Are there any potential equality issues that should be taken into account when considering this treatment?	
21b. Consider whether these issues are different from issues with current care and why.	
Key messages	

22. In up to 5 bullet points, please summarise the key messages of your submission.

- This technology would appear to be of significant benefit to patients with NTRK fusion-positive cancers
- Equitable access to screening would be required across tumour types to ensure that a test capable of detecting NTRK fusions is funded where appropriate (in liason with NHSE's genomics implementation unit)
- The above screening (and associated reflexing from pathology) is likely to already be embedded and funded in some tumour types but may need to be established in others in which fusion detection is not currently standard of care
-
-

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

.....

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

.....

NICE National Institute for Health and Care Excellence

Clinical expert statement

Larotrectinib for treating NTRK fusion-positive advanced solid tumours [ID1299]

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	**Alistair Reid**
2. Name of organisation	**Liverpool Clinical Laboratories/Liverpool Women's Hospital**

3. Job title or position	**Consultant Clinical Scientist**
4. Are you (please tick all that apply):	☐ an employee or representative of a healthcare professional organisation that represents clinicians? ☐ a specialist in the treatment of people with this condition? ☐ a specialist in the clinical evidence base for this condition or technology? ☒ other (please specify): A specialist in use of diagnostic genetic technologies that may be used to identify patients who may respond to targeted therapy in these conditions
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☒ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.)
6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the	☐ yes

rest of this form will be deleted after submission.)	
The aim of treatment for this condition	
7. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	To reduce tumour burden and delay progression
8. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	Outside my expertise.
9. In your view, is there an unmet need for patients and	Yes, therapeutic options are limited.

healthcare professionals in this condition?	
What is the expected place of the technology in current practice?	
10. How is the condition currently treated in the NHS?	The application covers a broad range of tumour types which generally would be treated with chemotherapy. In a minority of these conditions a proportion of patients may be eligible for immunotherapy. While a proportion of some of these cancers may be administered existing genomically targeted treatments, NTRK fusion-positive cancers in particular would not be expected to harbour other genomic biomarkers conferring eligibility for alternative targeted therapies.
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	Yes, tumour type-specific. Precise details outside my area of expertise.
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	The existing pathways are well defined.

• What impact would the technology have on the current pathway of care?	Following appropriate companion diagnostic test result, this drug would be administered instead of chemotherapy.
11. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Yes
• How does healthcare resource use differ between the technology and current care?	I would imagine it to be the same, although I am not involved in patient-facing oncology.
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Oncology clinics.
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	I would imagine that little additional investment is required, if any.
12. Do you expect the technology to provide clinically	Yes

meaningful benefits compared with current care?	
• Do you expect the technology to increase length of life more than current care?	Yes
• Do you expect the technology to increase health-related quality of life more than current care?	Yes
13. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Only patients whose tumours bear NTRK gene fusions. In certain rare tumour types these aberrations are relatively common, including but not restricted to certain paediatric tumours. However, these changes are also seen in common adult tumours, albeit relatively rarely.
The use of the technology	
14. Will the technology be easier or more difficult to use for patients or healthcare	Tumour types for which whole genome sequencing is not the current (or imminent) standard of care will require a test for NTRK fusions that is not currently standard of care in these diseases. Essentially

professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	therefore this additional test would be required in all relevant adult tumour types with the exception of sarcoma and acute leukaemia. In some diseases for which other gene fusions are already sought (eg lung cancer) an NTRK test could be combined with existing fusion tests on a "next generation sequencing" platform, for negligible additional expense. However, in diseases for which gene fusions are not currently sought, it seems likely that a "new" test would be required specifically for the NTRK fusions. Issues around additional clinical requirements are outside of my area of expertise.
15. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	See above (14) In addition it will be important to determine whether one particular type of test (IHC or genomics-based) is of greater clinical utility or preferable from an economic standpoint. As indicated above cost may vary according to tumour type depending on pre-existing need for screening of other gene fusion events.
16. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the	No

quality-adjusted life year (QALY) calculation?	
17. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	Yes
• Is the technology a 'step-change' in the management of the condition?	There is a precedent for the use of genomically-targeted therapies in some rare and common cancers, however due to mutual exclusivity of "driver" mutations these therapies would not be available to patients with NTRK-positive tumours, so in that sense the technology would be a step change for this genomic subtype of cancer.
• Does the use of the technology address any particular unmet need of the patient population?	Limited therapeutic options
18. How do any side effects or adverse effects of the	Outside my area of expertise

technology affect the management of the condition and the patient's quality of life?	
Sources of evidence	
19. Do the clinical trials on the technology reflect current UK clinical practice?	Yes, broadly
• If not, how could the results be extrapolated to the UK setting?	
• What, in your view, are the most important outcomes, and were they measured in the trials?	OS, PFS, tumour burden/response. Yes.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	
• Are there any adverse effects that were not	Not that I am aware of, although I am not an oncologist.

apparent in clinical trials but have come to light subsequently?	
20. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No
21. How do data on real-world experience compare with the trial data?	Anecdotally they seem comparable.
Equality	
22a. Are there any potential equality issues that should be taken into account when considering this treatment?	no
22b. Consider whether these issues are different from issues with current care and why.	

Topic-specific questions	
23. Is it appropriate to consider larotrectinib as a treatment option only after all other chemotherapy options have been exhausted for the following tumour types: sarcoma, thyroid, non-small cell lung, colorectal and gastrointestinal stromal tumours?	This would seem to be at odds with the current use of approved targeted therapies in metastatic cancer.
24. What other locally advanced or metastatic solid tumour types have NTRK fusions? At what point in the treatment pathway would it be appropriate to consider entrectinib as a treatment	secretory breast carcinoma Glioblastoma intrahepatic cholangiocarcinoma* congenital fibrosarcoma*

option for each of these solid tumour types?	*rare cases

Key messages

25. In up to 5 bullet points, please summarise the key messages of your statement.

- An effective targeted therapy would be of significant clinical benefit in these cancers.
- There is already a precedent for companion diagnostic testing of entire populations of advanced cancers to identify rare genetic subtypes with high likelihood of response to approved drugs
- Short-term cost implications of genomic screening vary according to tumour type from cost-neutral if full genomic screen already funded in that tumour type (eg paediatric) to significant cost implications if no fusion-based screening currently performed for a common tumour type (eg colorectal cancer)
- In medium to long term it seems likely that additional genomic biomarkers will emerge in most tumour types such that broad genomic screening would become standard of care to leverage access to a range of effective targeted therapies
-

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

……

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

.....

NICE National Institute for Health and Care Excellence

Clinical expert statement

Larotrectinib for treating NTRK fusion-positive advanced solid tumours [ID1299]

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	**Dr Charlotte Benson**
2. Name of organisation	**Sarcoma Unit, Royal Marsden Hospital**

3. Job title or position	**Consultant Medical Oncologist, Sarcoma Unit**
4. Are you (please tick all that apply):	☐ an employee or representative of a healthcare professional organisation that represents clinicians? X☐ a specialist in the treatment of people with this condition? ☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☐ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it X☐ other (they didn't submit one, I don't know if they submitted one etc.) I haven't yet seen it
6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☐ yes

The aim of treatment for this condition	
7. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	The main aim is of controlling and palliating advanced tumours that harbour NTRK mutations. Those of which I have expertise in are a proportion of soft tissue sarcomas and gastrointestinal tumours (GIST). I am an adult oncologist but there are a group of paediatric sarcomas (infantile fibrosarcomas) which harbour this mutation for which larotrectinib has been succesfully used in the pre- operative setting. (The use of neoadjuvant larotrectinib in the management of children with locally advanced TRK fusion sarcomas. DuBois SG et al Cancer. 2018 Nov 1;124(21):4241-4247)
8. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	Control of rate of growth of the tumour, reduction in size, improvement in disease related symptoms
9. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes- the soft tissue sarcomas which harbour this mutation have few if any effective medical therapies Eg inflammatory myofibroblastic tumour- current treatments include steroids, ALK inhibitors if ALK positive (not available on NHS) Malignant peripheral nerve sheath tumour- largely chemo resistant Spindle cell sarcoma NOS- uncertain response to chemo 'Wild type' GIST- no definite response to standard GIST treatments eg imatinib Infantile fibrosarcoma- few effective treatments

What is the expected place of the technology in current practice?	
10. How is the condition currently treated in the NHS?	Please see list above
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	Soft tissue sarcoma guidelines- Local and National British Sarcoma Group guidelines apply but these are general guidelines for a group of >70 tumour subtypes and don't take specific account of these rare subtypes GIST guidelines- as above
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	These are rare subtypes of a rare tumour group(sarcoma) that are treated in expert tertiary referral centres. Currently there is no standard of care for these patients if a NTRK fusion is identified
• What impact would the technology have on the current pathway of care?	Would offer a new and effective treatment for patients where none such option currently is available
11. Will the technology be used (or is it already used) in	yes

the same way as current care in NHS clinical practice?	
• How does healthcare resource use differ between the technology and current care?	Testing for NTRK fusion is by FISH, immunohistochemistry or next generation sequencing (NGS) . NGS is in the process of being rolled out for solid tumours and Sarcoma is at the forefront of this.
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Specialist sarcoma centres
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	Please see above
12. Do you expect the technology to provide clinically meaningful benefits compared with current care?	
• Do you expect the technology to increase	We don't have overall survival data yet, but the NEJM paper showed marked and durable anti tumour activity (NEngl J Med 2018; 378:731-9)

length of life more than current care?	
• Do you expect the technology to increase health-related quality of life more than current care?	Yes would expect symptomatic improvement
13. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Only those with identifiable NTRK fusions
The use of the technology	
14. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant	No- considerable experience already of using similar tyrosine kinase inhibitors, side effect profile shows manageable toxicity

treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	
15. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Standard oncologic principles
16. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	
17. Do you consider the technology to be innovative in	Yes- significant unmet need in current population of patients

its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	
• Is the technology a 'step-change' in the management of the condition?	yes
• Does the use of the technology address any particular unmet need of the patient population?	As outlined above
18. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	Majority of adverse events in NEJM paper were grade 1 or 2, any adverse events could be managed by dose reduction or interruption
Sources of evidence	

19. Do the clinical trials on the technology reflect current UK clinical practice?	yes
• If not, how could the results be extrapolated to the UK setting?	
• What, in your view, are the most important outcomes, and were they measured in the trials?	Phase 1-2 studies published so far. Due to rarity formal randomised Phase 3 in soft tissue sarcoma/GIST highly unlikely
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	Not that I am aware of, I don't have routine access to this drug currently
20. Are you aware of any relevant evidence that might	

not be found by a systematic review of the trial evidence?	
21. How do data on real-world experience compare with the trial data?	None yet published
Equality	
22a. Are there any potential equality issues that should be taken into account when considering this treatment?	Not that I am aware of
22b. Consider whether these issues are different from issues with current care and why.	
Topic-specific questions	
23. Is it appropriate to consider larotrectinib as a treatment option only after all other	No-would consider first line in those that harbour the mutation with soft tissue sarcoma due to poor efficacy of standard chemotherapy (single agent doxorubicin)

chemotherapy options have been exhausted for the following tumour types: sarcoma, thyroid, non-small cell lung, colorectal and gastrointestinal stromal tumours?	In small group of wild type GIST patients would consider in all lines of therapy- likely second line after imatinb
24. What other locally advanced or metastatic solid tumour types have NTRK fusions? At what point in the treatment pathway would it be appropriate to consider entrectinib as a treatment option for each of these solid tumour types?	
Key messages	

25. In up to 5 bullet points, please summarise the key messages of your statement.

- significant advance in treatment for those with soft tissue sarcoma and GIST that harbour NTRK mutation
- pre operative use in paediatric population with infantile fibrosarcoma
- good safety profile
-
-

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

.....

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

.....

NICE National Institute for Health and Care Excellence

Patient expert statement

Larotrectinib for treating NTRK fusion-positive advanced solid tumours [ID1299]

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	**Jayne Bressington**
2. Are you (please tick all that apply):	☐ a patient with the condition? ✔ a carer of a patient with the condition?

	✔ a patient organisation employee or volunteer? ☐ other (please specify):
3. Name of your nominating organisation	GIST SUPPORT UK & Sarcoma UK
4. Did your nominating organisation submit a submission?	✔ yes, they did ☐ no, they didn't ☐ I don't know
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	✔ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.)

6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☐ yes
7. How did you gather the information included in your statement? (please tick all that apply)	✔ I have personal experience of the condition ☐ I have personal experience of the technology being appraised ☐ I have other relevant personal experience. Please specify what other experience: ✔ I am drawing on others' experiences. Please specify how this information was gathered: As a patient carer and also Trustee and Vice Chair of GIST Support UK since 2010 I have engaged with many GIST patients and carers and as a result have an in-depth first-hand understanding of how patients' lives are impacted by GIST cancer. GIST is the most common of the cancers classified as a Sarcoma. I understand that in addition to GIST there are other Sarcomas that carry the NTRK fusion e.g. retroperitoneal and fibrosarcoma.
Living with the condition	
8. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	I have been a carer for my daughter who was diagnosed as a GIST patient in 2010 at the age of fifteen. This diagnosis prompted me to contact GIST Support UK for help as the information available via our doctors was minimal because they had never dealt with a GIST patient before. I understand that this is quite a common occurrence for GIST and Sarcoma patients. GIST Support UK were able to introduce me to other families who had children with his cancer and since then I have become a trustee and now vice chair of the charity.

	When I joined it was obvious that because it is a rare cancer there were no treatments available for my daughters' type of GIST. It made us all feel hopeless especially when the UK's leading specialists confirmed that there is no treatment or cure and at that time no research either, but to stay in that frame of mind was not tolerable, so we decided to try and find a way to change it and as a result our family life altered completely. My role became both carer and patient advocate working to improve the outlook for my daughter and others in the same situation as her. This has had a huge impact on every aspect of our life including income as I was not able to continue my career as a senior director in the corporate world of IT recruitment. This is obviously not every carers experience but very often family income is impacted because either the patient or carer are not able to work in the same capacity as before their diagnosis and there is the expense related to regular hospital visits which can be a long way from home to see the relevant specialist. Living with cancer is possible but every day that you wake up you hope that it was a bad dream and that it isn't real. I understand that this is a standard defence mechanism for cancer patients and their families. Learning to cope is something that you have to do and the last thing that you want to do as a carer is to give the impression that things will not be OK. You have to give your loved one hope. It is the standard pre-disposition for a mother to defend her young. A cancer diagnosis for your baby brings all of these instincts to the fore, beyond all imagination. My daughter had her stomach removed in 2013 and then went on to study a Business degree at Aston University, she coped with regular anaemia and blood transfusions, surgery to remove her stomach which then caused osteoporosis, bone fractures requiring surgery, dramatic weight loss and food intolerances. Eventually she died with a liver that was outgrowing the space in her body due to tumour burden and kidney failure as a result of contracting norovirus. The traumas and horrors of my daughter's life with GIST cancer have totally shattered what should have been the natural and happy course of my family's lives. Her father and brother, grandparents, aunts, uncles, cousins and friends and I have all been traumatised by the experience. Life will never be the same again. Our sense of loss is beyond comprehension. I know from first-hand experience and from engaging with other GIST patients, carers and families over the past nine years that they experience much the same. Carers take many forms, parents, partners, siblings, children and friends, all desperate to help and

	save the person that they love. A cancer diagnosis is the last thing that you think will happen to you or someone you love. Many GIST and Sarcoma patients manage, with effective treatment, to live relatively normal lives, continuing to work and play as best they can while manging the side effects of treatment. There are some who are fortunate that their GIST or Sarcoma is found early and before it has spread, they have it removed while still small and it does not return. This is as close to a cure as currently exists. For those where surgery does not totally remove the cancer, treatment side effects are invariably; recovery from surgery. Sometimes this can mean quite drastic interventions such as removal of the stomach, where the disease has reached an advanced stage before diagnosis. Toxic side effects are also encountered from anticancer therapies. Drug side effects include debilitating hand foot syndrome, diarrhoea, skin rashes the list is quite extensive but all of these are tolerated and coped with by patients for the chance to live longer and do the normal things you would expect to do when living. Patients are often forced to defer and put their lives on hold due to GIST and Sarcoma. Most cancer patients are courageous in adversity. The desire to live is overwhelmingly strong. This gives carers such as myself strength to carry on. In order to do this, most other things in my life took a back seat or disappeared and the hospital became our second home and the nurses and doctors our new family and friends**.** All carers experience stress and pressure when supporting a cancer patient. Most try hard to lead a normal life but hospital appointments and awaiting news which can be sometimes bad all takes its toll on the mental and physical health of carers as well as the patients.
Current treatment of the condition in the NHS	
9. What do patients or carers think of current treatments and care available on the NHS?	Patients and carers are very grateful for the treatments that are available via the NHS and eagerly seek out the possibility of new treatments that might be suitable for their type of cancer. Currently for GIST patients this consists of: • Surgery • Imatinib • Sutent

	• Regorafenib Unfortunately, not all GIST cancers are the same and there are many for whom the above treatments are not effective because either their primary mutation is not targeted by the above treatments or their disease metastasizes beyond the control of the above treatments. We are very grateful that there is some research happening in the world and that a treatment has been discovered for those GIST and Sarcoma patients who have an *NTRK* fusion mutation. Currently such a treatment is not available via the NHS but we hope that further to this appraisal that it will be available for patients with GIST and Sarcoma caused by *NTRK* fusion mutations.
10. Is there an unmet need for patients with this condition?	Yes. *NTRK* mutated GIST and Sarcoma patients do not currently have access to a targeted treatment for their type of mutation via the NHS.
Advantages of the technology	
11. What do patients or carers think are the advantages of the technology?	• The advantages of this technology are that it is a precision medicine designed to target the *NTRK* fusion mutation which is the cause of the cancer. It blocks the signalling pathway so that the tumours cannot grow and is effective even when patients have been heavily pre-treated with other therapies. • Larotrectinib is administered orally in capsule or liquid form and studies to date show that it is well tolerated and has shown encouraging anti-tumour activity in all patients with TRK fusion positive tumours. In one trial 75% of patients remained on Larotrectinib and some were able to have surgery with curative intent as their tumours had shrunk sufficiently to make them operable. • Larotrectinib is suitable for both adults and children. Drugs of this type are exactly what rare cancer patients are desperate to find and use to shrink and stop their tumours and get their life back on track.

Disadvantages of the technology	
12. What do patients or carers think are the disadvantages of the technology?	The only disadvantage that I can see is being an NTRK fusion cancer patient and not being able to access Larotrectinib. As with any drug there are side effects but those listed are tolerable and can be managed.
Patient population	
13. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	I understand from a GIST perspective that the patients currently classified as Quadruple negative GIST patients are the target group where NTRK fusion GIST's can be found. They have been named quadruple Wildtype GIST as they lack abnormalities in the four signalling pathways KIT, PDGFRA, SDH or RAS. We are already screening all quadruple negative GIST patients who attend the PAWS-GIST clinic at Addenbrookes Hospital in Cambridge to find the ones that have NTRK fusions. There are many different types of sarcoma and NTRK fusions have already been discovered in some of these including retroperitoneal and fibrosarcoma. When the whole genome sequencing starts as standard within the NHS this year it will speed up the discovery of patients whose cancer is caused by NTRK fusions for whom Larotrectinib is effective.

Equality	
14. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	The only inequality I can see currently, is that the world is fast tracking NTRK fusion inhibitors to be available to patients with NTRK fusions and until this appraisal has concluded Larotrctinib is not available to the patients who will benefit.
Other issues	
15. Are there any other issues that you would like the committee to consider?	
Topic-specific questions	
16. [To be added by technical team if required, after receiving the company submission. For example, if the company has deviated from the scope (particularly with respect to comparators) – check whether this is appropriate. Ask	

specific, targeted questions such as “Is comparator X [excluded from company submission] considered to be established clinical practice in the NHS for treating [condition Y]?”] **if not delete highlighted rows and renumber below**	

Key messages

17. In up to 5 bullet points, please summarise the key messages of your statement:

- NTRK fusions are the root cause of some GIST’s and Sarcomas.
- Larotrectinib is a precision medicine that targets NTRK fusion mutations.
- Trials have resulted in dramatic results for GIST and Sarcoma patients with NTRK fusions
- The NHS whole genome sequencing being launched this year will identify the patients with NTRK fusions
- Using Larotrectinib in GIST and Sarcoma patients with an NTRK fusion will reduce unnecessary expenditure on other ineffective therapies that are very expensive.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

……

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

NICE National Institute for Health and Care Excellence

National Institute for Health and Care Excellence

Cancer Drugs Fund Clinical Lead statement

Larotrectinib for treating NTRK fusion-positive solid tumours [ID1299]

Background

Tumour agnostic drugs

1. NTRK inhibitors are the first tumour agnostic drugs which are expected to be licensed in Europe but others are likely to follow in the next few years. There is evidence of benefit for anti PD-L1 immunotherapy in cancer patients whose tumours exhibit microsatellite instability-high or mismatch repair deficiency or high tumour mutational burden. There are clinical trials in other drugs targeting NTRK gene fusion cancers and also resistance to 1st generation NTRK inhibitors. A number of basket clinical trials are running in cancer patients with other mutations or gene fusions (e.g. RET, FAP etc).

Incidence of NTRK gene fusions

2. There is an emerging evidence base as to the incidence of NTRK gene fusions. Some very rare cancers have high (80-100%) proportions with NTRK gene fusions (e.g. the mammary analogue secretory variant of salivary gland cancer, the secretory variant of breast cancer, paediatric mesoblastic nephroma, infantile fibrosarcoma). Some rare cancers have modest (20-40%) proportions of NTRK gene fusions (e.g. paediatric non-brain stem glioblastoma, spitzoid melanoma) or low (2-12%) incidences (e.g. papillary thyroid cancer, some brain malignancies, cholangiocarcinoma, gastrointestinal stromal tumours). Most cancers and all the commoner

cancers have very low proportions of NTRK gene fusions of 1% or less.

3. NHS England and NHS Improvement notes that in the larotrectinib submission from Bayer, evidence is presented which shows NTRK gene fusion to be evident in <1% of patients with solid cancers.

4. NHS England and NHS Improvement concludes that on the current evidence and when all solid tumours are considered, it is reasonable to assume an incident proportion of between 0.5 and 1% with NTRK gene fusions. NHS England and NHS Improvement therefore considers that a base case figure of 0.5% should be used in this appraisal and a scenario analysis be done at a 1% incidence.

Natural history of cancers with NTRK gene fusions

5. Little is known as to the natural history of NTRK gene fusion positive varieties of solid tumours. Bayer in its submission mentions some preliminary evidence of worse prognosis with NTRK fusion positive disease in colorectal and papillary thyroid cancers. However, in the example of metastatic colorectal cancer, the outlook in patients with NTRK/ROS1/ALK genetic changes (n=27) is worse than those without such changes (n=319). However, this is not a pure NTRK gene fusion group and the incidence of NTRK gene fusion in colorectal cancer is thought to be <1%. The contribution of the ALK and ROS1 patients to this adverse outcome could explain much of this apparent difference. NHS England and NHS Improvement recognises that there may be a difference in outlook for incurable patients with metastatic cancer who have NTRK gene fusions but there is no robust evidence to support this at present.

Draft marketing authorisation

6. The draft marketing authorisation shows that larotrectinib is indicated in NTRK gene fusion positive solid cancers in adult and paediatric

patients who have locally advanced or metastatic disease and who have no further satisfactory systemic treatment options or where surgical resection is likely to result in severe morbidity. Similar wording was used in the main phase II studies of larotrectinib which delivered nearly all of the 102 analysed patients in the pooled analysis. The definitions of 'no further satisfactory' therapies and 'surgical resection likely to result in sever morbidity'' are very important. The phrase 'no further satisfactory therapies' is particularly open to potentially variable interpretation between oncologists and between oncologists and their patients.

Generalisability of the trial population as regards clinical benefit

7. NHS England and NHS Improvement notes that [redacted] of the 93 solid tumour patients were treatment naïve for chemotherapy. NHS England and NHS Improvement therefore considers that there is a potentially considerable bias in these early larotrectinib studies. This is as a result of the inclusion of patients who knew they had a NTRK fusion cancer and wished to have the opportunity of receiving entrectinib whilst the trial was open and they were eligible for treatment. It may be therefore that 'standard therapies' had not been fully explored. In addition, the larotrectinib studies have patients which are biased in terms of rare cancers figuring significantly e.g. sarcomas [redacted], salivary gland cancers [redacted], thyroid cancers [redacted] and infantile fibrosarcoma [redacted] and there is little representation from the common cancers such as non-small cell lung cancer ([redacted]) and colorectal cancer ([redacted]). There is therefore uncertainty as to the generalisability of the larotrectinib clinical data into widespread NHS use across all malignancies.

Activity and toxicity of larotrectinib

8. Larotrectinib is clearly a very active drug in NTRK fusion positive malignancy. It cannot be directly compared with entrectinib for

response rate, progression-free survival and overall survival in view of the differing case and age mix in the respective pooled analyses e.g. in terms of proportions of tumour types treated with larotrectinib versus entrectinib, non-small cell lung cancer contributes [redacted] breast cancer [redacted] infantile fibrosarcoma [redacted], melanoma [redacted] etc. Larotrectinib does have a substantially greater number of paediatric patients ([redacted]) in its pooled solid tumour analysis than Roche has for entrectinib ([redacted]).

9. The clinical impact of larotrectinib is striking but the median duration of follow-up is only [redacted], the number treated and evaluable is small and the numbers of patients with specific cancers are very small. Of note too is that larotrectinib is clearly active in patients with primary brain tumours although the number treated is very small ([redacted]).

10. Larotrectinib was reasonably well tolerated with rates of dose delays or interruptions of [redacted] and a treatment-related discontinuation in [redacted]

11. Currently, systemic therapy is organised around tumour site-specific teams as knowledge and experience of the natural history of individual cancers is very important in the optimal care of patients. The rarity of NTRK gene fusions in most cancers means that individual oncologist experience in the use of larotrectinib will be very small. Consideration will therefore have to be given within cancer centres of sharing experience of larotrectinib use in order to assist in the best management of side-effects.

The treatment pathway and comparators

12. The issue of where in the treatment pathways patients would be treated with larotrectinib is an important one, partly as it determines what the comparator costs should be but mainly because it resolves what the comparator durations of survival should be. This is because Bayer has submitted a naïve weighted comparison of outcomes with

larotrectinib versus what it believes to be the correct comparator albeit in populations of patients with unknown NTRK gene fusion status. The company has chosen best supportive care as the comparator for lung, colorectal and thyroid cancers as well as most adult sarcomas and NHS England agrees with this approach. However, for example, to compare larotrectinib in cholangiocarcinoma on the basis of [redacted] treated with the standard therapy of cisplatin plus gemcitabine is unrealistic, as is to compare larotrectinib in breast cancer on the basis of [redacted] treated with the standard therapy of taxanes.

13. NHS England and NHS Improvement has set out these details as a weighted cost, progression free (PFS) and overall survival (OS) analyses have been used in the comparison with larotrectinib. Whilst the costs of the comparator arm should be reduced to reflect best supportive care where appropriate, so should the survival outcomes be used for best supportive care where appropriate. It is difficult to follow in Bayer's submission to NICE that the rigour of consistent analysis has been applied.

Pooling of the larotrectinib studies

14. NHS England and NHS Improvement supports the pooling of the 4 larotrectinib studies in order to maximise the patients included in the analyses of clinical and cost effectiveness. This pooling has consequences for the case mix of tumours (see later).

Cost effectiveness

Parametric extrapolation

15. Given the immaturity of the larotrectinib data, NHS England and NHS Improvement recognises the need for parametric extrapolation of the data on progression-free and overall survivals. NHS England and NHS Improvement considers that the exponential extrapolation for

both progression free and overall survivals is just as clinically plausible as the company-choice of the Weibull. NHS England and NHS Improvement notes that use of the exponential for overall survival would significantly increase the ICER.

Generalisability as regards costs

16. NHS England and NHS Improvement again notes that the cost effectiveness analysis for the comparator population is based on which cancers were treated in the larotrectinib pooled analysis. Given the tumour agnostic marketing authorisation that is expected for larotrectinib, it is highly likely that the case mix of the NHS England and NHS Improvement treated population will significantly differ from the biased case mix of the pooled larotrectinib analysis e.g. the real world NHS England and NHS Improvement population will not be constituted by a [redacted] proportion made up by patients with the paediatric soft tissue sarcoma or a [redacted] proportion of salivary gland carcinoma or a [redacted] proportion of lung cancer or a [redacted] proportion of breast cancer. Put another way and to illustrate the same point, the economic model assumes a patient case mix of [redacted] adults and [redacted] paediatrics. It is likely that a real-world patient case mix would increase the ICER by both reducing the incremental survival and by increasing the costs of ascertaining NTRK gene fusion e.g. from a very small cost for testing such salivary gland tumours (NTRK gene fusion present in 90-100%) versus lung cancer (NTRK gene fusion present in 1% or so).

Utilities

17. NHS England and NHS Improvement notes that the mean utility values gained from the larotrectinib pooled analysis for the progression free and post progression survival states of the economic model were [redacted] and [redacted], respectively, yet the corresponding figures for the comparator population were [redacted] and [redacted] , respectively. It is counterintuitive for the progression free state utility values for two so-

called identical populations to differ so much and then for this divergence in utilities to increase at progression is also stretching credulity. These differing utility values between the two types of treatment greatly increase the QALY gain for larotrectinib when the most likely interpretation of these divergent results is that like is not being compared with like. It is NHS England and NHS Improvement's view that Bayer recognises the weakness of their utility data and do the appropriate scenario analyses with similar utility values for each respective health state.

Costs of chemotherapy

18. In terms of drug administration costs, Bayer has omitted any chemotherapy tariff costs for any oral treatment: this is important for larotrectinib which has a substantial mean treatment duration. The SB11Z oral chemotherapy tariff (£120 per visit) should have been used and this incremental cost applies almost completely to the larotrectinib arm.

19. In addition, Bayer estimates the weekly comparator chemotherapy costs for paediatric soft tissue sarcoma and infantile sarcoma as ranging from [redacted]. This is a very wide range and requires detailed justification. In addition, the treatment duration for this group of patients is 350 days which seems remarkably long for a non-standard therapy which is 'not satisfactory'.

NTRK gene fusion testing

20. Proof of a NTRK gene fusion requires either whole genome sequencing (WGS) or next generation sequencing (NGS), the latter providing the technology for multigene panels (which provide testing for anything between 5 and 500 genes). There are some screening TRK immunohistochemistry tests which greatly reduce the need for NGS but these also have a significant false negative rate.

21. As part of the establishment of the NHS Genomic Medicine Service (including the Genomic Laboratory Hubs), NHS England and NHS Improvement are making fundamental changes to how cancer genomic testing is provided, commissioned and funded. A national service has been created and is regionally organised by 7 Genomic Laboratory Hubs. The hubs are responsible for processing samples for WGS, performing NGS testing and interpreting all NGS and WGS results before returning the results to the requesting clinician. The WGS is done by Genomics England which receives samples from and returns WGS results to the hubs. 2019-20 is a critical set up year for the Genomic Laboratory Hubs for both their establishment and for the diversion of previous genomic funding from the many hospitals who have done a variety of gene testing until now. The NHS England and NHS Improvement Genomics Medicine Service is the first national service to be set up in the world: its ambition is matched by the revolution occurring in the organisation and funding of the 7 Genomic Laboratory Hubs and in the types of NGS now becoming available.

22. During 2019, the NHS will start to offer whole genome sequencing for patients with paediatric cancer and for adults with all types of sarcoma. The current timeline for the start of the WGS operation is end of the summer of 2019, however full implementation will take time (NHS England and NHS Improvement's working assumption is that it will be the autumn of 2020 before all WGS pathways are fully operational across the country). Full implementation requires significant changes to the diagnostic pathway including the establishment of pathways of care such that fresh frozen tissue can be processed by the Genomic Laboratory Hubs in a timely fashion so that DNA of the appropriate quality is obtained before then being sent to Genomics England for testing. Funding from NHS England and NHS Improvement is in place for the provision of WGS for paediatric cancer and sarcoma although it is recognised that NGS may be

necessary for NTRK fusion testing in the short term until WGS is fully operational.

23. For some rare tumours, such as mammary analogue secretory carcinoma of the salivary gland and the secretory variant of breast cancer, the National Genomic Test Directory for 2019 already sets the expectation that NTRK testing should be performed. Although NHS England and NHS Improvement does not have robust data about existing testing activity, it is aware that cancer genomic testing for such a test is not currently performed systematically across the country. Funding, however, is place for the NTRK gene fusion testing for these 2 rare cancers.

24. In all other adult solid cancers, NTRK gene fusion testing is not currently required by the National Genomic Test Directory and is not systematically performed. However, by the end of the 2019/20 financial year, the Genomic Laboratory Hubs plan to introduce gene panels for solid tumour testing, which will include the capability to identify NTRK gene fusions. This could be for example with a 50-60 gene panel (cost ~£250) or a 500 gene panel (cost ~£400). To facilitate testing for NTRK gene fusion in solid tumours, NHS England and NHS Improvement will need to include NTRK gene fusion testing in the National Genomic Test Directory and determine the funding required. Some of the Genomic Laboratory Hubs are currently more advanced in their ability to deliver NGS multigene panel testing and hence there is likely to be some initial sharing of NGS testing until all 7 of the hubs are fully operational.

25. As is clear from the preceding paragraphs and apart from the rare cancers in which NTRK gene fusions are more commonly expressed, large numbers of patients have to be screened to find the NTRK gene fusion. For a tumour agnostic drug which has a high chance of benefitting patients who harbour the NTRK gene fusion, the logical potentially eligible population is in all patients with solid cancers which

are incurable (i.e. the patients who have locally advanced or metastatic disease). Some cancers already have some genetic testing embedded in the treatment pathway (e.g. melanoma and lung, colon, thyroid, breast and ovarian cancers). For patients and clinicians to be able to best use the information of NGS panel testing, such testing has to be done prior to the initiation of all systemic therapy for the locally advanced/metastatic disease. The cost of NGS panel testing is therefore very great as NHS England and NHS Improvement estimates that it would need to test approximately 100,000 patients in all. About 3,000 will be eligible for WGS and 30,000 already receive some genomic testing as part of existing standard of care (and this is assumed to cover the cost of NGS panel testing at least in melanoma and lung and colorectal cancers). Thus 67,000 patients represent additional and new activity. The estimated assay cost of this new activity would be £16.8m if the 50-60 gene panel is used, £26.8m if the 500 gene panel is used and £21.8m if an average cost of £325 per multigene test is used. If the £325 figure is used and since 33% of the total testing cohort is assumed to already receive testing, this means that the average incremental diagnostic cost per patient tested is £218. If the incidence of NTRK gene fusion is 1 in 200, then the total cost per positive NTRK gene fusion patient is £43,500. If the incidence of NTRK gene fusion is 1 in 100, then the total cost per positive NTRK gene fusion patient is £21,800. If WGS initially does not deliver information within a timetable required for clinical decision making and at first all 3000 patients have to have NGS, then the average incremental diagnostic cost per patient tested would initially be £227 with a cost per positive NTRK fusion of £45,000 if its incidence is 1:200 and £21,500 if 1:100.

26. In addition to the costs of WGS and gene panel testing, there are capital costs to consider: laboratory equipment, bioinformatics and the increased need for expert interpretation of results to aid clinical decision-making. NHS England and NHS Improvement is currently

working through these issues as the 7 Genomic Laboratory Hubs are starting from different baseline positions.

27. In summary, it is anticipated that WGS will be fully operational by Q2 2020/21 and panel testing will be available by Q1 2020/21. Uptake of molecular testing across the 7 genomic hubs will increase during 2020/21 as genomic pathways are embedded and links are made with the clinical teams. Given the complexity of implementation, it may take a further 12 months for molecular testing to become fully embedded in practice.

Costing of NTRK gene fusion testing for this appraisal

28. The established approach in NICE technology appraisals of cancer drugs which require genomic testing has been to ensure that the full cost of the testing has been included in the cost effectiveness analysis. In the appraisal of larotrectinib, there are 4 important differences to previous appraisals of targeted cancer drugs which have required genomic testing. Firstly, larotrectinib is a tumour agnostic drug and hence all cancers have to be tested as there is currently no evidence to indicate that certain cancers never have NTRK gene fusions. The consequence of this is that the number of patients to be tested is very great. Secondly, the incidence of NTRK gene fusions in most solid tumours is very low. Thirdly, the need for NTRK fusion testing is coming at a critical set up time for a new and national genomic medicine service in England. Fourthly, this new service must embed at set-up the technologies which will it will need to provide the huge benefits of such a national service i.e. a service for WGS and NGS panel tests has to be built.

29. NHS England and NHS Improvement recognises that the national availability of WGS and NGS multi-gene panel testing for patients with incurable solid tumours will bring many future treatment opportunities: not only for NTRK fusion inhibitors but for other tumour agnostic

cancer drugs (several of which are likely in the next few years), for the many expected future targeted drugs which will require genomic testing in patients with specific tumours and for greater patient entry into clinical trials. It should be noted that the current plans for NHS investment in these genomic services is primarily for improving geographical equity of access and pump priming the new genomics infrastructure. NHS England and NHS Improvement therefore considers that it is appropriate that at least part of the cost for multi-gene panel testing be covered by each company that benefits from this new service provision (in line with the standard approach employed in technology appraisals). As a consequence, NHS England and NHS Improvement would wish NICE to explore scenario analyses in its appraisal of the cost effectiveness of larotrectinib in which various percentages of the costs of multi-gene panel testing are borne by larotrectinib: 100%, 50%, 33%, 25% and 0%.

30. To reach a proportionate and reasonable position on how much of this cost should be borne by the NHS vs an individual company with a tumour agnostic product, NHS England and NHS Improvement will wish to see these scenario analyses and will decide on the appropriate level of contribution by October 2019, i.e. in advance of the final point of submission before the NICE committee considers larotrectinib in its November 2019 meeting.

Bayer costing of detecting NTRK gene fusions

31. Bayer has not included any cost of testing for NTRK gene fusions in its base case as it considers that all testing is routine as it has been included in the NHS Long Term Plans. This is an unrealistic assumption.

Base case modelling by health state

32. [redacted]

End of life cost effectiveness threshold

33. NHS England and NHS Improvement agrees that larotrectinib would satisfy NICE's End of Life threshold criteria given that using a weighted average of survival for the comparator is a reasonable approach. NHS England and NHS Improvement also believes that survival may be overestimated as Bayer has assumed use of larotrectinib at earlier points in the treatment pathway in some diseases (see above for discussion of this).

Cancer Drugs Fund

34. NHS England and NHS Improvement supports Bayer's aim for larotrectinib to enter the Cancer Drugs Fund. NHS England and NHS Improvement regards larotrectinib as a highly promising drug which needs clinical data of much greater maturity and testing in a real world setting across many cancers and in much greater numbers. NHS England and NHS Improvement is baffled therefore that in Bayer's own economic analysis, it has [redacted]: with the full QALY weighting of the End of Life threshold (i.e. at £50,000 per QALY), Bayer's base case deterministic ICER is [redacted] at the discounted larotrectinib price.

Implementing a positive NICE recommendation

NICE recognises that in the event of a positive recommendation, more prescriptive clinical commissioning criteria for treatments commissioned via Specialised Services will be implemented by NHS England and NHS Improvement to ensure appropriate use within the NHS.

NHS England and NHS Improvement is responsible for ensuring that the final clinical commissioning criteria are aligned with final guidance (section 1 – recommendation and section 3 – committee discussion).

Draft commissioning criteria

35. If larotrectinib for treating NTRK gene fusion locally advanced/metastatic solid tumours is recommended for use within its marketing authorisation, NHS England and NHS Improvement proposes to use the following commissioning criteria:

- The patient's cancer must have the presence of an NTRK gene fusion as determined by WGS or following a NHS multigene panel test
- The patient must have locally advanced or metastatic disease
- The patient must have progressed following treatment with all NICE-recommended systemic therapies or established standard therapies in clinical practice or have a documented ineligibility for such treatments or where surgical resection is likely to result in severe and permanent morbidity (such as limb amputation, facial disfigurement and a paralysis-causing procedure)
- The patient must have an ECOG performance score of 0-2
- If the patient has metastases in the central nervous system, then these must be asymptomatic if untreated or treated and controlled
- Larotrectinib is to be used as monotherapy
- The prescription of larotrectinib and care of the patient on larotrectinib to be by a consultant oncologist specifically trained and accredited in the use of systemic anticancer therapy
- The patient is to be treated until progressive disease or unacceptable toxicity or the patient choice to discontinue treatment or surgical resection, whichever is the sooner.

If this technology is recommended for routine commissioning in a subpopulation or with certain specifications (for example, a treatment continuation rule), the final commissioning criteria will reflect these conditions.

36. If larotrectinib for treating NTRK gene fusion positive locally advanced or metastatic cancer is recommended for use in the Cancer Drugs Fund, the final commissioning criteria will reflect the patient eligibility criteria in the managed access agreement.

Issues for discussion

37. These have all been outlined above.

Issues for decision

38. These relate to the above and principally relate to:

- Incidence of NTRK gene fusion cancers
- Interpretation of the wording of the marketing authorisation
- Generalisability of the trial population
- Treatment pathway and comparators
- Parametric modelling of progression free and overall survivals
- Utilities in the progression free and post progression health states
- Costs of chemotherapy
- NTRK gene fusion testing, implementation and costs
- CDF entry

Equality

39. NHS England and NHS Improvement recognises that the 7 Genomic Laboratory Hubs are at different stages of being able to implement NGS multigene panel testing and this variation will be resolved over the next 1-2 years. NHS also recognises that WGS will take time to embed within clinical treatment pathways, particularly in respect of the need for the collection and processing of fresh tissue.

Author

Professor Peter Clark, NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund

July 2019

CONFIDENTIAL UNTIL PUBLISHED
Evidence Review Group's Report

Larotrectinib for treating NTRK fusion-positive advanced solid tumours

Produced by	CRD and CHE Technology Assessment Group, University of York, Heslington, York YO10 5DD
Authors	Ana Duarte, Research Fellow, CHE Mark Corbett, Research Fellow, CRD Alessandro Grosso, Research Fellow, CHE Ruth Walker, Research Fellow, CRD Melissa Harden, Information Specialist, CRD Laura Bojke, Reader in Health Economics, CHE Mark Simmonds, Senior Research Fellow, CRD
Correspondence to	Mark Simmonds, Senior Research Fellow, CRD, University of York, York YO10 5DD
Date completed	Date completed 23/08/2019

Source of funding

This report was commissioned by the NIHR HTA Programme as project number 127895. NICE project ID 1299.

Declared competing interests of the authors

None

Acknowledgements

We would like to thank Dr Robert Metcalf, Honorary Consultant in Medical Oncology at The Christie NHS Foundation Trust, Professor Christian Ottensmeier, Chair in Experimental Cancer Medicine at University Hospital Southampton, Dr Helene Schlecht, Principal Clinical Scientist at Manchester University NHS Foundation Trust, and Marta Pereira, Clinical Scientist at Manchester University

NHS Foundation Trust for their expert advice. We would also like to thank Professor Stephen Palmer of the Centre for Health Economics for his advice.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Duarte A, Corbett M, Grosso A, Walker R, Harden M, Bojke L, Simmonds M. Larotrectinib for treating NTRK fusion-positive advanced solid tumours: A Single Technology Appraisal. CRD and CHE, University of York, Technology Assessment Group, 2019

Contributions of authors

Mark Corbett, Ruth Walker and Mark Simmonds wrote the clinical effectiveness sections of the report. Ana Duarte, Alessandro Grosso and Laura Bojke wrote the cost effectiveness sections and conducted the ERG economic analyses. Melissa Harden wrote the search strategy sections. Mark Simmonds took overall responsibility for the clinical effectiveness sections. Laura Bojke took overall responsibility for the cost effectiveness sections.

Note on the text

All commercial-in-confidence (CIC) and academic-in-confidence (AIC) data have been redacted.

Copyright statement

Table of Contents

List of abbreviations 11
1 Summary 13
1.1 Critique of the decision problem in the company's submission 13
1.2 Summary of clinical effectiveness evidence submitted by the company 14
1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted 15
1.4 Summary of cost effectiveness submitted evidence by the company 16
1.5 Summary of the ERG's critique of cost effectiveness evidence submitted 18
1.6 ERG commentary on the robustness of evidence submitted by the company 19
1.6.1 Strengths 19
1.7 Weaknesses and areas of uncertainty 20
1.7.1 Trials of larotrectinib 20
1.7.2 Indirect comparisons 20
1.7.3 Identifying NTRK fusions 21
1.7.4 Uncertainty in the extrapolation of larotrectinib survival outcomes 22
1.7.5 Potential confounding of subsequent lines of therapy 22
1.7.6 Uncertainty surrounding the homogeneity of the treatment effect 22
1.8 Summary of exploratory and sensitivity analyses undertaken by the ERG 23
1.8.1 Re-analysis of included trials 23
1.8.2 Adjustments to the cost-effectiveness model 23
2 Background 26
2.1 Critique of company's description of underlying health problem. 26
2.1.1 Eligible population and prevalence of NTRK 26
2.2 Critique of company's overview of current service provision 30
2.2.1 Treatment pathways 30
2.2.2 NTRK fusion diagnostic pathways 30
2.2.2.1 Current provision for NTRK fusion screening 31
2.2.2.2 Testing strategies for detection of NTRK fusions 33
2.2.2.3 Diagnostic accuracy for detection of NTRK fusions 34
3 Critique of company's definition of decision problem 37
3.1 Population 37
3.2 Intervention 39
3.3 Comparators 39
3.4 Outcomes 40
4 Clinical Effectiveness 41
4.1 Critique of the company's systematic review methods 41

4.1.1 Searches 41
4.1.2 Inclusion criteria 42
4.1.3 Critique of data extraction 42
4.1.4 Quality assessment 42
4.1.5 Evidence synthesis 43
4.2 Critique of trials of the technology of interest, their analysis and interpretation 44
4.2.1 Design and analysis of basket trials 44
4.2.2 Details of the included trials 46
4.2.3 Summary of the pooled dataset effectiveness results 51
4.2.3.1 Overall response rate 51
4.2.3.2 Time to response and duration of response 52
4.2.3.3 Progression free survival and overall survival 53
4.2.3.4 Health-related quality of life 56
4.3 Critique of systematic review for indirect comparison analyses 56
4.3.1 Searches 57
4.3.2 Selection of TAs and trials 58
4.4 Critique of the indirect comparison analyses 58
4.4.1 Comparison to previous NICE TAs and trials 58
4.4.2 Responder/non-responder analysis 60
4.4.3 Growth Modulation Index (GMI) analysis 62
4.5 Adverse events 63
4.5.1 Adverse events of any cause 63
4.5.2 Adverse events judged to be related to larotrectinib 63
4.5.3 Adverse events in comparator trials 64
4.6 Additional work on clinical effectiveness undertaken by the ERG 64
4.6.1 Modelling response rates across tumour sites 64
4.6.1.1 Results 66
4.6.1.2 Primary CNS tumours 68
4.6.1.3 Regression models 68
4.6.2 Exploring heterogeneity in time-to-event outcomes across tumour sites 68
4.6.2.1 Progression-free survival 69
4.6.2.2 Overall survival 71
4.7 Conclusions of the clinical effectiveness section 73
4.7.1 Systematic review and trials of larotrectinib 73
4.7.2 Indirect comparisons 74
4.7.3 Analyses conducted by the ERG 75

4.7.4 Identifying NTRK fusions 76
4.7.5 Summary 76
5 Cost Effectiveness 78
5.1 ERG comment on company's review of cost-effectiveness evidence 78
5.1.1 Searches for cost-effective analyses of treatments for patients with TRK-Fusion cancer 78
5.1.2 Searches for cost-effectiveness evidence of treatments for patients with solid tumours that are known to harbour NTRK gene fusions 78
5.1.3 Inclusion/exclusion criteria used for study selection 80
5.1.4 Studies included and excluded in the cost effectiveness review 80
5.1.5 Conclusions of the cost effectiveness review 81
5.2 ERG's summary and critique of company's submitted economic evaluation 81
5.2.1 Model structure 81
5.2.2 The company's economic evaluation compared with the NICE reference case checklist 89
5.2.3 Population 90
5.2.3.1 Larotrectinib population 90
5.2.3.2 Comparator population 92
5.2.4 Intervention and comparators 94
5.2.4.1 Intervention 94
5.2.4.2 Comparator 94
5.2.5 Perspective, time horizon and discounting 101
5.2.6 Treatment effectiveness and extrapolation 102
5.2.6.1 Larotrectinib treatment effect 102
5.2.6.2 Comparator treatment effect 104
5.2.6.3 Progression free survival 105
5.2.6.4 Overall survival 112
5.2.6.5 All-cause mortality 122
5.2.6.6 Adverse events 123
5.2.7 Health related quality of life 126
5.2.7.1 Systematic review of HRQoL studies 126
5.2.7.2 Health State Utilities 127
5.2.7.3 Adverse Events Utilities 131
5.2.8 Resources and costs 132
5.2.8.1 Systematic review of resource use and costs 132
5.2.8.2 Drug acquisition and administration costs 134
5.2.8.3 Health state unit costs and resource use 138
5.2.8.4 Adverse events 142

5.2.8.5 NTRK fusion testing 142
5.2.9 Cost effectiveness results 143
5.2.9.1 Base-case results 143
5.2.9.2 Sensitivity analyses 145
5.2.10 Model validation and face validity check 156
5.3 Conclusions of the cost effectiveness section 157
6 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG 161
6.1 Overview 161
6.2 ERG adjustments to the company's base case model 161
6.3 Additional ERG analyses 162
6.3.1 ERG adaptation of the company cost-effectiveness model 162
6.3.2 ERG additional analyses 164
6.3.3 Extrapolation of survival outcomes 167
6.3.4 Larotrectinib post-progression survival gains 172
6.3.5 Post-progression utility for larotrectinib 175
6.3.6 Heterogeneity in response rates 176
6.4 ERG base case model 177
6.5 Further exploratory analysis of heterogeneity using the ERG base-case analysis 177
6.5.1 Why heterogeneity matters 177
6.5.2 Data required to illustrate heterogeneity 180
6.5.3 Exploring heterogeneity in response rates according to tumour type 181
6.5.4 Exploring heterogeneity in testing costs according to tumour type 183
6.6 Conclusions from ERG analyses 183
7 End of life 186
8 Overall conclusions 188
8.1 Clinical effectiveness 188
8.1.1 Trials of larotrectinib 188
8.1.2 Indirect comparisons 189
8.1.3 Identifying NTRK fusions 189
8.2 Implications for research 190
8.3 Cost-effectivness 191
9 References 193
10 Appendices 198
10.1 Philips checklist 198
10.2 Company base-case analysis with PAS price 205
10.3 Response-based survival models 206

10.4 NTRK fusion testing costs 208
10.5 ERG's exploratory analyses: PAS price 211

Table of Tables

Table 1 Summary of ERG exploratory analyses 24

Table 2. Number of patients who have a NTRK fusion, who are eligible for testing and treatment with larotrectinib and the number needed to screen (NNS), by cancer type. 29

Table 3. Testing methods currently available of the NHS for those cancers that harbour NTRK fusions 31

Table 4 Impact of imperfect diagnostic testing on larotrectinib efficacy 36

Table 5 Summary of the larotrectinib trials methods (adapted from Table 5 of the CS) 47

Table 6 Baseline characteristics for the efficacy evaluable patients: pooled analysis datasets and individual larotrectinib studies (adapted from Table 8 of the clarification response) 49

Table 7 Number of responders by tumour type (adapted from CS, Appendix E, Table 82) 65

Table 8 Probabilities that response rate exceeds 10% or 30% from BHM 66

Table 9 Overview of the company's economic evaluation 82

Table 10 Comparison of company's economic evaluation with NICE reference case 89

Table 11 Previous lines of treatment 90

Table 12 Baseline patient characteristics in the model 91

Table 13 Patient characteristics for the non-responder control (Adapted from Table 114, Appendix L) 93

Table 14 List of comparator therapies by tumour site 96

Table 15 Summary of company's justification for PFS extrapolation curves selection – larotrectinib 105

Table 16 Estimated PFS by parametric form (years from start of treatment) (Appendix L, p399) 106

Table 17 Summary of company's justification for PFS extrapolation curves selection – historical comparator 108

Table 18 Summary of company's justification for PFS extrapolation curves selection – Responder/Non-Responder Analysis 111

Table 19 Estimated PFS by parametric form (years from start of treatment) (Appendix L, p460) 111

Table 20 Summary of company's justification for OS extrapolation curves selection – larotrectinib 112

Table 21 Summary of company's justification for OS extrapolation curves selection – pooled historical comparator 116

Table 22 Summary of company's justification for OS extrapolation curves selection – Responder/Non-Responder Analysis 118

Table 23 Estimated OS by parametric form (years from start of treatment) (Appendix L, p466) 119

Table 24 Comparator adverse event rates by tumour site and weighted 125

Table 25 Health state utility values in the model 128

Table 26 Tumour site specific health state utilities in the model 130

Table 27 Utility decrements applied for each treatment and tumour site in the company's base-case 131

Table 28 Larotrectinib's expected cost 135

Table 29 Larotrectinib acquisition costs based on initially expected list price (adapted from Table 42, CS) 135
Table 30 Comparator treatment costs per cycle 138
Table 31 Health state costs per cycle (adapted from p98, company's response to Points for Clarification, and Appendix M) 140
Table 32 Adverse event costs and rates applied in the company's base-case analysis 142
Table 33 Company base-case deterministic cost-effectiveness results (adapted from Table 53, CS) exclusive of the confidential PAS 143
Table 34 Summary of scenario analysis (Table 52, CS) 147
Table 35 Scenario analysis results for alternative survival models for larotrectinib OS and PFS 148
Table 36 Alternative modelling methods: using non-responding patients as a control (Table 58, CS) 149
Table 37 Alternative modelling methods: using previous line of therapy to adjust control (Table 59, CS) 150
Table 38 Other scenario results in the CS (adapted Table 1, response to clarification, and Tables 60 & 61, CS) 152
Table 39 Company base-case probabilistic cost-effectiveness results (adapted from Table 55, CS) 155
Table 40 ERG adjusted company base-case cost-effectiveness results 162
Table 41 Overview of ERG's additional analyses 165
Table 42 Cost effectiveness results for alternative extrapolation assumptions using response-based survival approach 168
Table 43 Cost effectiveness results for alternative extrapolation assumptions using the company's survival approach 171
Table 44 Cost effectiveness results for post-progression survival equal for larotrectinib and comparator 174
Table 45 Cost effectiveness results for post-progression survival for larotrectinib equal to OS for comparator 174
Table 46 Cost effectiveness results assuming post-progression utility is independent of treatment 175
Table 47 Cost effectiveness results assuming alternative response rates from Bayesian hierarchical model 176
Table 48 ERG base-case cost-effectiveness results 177
Table 49 Cost effectiveness results for IFS scenario (87% ORR) 181
Table 50 Cost effectiveness results for colorectal cancer scenario (43% ORR) 182
Table 51 Cost-effectiveness results including NTRK fusion testing costs 183
Table 52 Assessing end-of-life criteria by tumour site 187
Table 53 Company base-case deterministic cost-effectiveness results with PAS price 205
Table 54 Company base-case probabilistic cost-effectiveness results with PAS price 205
Table 55 NTRK testing scenario – testing strategies per tumour site 209
Table 56 Unit costs of NTRK fusion testing 210
Table 57 Cost per strategy 210

Table of Figures

Figure 1. A rational approach for the detection of NTRK1/2/3 fusion genes in human malignancies Reprinted from C Marchiò et al. 34

Figure 2 PFS evaluation in the N=137 cohort (cut-off 30 July 2018) .. **Error! Bookmark not defined.**

Figure 3 Overall survival for N=137 cohort (cut-off 30 July 2018) ..... **Error! Bookmark not defined.**

Figure 4 Pooled PFS and OS curves for the comparator data from past NICE TAs 60

Figure 5 Kaplan-Meir curve of PFS by responders (purple), non-responders (blue) and comparator data (green) ..... **Error! Bookmark not defined.**

Figure 6 Kaplan-Meir curve of OS by responders (purple), non-responders (blue) and comparator data (green) 62

Figure 7 Waterfall plot of Growth Modulation Index by primary diagnosis (PFS)... **Error! Bookmark not defined.**

Figure 8 Predicted response rate distributions from BHM 66

Figure 9 Predicted mean response rates from BHM 67

Figure 10 Predicted response for an unevaluated tumour site 68

Figure 11 Sampled PFS Kaplan-Meier curve, stratified by tumour site 69

Figure 12 Distributions of plausible median PFS by tumour site 70

Figure 13 Potential improvement in PFS for larotrectinib versus comparator, by tumour site 70

Figure 14 Distributions of plausible median OS by tumour site 71

Figure 15 Potential improvement in OS for larotrectinib versus comparator, by tumour site 73

Figure 16 Model structure (CS, p140) 88

Figure 17 Larotrectinib PFS extrapolated curves (Appendix L, p391) 106

Figure 18 Parametric model fittings: PFS stratified by response status (Appendix L, p459) 112

Figure 19 Larotrectinib OS extrapolated curves (Appendix L, p413) 114

Figure 20 Observed and extrapolated survival curves for larotrectinib with company's preferred assumptions (adapted from company's model) 115

Figure 21 Parametric model fittings: OS stratified by response status (Appendix L, p466) 119

Figure 22 OS Kaplan Meir curves by response status (Figure 111, Appendix L) 121

Figure 23 Observed and extrapolated OS curves for larotrectinib by response status (adapted from company's model) 122

Figure 24 Deterministic sensitivity analysis tornado diagram (Figure 32, CS) 146

Figure 25 Generic functional forms for the response-based survival models (from the model) 206

Figure 26 R regression output for PFS response-based survival models (from the model) 206

Figure 27 R regression output for OS response-based survival models (from the model) 207

List of abbreviations

AE	Adverse events
ALK	Anaplastic lymphoma kinase
BHM	Bayesian hierarchical model
BIA	Budget Impact Assessment
BSC	Best supportive care
CDF	Cancer Drugs Fund
CI	Confidence interval
CNS	Central nervous system
CR	Complete response
CRC	Colorectal cancer
CrI	Credible interval
CS	Company submission
CSR	Clinical study report
DoR	Duration of response
EMA	European Medicines Agency
eMIT	Electronic market information tool
ERG	Evidence review group
ESMO	European Society for Medical Oncology
FISH	Fluorescence in situ hybridisation
HRQoL	Health related quality of Life
HTA	Health Technology Assessment
ICER	Incremental cost-effectiveness ratio
IHC	Immunohistochemistry
IPD	Individual participant data
ITT	Intention to treat
KM	Kaplan Meier
LYG	Life years gained
MASC	Mammary-analogue secretory cancer
NGS	Next generation sequencing
NICE	National Institute for Health and Care Excellence
NNS	Number needed to screen
NSCLC	Non-small cell lung cancer
NTRK	Neurotrophic tyrosine receptor kinase
ORR	Objective response rate

OS	Overall survival
PAS	Patient access scheme
PFS	Progression free survival
PR	Partial response
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QALY	Quality adjusted life-year
RCT	Randomised Controlled Trial
RT-PCR	Real-time polymerase chain reaction
Trk	Tropomyosin receptor kinase
WGS	Whole-genome sequencing

1 Summary

1.1 Critique of the decision problem in the company's submission

The population addressed in the CS matched the NICE scope and covers adult and paediatric patients with solid tumours with a Neurotrophic Tyrosine Receptor Kinase (NTRK) gene fusion, and a disease that is locally advanced, metastatic or where surgical resection is likely to result in severe morbidity with no satisfactory treatment options. The company did not define 'no satisfactory treatment options'.

There is substantial uncertainty around exactly which patients will satisfy these conditions. The exact position in the treatment pathway where larotrectinib will be given in clinical practice is unclear, and will vary by tumour type. The prevalence of NTRK fusion by tumour type is also uncertain, with varying estimates both within and across tumour types. Hence the number of patients eligible for larotrectinib is unclear, but appears to be fewer than [redacted] patients per year.

The genomic and histological testing required to identify patients eligible for larotrectinib is currently not routine for all and is currently only used for some patients with some cancer types. The testing methods used vary by cancer type and between laboratories. Where NGS RNA assays are available, testing for NTRK gene fusions could be added at negligible cost. Any implementation of alternative testing methods or an increase in the volume of testing would be incurred as a further cost to the NHS. The diagnostic accuracy of tests to detect NTRK fusions is uncertain and could have a substantial impact on the efficacy of larotrectinib. Tumour sites with NTRK fusion prevalence below 1% mean that, even with a very high specificity, the number of "false positives" may exceed the number of genuine NTRK fusion cases and people without an NTRK fusion may be offered larotrectinib.

The intervention is larotrectinib, in line with the NICE scope. Broadly, larotrectinib was given to patients at either 100mg or 150mg orally b.d.

The comparator in the NICE final scope is "Established management without larotrectinib". In the absence of a control arm within the clinical trial evidence, the company identified comparators by tumour type through systematic literature reviews, prioritising NICE technology appraisals (TA) sources. The identified comparators included a mix of best supportive care and active comparators. The ERG consider these approaches to be reasonable, given the lack of suitable comparator data within the larotrectinib trials, but still inferior to the use of actual patient data on patients receiving best supportive care.

The outcomes specified in the CS matched the NICE final scope and include overall survival (OS), progression free survival (PFS), and overall response rate (ORR), duration of response (DoR), adverse effects and health-related quality of life.

1.2 Summary of clinical effectiveness evidence submitted by the company

Larotrectinib is described by the company as being 'tumour agnostic' because it targets specific genomic alterations, irrespective of tumour histology. The larotrectinib efficacy and safety analyses were based on three ongoing, single-arm, open-label trials, with the largest trial, NAVIGATE (n=82), using a basket trial design. Using a cut-off date of 30th July 2018, data from the three trials were pooled, which the company presented as an efficacy analysis dataset of 102 patients. This was comprised of 93 patients who had tumours *other* than primary CNS tumours (ePAS2 dataset) and 9 patients with primary CNS tumours (SAS3 dataset). The safety analysis dataset included 137 patients. Fifteen different cancers were represented, with the most common being soft tissue sarcoma (n=20), salivary gland tumour (n=17), infantile fibrosarcoma (n=13), thyroid cancer (n=10) and primary CNS tumours (n=9). Most patients had metastatic disease and an ECOG performance status of 0 or 1. For 23% of the 93 non-CNS tumour patients, larotrectinib was the initial systemic therapy as there was no standard of care.

For ePAS2 patients the overall response rate (the primary outcome) was 72% (95% CI 62 to 81); 15 of the 67 responders had complete responses. For the cohort with primary CNS tumours one out of nine patients (11%) responded. Twenty-eight patients were excluded from the ePAS2 dataset because of insufficient follow up to permit independent review committee assessment of tumours and seven patients were excluded for having a solid tumour with NTRK fusion but no measurable lesion.

In the full N=137 cohort the median duration of OS [redacted] (median follow-up [redacted] months). The 12-month OS rate was [redacted]). The median duration of PFS in the full [redacted] had progressed disease by the 30 July 2018 data cut-off. The median follow-up for PFS was [redacted] months. Of the 93 ePAS2 patients, 34 (37%) had progressed - with mutations identified as a mechanism for resistance in [redacted] patients; the 6-month PFS rate was 77%([redacted]) and the 12-month rate 64% (95% CI 53 to 75).

[redacted] patients had surgery after achieving a partial response: [redacted] post-progression patients received LOXO-195, an experimental therapy manufactured by the company for patients who become resistant to TRK inhibitors.

The company used summary data from previous STAs and published trials as the basis for forming an outcomes dataset for comparator treatments. Comparator data were calculated for all tumour types, and then pooled, weighted by tumour site, with proportions based on the tumour types seen in the larotrectinib trials. Kaplan-Meier data were then digitalised and parametric survival curves were fitted to estimate comparator PFS and OS outcomes over time for each tumour site. The company also undertook two alternative approaches to constructing comparator datasets for scenario analyses - using outcome data from larotrectinib non-responders and outcome data from patients' previous line of therapy.

Most patients (83%) had at least one treatment-emergent adverse event (TEAE) thought to be related to larotrectinib. A conference abstract of the most recent data cut (February 2019) reported that [redacted] of patients experienced a grade 3 or 4 TEAE thought to be related to larotrectinib.

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

Most of the systematic review methods used by the company to identify larotrectinib studies were appropriate and the recruited patients appear to be representative of patients who might be eligible for larotrectinib in the NHS. However, the larotrectinib trials were not designed or sufficiently powered to test the assumption of heterogeneity of response across subgroups. The ERG consider it inappropriate to assume a common response rate independent of tumour histology and find this to be a limitation of the submission. The division between the submission's efficacy datasets ('ePAS2', and 'SAS') appears quite arbitrary and was not clearly justified in the submission. It was also unclear why the safety analysis set (i.e. the modified intention-to-treat dataset) was not used for analysing PFS and OS, although the two datasets yielded very similar results (albeit with immature data).

The company declined to provide the ERG with PFS and OS results for subgroups which the ERG requested and considered might be important for investigating possible treatment effect heterogeneity. The ERG disagrees with the company's view that there is no evidence of effect heterogeneity. From the data which *were* available this heterogeneity was most clearly evident when comparing the ORR for patients with primary CNS tumours with the ORR for the ePAS2 cohort. Moreover, the observation that [redacted] had surgery following a partial response and that these patients had either [redacted] also suggests that heterogeneity in PFS and OS across tumours sites is likely.

[redacted] post-progression patients received an experimental therapy called 'LOXO-195' which was developed for treating patients who become resistant to TRK inhibitors, and which is produced by the manufacturer of larotrectinib. Whilst the ERG acknowledges that some resistant patients would receive experimental treatments in the NHS, the number of patients specifically receiving LOXO-195

(as opposed to other experimental treatments) suggests that the impact of LOXO-195 on survival after progression following treatment with larotrectinib should not be ignored (as was suggested by the company). Furthermore, [redacted] patients continued to receive larotrectinib post-progression, which would not happen in clinical practice.

Although the company undertook a comprehensive systematic review of comparator therapies, the submission did not present how the comparator baseline characteristics or outcome data for individual tumour sites compared with the corresponding data for larotrectinib, so it was not possible to evaluate the appropriateness of the comparator data sets selected (in the systematic review). However, given both the historical and broad nature of the comparator datasets (many patients will not have had NTRK gene fusions), it is likely that no or very few patients who progressed would have received targeted experimental therapy – such as LOXO-195 – a bias which would favour larotrectinib.

The ERG considers the validity of pooling across multiple tumour sites to be highly uncertain since they may have varying expected survival times, so combining them in an average survival curve may not be meaningful. Also, patients in comparator treatment trials may not have had NTRK fusion, and so might have a different prognosis to patients eligible for larotrectinib. The ERG therefore considers that an analysis accounting for potential heterogeneity across tumour sites would have been more appropriate. The ERG agrees with the key limitations noted by the company about the two alternative approaches to constructing comparator datasets. However, although none of the comparator datasets are ideal, their results are broadly consistent though limited by not accounting for tumour-site heterogeneity.

1.4 Summary of cost effectiveness submitted evidence by the company

The company's economic submission included systematic reviews of published evidence on 1) the cost-effectiveness analyses of treatments for patients with NTRK-fusion cancer and 2) the cost-effectiveness evidence of treatments for patients with solid tumours that are known to harbour NTRK gene fusions. The company did not identify any published cost-effectiveness studies on the treatment of NTRK fusion-positive cancer. For the second search the company present the results by tumour site. The CS states that 98 studies were identified across all tumour sites. These studies were used in the development of the model structure, and assumptions used in model development.

The CS presented a *de novo* cohort cost-effectiveness model to estimate the cost-effectiveness of larotrectinib compared with established practice in a population of adult and paediatric patients with *NTRK* fusion-positive solid tumours. In the base case analysis, established practice consisted of a composite comparator represented by a weighted average of comparators from the tumour types represented in the integrated efficacy analysis for larotrectinib, this is referred to as the historical

comparator in this ERG report. Cost-effectiveness was assessed over a lifetime time horizon (40 years for model engines considering adult patients, 80 years for engines considering pooled populations including both adult and paediatric patients), with a 3.5% discount rate applied to both costs and quality adjusted life years (QALYs).

The model structure is based on a partitioned survival model (PSM) or "area under the curve" analysis comprising of three mutually exclusive health states: (i) PFS (progression free), (ii) progressive disease (PD; progression), and (iii) death. The base case assumed treat-to-progression for the larotrectinib arm and followed the treatment duration for comparator treatments in the source documents. A scenario was available to apply the time to treatment discontinuation curve from the larotrectinib clinical trial programme. . The model predicted the total costs and QALYs separately for the larotrectinib arm and the historical comparator arm. The distribution of patients in each health state was determined by using estimates of PFS and OS. Two alternative comparator scenarios were also explored using larotrectinib trial data to generate a control group: 1) using non-responders (response-based model), 2) using within-study previous line of treatment.

For larotrectinib, the distribution of patients in health states were determined by extrapolating KM data from the integrated efficacy analysis. Estimates of time in PFS and PD were then used to estimate total costs and QALYs for larotrectinib. For the historical comparator, for each tumour histology, the comparator engine is informed by extrapolated OS and PFS curves derived from published literature specific to that treatment, including NICE TAs. The OS and PFS data were extrapolated according to the preferred approach in the published analysis. Total costs and QALYs are generated for 12 different tumour types. Weighted total costs and QALYs for the historical comparator were then estimated using the distribution of tumours in the integrated efficacy analysis of larotrectinib.

The OS and PFS extrapolations for larotrectinib were based on the integrated efficacy analysis which pooled data from three trials: LOXO-TRK-14001, LOXO-TRK-15002 (NAVIGATE) and LOXO-TRK-15003 (SCOUT). The integrated efficacy analysis set included 102 patients across 12 different tumour types. The data-cut off used in the economic model was the 30th July 2018. To extrapolate the observed OS and PFS data, the company fitted a number of standard parametric models. The models selected for the company's base-case analysis were extrapolated [redacted] OS and PFS survival functions. The model does not consider stratification by tumour site or any other subgroup (e.g. children vs adults) of larotrectinib survival outcomes. This is justified in the CS on the basis of the small number of events (overall and by tumour site).

The estimates used in the company's base-case analysis for health-related quality of life of patients in the PFS and progressive disease health states for larotrectinib were derived from EQ-5D and PedsQL data collected in the larotrectinib clinical trial programme. A different set of health state utilities was applied by tumour site for the comparator, and was informed by previous NICE TAs and targeted literature searches. Utility decrements for adverse events were included and were tumour site specific. The comparator adverse event disutilities were informed by previous NICE TAs and targeted literature searches. Utility decrements for adverse reactions for larotrectinib were assumed to be the maximum disutility for the event across all tumour sites.

Resource use and costs included: drug acquisition and administration costs, costs related to health states and adverse events. Comparators administered orally, and larotrectinib, were assumed to have no administration costs. The duration of larotrectinib treatment was assumed to be until progression (and varied on scenario analysis to reflect treatment duration as in the clinical studies). Health care resource use and costs for the comparator was modelled separately by tumour site and sourced from previous NICE Technology Appraisals (TAs), when available, or from other published sources. Testing costs, to determine NTRK fusion, were not included.

The company found larotrectinib to be more costly (cost difference of [redacted]) and more effective ([redacted] QALYs gain) compared with established management. The pricing structure was subsequently updated for larotrectinib and the company submitted an amended base case. This resulted in a lower incremental cost difference ([redacted]). The deterministic base case incremental cost-effectiveness ratio (ICER), using the updated pricing structure, was [redacted] per QALY. Probabilistic results were [redacted] and [redacted] per QALY gained for the original and updated pricing structure, respectively. The company reported that the most influential parameter in the one-way sensitivity analysis was the parametric distribution used to extrapolate larotrectinib OS estimates.

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

The ERG validated the original company model and found a number of small errors and inconsistencies with the CS. These are unlikely to be key drivers of cost-effectiveness.

The ERG highlights that there are significant number of issues that contributed to uncertainty in the cost-effectiveness results presented by the company. These focus on the choice of historical control in the company base case, the extrapolation of larotrectinib survival outcomes, potential confounding of subsequent lines of therapy, heterogeneity in treatment effect and the lack of NTRK fusion testing costs included.

The analysis of PFS and OS for larotrectinib, given small numbers and data immaturity, means that estimates are subject to considerable uncertainty and are likely to be sensitive to assumptions made to

extrapolate the integrated efficacy analysis results over the model time horizon. In addition, the absence of a control group for larotrectinib makes estimation of the treatment effect difficult. The company sought to explore alternative scenarios to generate a comparator arm, however the ERG's general view is that all approaches have limitations and may result in biased estimates of treatment effectiveness. The choice of the historical comparator for the base case analysis is likely to be subject to confounding bias, due to the unknown NTRK status of patients in the comparator studies. In generating estimates of OS, the ERG also noted that the post progression survival may be affected by treatments available in the larotrectinib arm that are not available in routine practice, i.e. are not available to the historical comparator. The response-based model is less affected by confounding from subsequent lines of treatment and imbalance of patients characteristics. The previous line of treatment control is also a valid approach to reduce the confounding of treatment effect, however it was not appropriately applied in the company submission.

The company present a single ICER across all tumour types. The ERG's view is that this potentially conceals significant variation in the tumour specific ICERs, driven by a combination of factors, particularly variability in relative effectiveness between tumour types, as well as across other clinical characteristics such as age (paediatric vs adults). The implications of this heterogeneity for cost-effectiveness results is unknown. The ERG conduct exploratory analyses on response data, which suggests that there is evidence that the treatment effect is heterogeneous across tumour types. The company do not explore the issue of heterogeneity in response and/or survival times, nor do they consider heterogeneity in ORR between the ePAS2 and the full population or by individual study in the integrated efficacy analysis. The ERG consider the issue of heterogeneity in treatment effect to be a fundamental issue in determining the cost-effectiveness of larotrectinib and has concerns regarding the validity of the ICERs generated by the company, given the lack of analysis to reflect these sources of heterogeneity.

The ERG also has substantive concerns regarding the lack of NTRK fusion testing costs assumed in the company submission. NTRK gene fusion testing is currently not performed routinely in the UK for all tumour sites, and it is unclear how adding tests costs for these populations will impact on the ICER for larotrectinib. The cost of testing is also an important sources of heterogeneity, as different tumour types require different numbers of patients to test, to identify NTRK fusions, given differences in prevalence rates.

1.6 ERG commentary on the robustness of evidence submitted by the company

1.6.1 Strengths

The decision problem considered largely matched the NICE scope and was appropriate. The three trials considered were well-conducted single arm trials of larotrectinib, covering the broad range of

tumour sites where NTRK fusion might occur, including both adult and paediatric cancer cases. The trials combined had sufficient power to detect a meaningful overall response rate of 30% or higher.

The company explored multiple scenarios to generate a comparator for larotrectinib. These three methods gave broadly the same conclusions. In the base case analysis the use of the historical control was preferred. The company conducted systematic reviews to identify survival outcomes for the 15 tumour types included. Efforts were made by the company to replicate the preferred approach to survival analysis in the relevant publication informing each tumour type model engine.

1.7 Weaknesses and areas of uncertainty

1.7.1 Trials of larotrectinib

The primary analysis included only 93 patients, which is a small sample on which to base any assessment, and samples are even smaller once the number within any tumour site included is considered; with a maximum of [redacted] patients in any tumour site group, or 17 when considering adults and children separately.

The analysis in the CS made the assumption that larotrectinib was equally effective in all tumour types, and all patients were analysed together without differentiating between trials, patient ages or tumour sites. The ERG considers this to be inappropriate, as heterogeneity is at least plausible, and basket trials are designed to investigate heterogeneity across tumour sites. The ERG's analysis accounting for heterogeneity found that larotrectinib was less effective than estimated in the CS, with clear evidence of heterogeneity across tumour sites. Therefore the ERG considers the analysis provided in the CS to be unsound.

1.7.2 Indirect comparisons

As all three trials had single arm designs, no direct comparison of larotrectinib with other interventions was possible. The three approaches for indirect comparison used in the CS all have substantial limitations.

Survival curves were estimated for each tumour site using data from past NICE TAs or, in the absence of suitable data, other literature sources or by assuming equivalent survival across similar tumour sites. These TAs varied in what intervention was used; some used best supportive care, other an active intervention; there were some tumour sites with no TA, so assumptions were made by equating survival across similar tumour sites. Because of these assumptions, the somewhat arbitrary selection of previous TAs, the variation in interventions considered, and the likelihood that people in previous TAs did not have NTRK fusions, the ERG considers this approach to have substantial limitations, and may not truly represent the survival expectations of people with NTRK fusion who do not receive

larotrectinib. The response-based model, however, allows for further exploration of assumptions regarding treatment effect homogeneity.

Patients who did not respond to larotrectinib were considered as a proxy for patients not receiving larotrectinib. Such non-responders appeared to have a worse survival profile than patients who never receive larotrectinib. For overall survival results may be biased because many non-responder continued to receive larotrectinib or were recruited into the trial of LOXO-195.

The third approach compared time to progression on previous lines of therapy to progression-free survival on larotrectinib. The ERG considers this to be of limited vale as it cannot inform overall survival, data were not available for all patients, and there was considerable heterogeneity across patients and tumour sites. This approach was not implemented appropriately in the cost-effectiveness model.

The ERG considers that the company did not fully explore the implications of the alternative approaches to model the cost-effectiveness results.

1.7.3 Identifying NTRK fusions

The ERG notes that there is considerable uncertainty in how many people will be eligible for larotrectinib. Estimates of the prevalence of NTRK fusion are highly uncertain, with varying estimates from different sources, heterogeneity across tumour sites, and uncertainty over exactly which types of tumour might harbour NTRK fusions. There is also uncertainty as to where in the clinical pathway larotrectinib will be used. The approval is for whenever no "suitable" alternative exists; hence use of larotrectinib may vary between tumour sites, where the range of available treatments will vary.

The CS did not consider screening to identify NTRK fusions. The ERG considers this to be a critical omission, because successful screening for NTRK fusion is essential to identify patients eligible for larotrectinib. There currently appears to be little consensus on how genetic/genomic screening should be used. In some tumour sites it may already be widely used (but may require extra panels to identify NTRK fusion specifically); in others it may not be used at all. The ERG found that the numbers needed to screen to identify one NTRK fusion may be very high in tumour sites with low NTRK fusion prevalence, raising doubt as to the practicality of screening in such tumour sites.

There also appears to be uncertainty as to the accuracy of screening to identify NTRK fusions. The ERG notes that, for tumour sites with low NTRK fusion prevalence, even with a near perfect test (e.g. of 99% accuracy) the number of "false positives" (people who test positive for NTRK fusion but do not have it) may outnumber the true NTRK fusion cases. This will substantially reduce the observed

effectiveness of larotrectinib, and casts doubt on whether larotrectinib can be used ethically in tumour sites with low NTRK fusion prevalence.

The provision of larotrectinib on the NHS is likely to substantially increase the number of patients requiring molecular testing. The cost of testing is also an important sources of heterogeneity, as different tumour types require different numbers of patients to test, to identify NTRK fusions, given differences in prevalence rates. The ICER may therefore vary widely by including testing costs weighted according to the tumour types observed in the integrated efficacy analysis.

1.7.4 Uncertainty in the extrapolation of larotrectinib survival outcomes

The choice of parametric model to extrapolate from observed data for larotrectinib, is a key driver of the survival gains. The ERG highlights that the observed data for larotrectinib is immature and there are small numbers of events. As such, there is significant uncertainty regarding the longer-term survival benefits of larotrectinib. The ERG considers that, by only exploring alternative parametric distribution for the historical comparator analysis and not the response based model, the company do not fully explore the implications of the large variation in survival gains for larotrectinib and the ICER according to different parametric distributions.

1.7.5 Potential confounding of subsequent lines of therapy

The ERG have concerns that post-progression gains in survival for larotrectinib, may be driven, at least in part, by treatments only available to patients in the larotrectinib comparator: continued larotrectinib post progression and LOXO-195. The bias is likely to be in favour of larotrectinib. The company do not explore the implications of this bias for the cost-effectiveness results.

1.7.6 Uncertainty surrounding the homogeneity of the treatment effect

The ERG consider that there is potential for heterogeneity in the treatment effect across tumour types, as well as across other clinical characteristics such as age (paediatric vs adults) and fusion type. The company do not explore the issue of heterogeneity in response and/or survival times, nor do they consider heterogeneity in ORR for the populations included in the integrated efficacy analysis.

The ERG consider the issue of heterogeneity in treatment effect to be a fundamental issue in determining the cost-effectiveness of larotrectinib and has concerns regarding the validity of the ICERs generated by the company, given the lack of analysis to reflect these sources of heterogeneity.

1.8 Summary of exploratory and sensitivity analyses undertaken by the ERG

1.8.1 Re-analysis of included trials

The ERG disagreed with the decision in the CS to analyse the trials without accounting for possible heterogeneity across patients and tumour sites. The ERG requested response and survival data by trial and tumour site in order to investigate heterogeneity, but the company declined to provide this. To investigate and account for heterogeneity the ERG fitted a Bayesian Hierarchical model (BHM) to the overall response rate (ORR) data, which was reported by tumour site in the CS.

This model found that the summary overall response rate was 64% (95% CrI 29 to 83), lower than that reported in the CS. If primary CNS tumours were included the estimated response rate became 57% (95% CrI 23 to 80). The model identified evidence of heterogeneity in response across tumour sites. Some tumour sites had near-certainty that the ORR exceeded 30% (including IFS, MASC, GIST, thyroid, lung), but in other sites there was some chance that the ORR was under 30% (including appendix, breast, colon, pancreas). Further regression analysis found that the predicted ORR was higher in tumour sites where NTRK fusion was common, and may be higher in paediatric cases.

The ERG performed a speculative analysis to investigate possible heterogeneity in progression-free and overall survival. This assumed that all patients with a response to larotrectinib (and all with no response) had the same survival distribution, regardless of tumour site. The analysis found that substantial heterogeneity in survival times across tumour types was plausible.

1.8.2 Adjustments to the cost-effectiveness model

The key uncertainties addressed by the ERG scenario analyses relate to:

- Parametric distribution fitted to larotrectinib PFS and OS data, using the response based model and the historical comparator.
- Assumptions regarding the gains in post-progression survival with larotrectinib
- The use of a Bayesian Hierarchical model to estimate overall response rate, using the ePAS2 and the full integrated efficacy analysis populations.

The ERG also adjusts the paediatric does for larotrectinib to accord with the average BSA in the larotrectinib integrated efficacy analysis. This generates an ERG adjusted company base case ICER of [redacted] per additional QALY.

The results of these scenario analyses including the ERG's base-case are summarised in Table 1. Deterministic ICERs are presented throughout. The ERG conclude that scenario 4 may provide more

clinically plausible projections of post-progression survival for larotrectinib. Therefore, the ERG builds upon this scenario from scenario 9 onwards. All of the scenarios using the company's preferred survival approach (5 to 8) generate long and clinically implausible projections of post-progression survival for larotrectinib. Under the ERG's alternative set of assumptions, the ICER for larotrectinib versus established care is [redacted] per QALY for the ePAS2 and the full integrated efficacy analysis population, respectively.

Table 1 Summary of ERG exploratory analyses

	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG adjusted company base case	[redacted]	[redacted]	[redacted]
Scenario 1: Response based model, Weibull for OS and PFS	[redacted]	[redacted]	[redacted]
Scenario 2: Response based model, Exponential for OS and PFS	[redacted]	[redacted]	[redacted]
Scenario 3: Response based model, Gompertz for OS and PFS	[redacted]	[redacted]	[redacted]
Scenario 4: Response based model, Gompertz for PFS and Weibull for OS	[redacted]	[redacted]	[redacted]
Scenario 5: Historical comparator, Weibull for OS and PFS	[redacted]	[redacted]	[redacted]
Scenario 6: Historical comparator, Exponential for OS and PFS	[redacted]	[redacted]	[redacted]
Scenario 7: Historical comparator, Gompertz for OS and PFS	[redacted]	[redacted]	[redacted]
Scenario 8: Historical comparator, Gompertz for PFS and Weibull for OS	[redacted]	[redacted]	[redacted]
Scenario 9: Post-progression survival equal for larotrectinib and comparator	[redacted]	[redacted]	[redacted]
Scenario 10: Post-progression survival for larotrectinib equal to OS for comparator	[redacted]	[redacted]	[redacted]
Scenario 11.1: Post-progression utility is independent of treatment (same progression utility)	[redacted]	[redacted]	[redacted]
Scenario 11.2: Post-progression utility is independent of treatment (same post progression utility and same post-progression survival as comparator OS)	[redacted]	[redacted]	[redacted]
ERG alternative base-case analysis:			
Scenario 4 + 64% ORR (ePAS2 population)	[redacted]	[redacted]	[redacted]
Scenario 4 +57% ORR (full population)	[redacted]	[redacted]	[redacted]

In further exploratory analysis the ERG explore the potential impact of heterogeneity in response rate, using two extreme response tumour types, with ORR estimated from the BHM. Two tumour types are chosen for scenarios: IFS which has a very high response rate (87%) compared to the overall, and colorectal cancer which has a very low response rate (43%) compared to the overall rate. For these

two tumour types, tumour specific ICERs are estimated. The ICER for IFS is [redacted] per additional QALY and the ICER for colorectal cancer is [redacted] per additional QALY. These scenarios are likely to underestimate the impact of tumour type on the ICERs, as a common distribution of PFS and OS conditional on response is applied.

The ERG finally explores a scenario including testing costs and examines its likely impact on cost-effectiveness. The weighted overall cost of testing for larotrectinib applied in the model is £18,670. In this scenario, the cost of testing was added as a one-off cost to the total costs of larotrectinib. The ICER for larotrectinib, including this testing cost is [redacted] per QALY gained when assuming a 64% ORR (ePAS2 population estimated from the BHM), and [redacted] per QALY gained, when assuming a 57% ORR (full integrated efficacy analysis population estimated from the BHM).

2 Background

2.1 Critique of company's description of underlying health problem.

This appraisal concerns the treatment of solid tumours with a Neurotropic Tyrosine Receptor Kinase (NTRK) gene fusion that is locally advanced, metastatic or where surgical resection is likely to result in severe morbidity and no satisfactory treatment option is available. The use of larotrectinib will be based solely on the presence NTRK fusions as a primary oncogenic driver, across a wide variety of solid tumour sites in both children and adults. As such, the company focus on NTRK fusion-positive cancer as the disease entity, independent of tumour site.

The ERG considers the company's description of the underlying health problem to be appropriate and relevant to the decision problem under consideration. The company describe the role of tropomyosin receptor kinase (TRK) proteins, encoded by the NTRK gene family (NTRK1, NTRK2 and NTRK3), in regulating the proliferation, growth and survival of neurons in the central and peripheral nervous system. NTRK-fusion positive tumours arise from a gene rearrangement involving fusion of a portion of the NTRK1, NTRK2 or NTRK3 gene with another unrelated gene (a fusion partner). The company describe how these fusions promote cancer formation through the activation of the TRK fusion protein that drives unchecked cell proliferation and tumour growth though the TRK pathway[1].

The CS states the results of studies regarding the prognosis of patients with an NTRK fusion on page 21. The company provides examples of studies that suggest an association between the presence of an NTRK fusion and unfavourable disease presentation [2, 3] and better prognosis in patients with congenital mesoblastic nephroma (CMN), who harbour an NTRK fusion compared to those without the genetic abnormality[4]. The ERG note that evidence concerning the prognosis of patients with an NTRK fusion is generally weak and supported by data from fewer patients. However, there is a suggestion that the prognosis of patients with NTRK fusions varies between cancer types and that variation may also exist between NTRK fusion types. From the evidence available, it is unclear whether NTRK fusions are in themselves prognostic or whether it is their association with specific prognostic factors (such as ECOG status or age) that drives the different prognosis.

2.1.1 Eligible population and prevalence of NTRK

In the budget impact assessment (BIA) submitted by the company, the eligible patient population was estimated to be [redacted] in 2020 and was based on the sum of the number of patients receiving last line of cancer therapy for various tumour sites harbouring NTRK fusion. The company note that not all of these patients would harbour NTRK fusion or be appropriate for larotrectinib therapy.

The CS states that less than 1% of solid tumours harbour NTRK fusions. Estimates range from <0.1% to 3% in common histologies, such as non-small cell lung cancer (NSCLC) and colorectal cancer

(CRC), to over 90% in several uncommon tumours, such as secretory breast carcinoma and infantile fibrosarcoma (IFS).

Using data from the BIA and from additional investigations for estimates of NTRK fusion prevalence, the ERG has derived estimates by cancer type of the number of patients who have the NTRK fusion, the number of patients who are eligible for testing, the number of patients who are eligible for testing and have the NTRK fusion and would therefore be treated with larotrectinib in clinical practice and the number needed to screen to identify one patient with an NTRK fusion cancer.

Additional investigations by the ERG identified the percentage of those with the NTRK fusion from the larotrectinib NDA Multidisciplinary review and evaluation document submitted to the FDA.[5] The case per year were derived from the BIA, except those for MASC and CNS paediatric cancers which were not provided in the BIA and derived from the literature by the ERG.[6 7] As the true incidence of MASC is unknown, an estimate was derived from the incidence of salivary gland cancer by taking the proportion of those known to be MASC.[8]

The percentages eligible for larotrectinib were derived from the BIA, except those for MASC, STS infantile sarcoma, secretory breast cancer and CNS paediatric that were not provided in the BIA. For the percentage eligible for larotrectinib within these cancer types, the ERG identified the percentage of patients at stage three or four for each cancer type,[6, 7, 9] and assumed based on clinical advice that 30% of these patients would be eligible for further line treatments and therefore eligible for larotrectinib.

The number of patients who have the NTRK fusion was derived from multiplying the number of cases per year by the percentage of those with the NTRK fusion. The number eligible for testing was derived from multiplying the number of cases per year by the percentage eligible for larotrectinib (proportion at relevant line of therapy). The number who are eligible for testing and have the NTRK fusion was derived from multiplying the number eligible for testing by the percentage of those with the NTRK fusion, whilst the number needed to screen was derived from dividing the number eligible for testing by the number eligible for testing and has an NTRK fusion.

Estimates are presented for those cancers presented in the BIA (Table 2), with estimated data for cancers included in the CS but not in the BIA. The ERG note that further tumour sites unrepresented in the CS could contribute to the number of patients receiving larotrectinib in clinical practice, however these numbers are expected to be very low.

There is uncertainty around the proportion at relevant lines of therapy, as it is currently unknown for given tumour sites, the exact position in the treatment pathway larotrectinib will be given in clinical

practice; see section 2.2.1 for further details. Based on estimates presented in Table 2, the number of patients who receive larotrectinib in clinical practice may be substantially lower than the number of patients receiving last line of cancer therapy for various tumour sites harbouring NTRK fusion.

Table 2. Number of patients who have a NTRK fusion, who are eligible for testing and treatment with larotrectinib and the number needed to screen (NNS), by cancer type.

Cancer type	Cases per year	% eligible for larotrectinib	% of tumours with NTRK fusion	Has NTRK fusion	Eligible for testing	Eligible and has NTRK fusion	Number Needed to Screen
NSCLC cancer patients	[redacted]	[redacted]	0.09%	[redacted]	[redacted]	[redacted]	[redacted]
Salivary cancer patients (non-MASC)	[redacted]	[redacted]	1.72%	[redacted]	[redacted]	[redacted]	[redacted]
MASC cancer patients	[redacted]	[redacted]	100.00%	[redacted]	[redacted]	[redacted]	[redacted]
Melanoma cancer patients	[redacted]	[redacted]	0.21%	[redacted]	[redacted]	[redacted]	[redacted]
Colorectal cancer patients	[redacted]	[redacted]	0.12%	[redacted]	[redacted]	[redacted]	[redacted]
Appendix cancer patients (assumed same as pancreatic)	[redacted]	[redacted]	4.00%	[redacted]	[redacted]	[redacted]	[redacted]
STS adults (GIST) cancer patients	[redacted]	[redacted]	1.28%	[redacted]	[redacted]	[redacted]	[redacted]
STS adults (non-GIST) cancer patients	[redacted]	[redacted]	0.56%	[redacted]	[redacted]	[redacted]	[redacted]
Bone Sarcoma cancer patients (assumed same as STS adults)	[redacted]	[redacted]	1.00%	[redacted]	[redacted]	[redacted]	[redacted]
STS paediatrics cancer patients	[redacted]	[redacted]	0.56%	[redacted]	[redacted]	[redacted]	[redacted]
STS infantile sarcoma cancer patients	[redacted]	[redacted]	90.90%	[redacted]	[redacted]	[redacted]	[redacted]
Secretory breast cancer patients	[redacted]	[redacted]	91.70%	[redacted]	[redacted]	[redacted]	[redacted]
Cholangiocarcinoma cancer patients	[redacted]	[redacted]	0.10%	[redacted]	[redacted]	[redacted]	[redacted]
CNS cancer patients (assumed same as brain)	[redacted]	[redacted]	0.05%	[redacted]	[redacted]	[redacted]	[redacted]
CNS paediatric cancer patients	[redacted]	[redacted]	5.30%	[redacted]	[redacted]	[redacted]	[redacted]
Pancreas cancer patients	[redacted]	[redacted]	0.26%	[redacted]	[redacted]	[redacted]	[redacted]
Thyroid cancer patients	[redacted]	[redacted]	3.96%	[redacted]	[redacted]	[redacted]	[redacted]
				[redacted]	[redacted]	[redacted]	[redacted]

.1 Grey cells indicate where data was not reported in the BIA, and has been identified by the ERG. * The overall number needed to screen is derived from dividing the total number eligible for testing by the total number eligible for testing and has an NTRK fusion.

2.2 Critique of company's overview of current service provision

2.2.1 Treatment pathways

The company state that there are no approved treatment options in the UK specifically for patients with NTRK fusion-positive solid tumours and, to date, treatment recommendations regarding NTRK fusion-positive cancer have not been included within any UK guidelines. Patients are currently treated per treatment guideline recommendations for the specific tumour site, irrespective of NTRK status. The company describe that treatment recommendations vary by tumour site, with the more common tumours having guideline recommendations for multiple lines of therapy and less frequent tumours sites having limited or no treatment guidelines.

The position at which NTRK fusion-positive cancer patients would be offered larotrectinib is, therefore, likely to vary by tumour site and be dependent on the availability of effective treatments in each tumour. This is reflected in the marketing authorisation, which covers Vitrakvi (larotrectinib) for the "treatment of adult and paediatric patients with solid tumours that display a Neurotrophic Tyrosine Receptor Kinase (NTRK) gene fusion. Treatment with Vitrakvi is recommended for patients whose disease has spread or cannot be surgically removed, and who have no other satisfactory treatment options." In the clarification response the company stated that the term [redacted] Information regarding which therapies are likely to fall into the category of 'unsatisfactory' was not provided.

Clinical advice to the ERG agreed that the position at which larotrectinib would be offered to NTRK fusion- positive cancer patients would vary by tumour site and noted that for tumours where there is no recommended standard of care or standard of care is poor, larotrectinib may be used as a first line therapy over standard chemotherapy. Furthermore, the threshold at which alternative treatments are deemed unsatisfactory would involve assessment of response rates and adverse events burden and a discussion of alternatives with patients. Therefore, it is likely this threshold would vary by both clinicians and patients.

2.2.2 NTRK fusion diagnostic pathways

The company explain that multiple testing methods are available to identify patients with tumours harbouring NTRK fusions. These are next generation sequencing (NGS), fluorescence in situ hybridisation (FISH), immunohistochemistry (IHC), reverse transcription polymerase chain reaction (RT-PCR), and whole-genome sequencing (WGS). The company describe that WGS will be routinely available to cancer patients as part of the national NHS Genomic Medicine Service and

reference the NHS Long Term Plan[10] section 3.63 which states that "the NHS will routinely offer genomic testing to all people with cancer for whom it would be of clinical benefit…".

Clinical advice to the ERG estimated a minimum of five to ten years before cancer diagnostic services expand so that all patients with cancer for whom it would benefit, would receive WGS. The ERG consequently notes that it cannot be assumed that genetic screening is currently available to all patients potentially eligible for larotrectinib.

2.2.2.1 Current provision for NTRK fusion screening

Considering current practice, clinical advice to the ERG suggested that next generation sequencing (NGS), which permits the reading of single genes or panels of genes, is mainly conducted on the NHS for prognostic purposes and that patient eligibility for NGS is criteria driven. The identification of those who are eligible to receive larotrectinib would require the implementation of routine testing for the NTRK fusion, for diagnostic purposes.

Currently, testing for NTRK fusions for diagnostic purposes is more common in paediatric patients, with high NTRK fusion prevalence cancers. However, clinical advice to the ERG noted that the current WGS used on the NHS may not detect NTRK fusions as it may lack the specific testing panel.

For some cancers that harbour NTRK fusions, genomic and histological testing methods are available on the NHS and are detailed in Table 3. Clinical advice informed the ERG that, for cancers where testing methods are available, they are often not available for all; those who do receive them, do so based on specific criteria.

Table 3. Testing methods currently available of the NHS for those cancers that harbour NTRK fusions

Tumour Type *	IHC	Current Molecular Testing[11]
Salivary gland (MASC)	-	FISH (ETV6-NTRK3)
NSCLC Lung (Adenocarcinoma & squamous cell carcinoma)	IHC 22C3 (doesn't seem that common) [12]	Multi-target NGS panel (EGFR)
Breast cancer (not specified) Secretory breast carcinoma	IHC (HER2) [13]	Oncotype DX: multi-target
Papillary thyroid tumour Thyroid tumour (NOS)	IHC [14]	Multi-target NGS Panel (BRAF, KRAS, NRAS, HRAS) [15] Multi-target NGS Panel (KRAS, NRAS, HRAS, RET)
Colon/colorectal	IHC for Lynch Syndrome (hereditary CRC) [16]	Multi-target NGS Panel (BRAF, KRAS, NRAS)

Melanoma (NOS) Spitzoid Melanoma	-	Multi-target NGS Panel (BRAF, NRAS, KIT)
Neuroendocrine (NOS)	-	-
Gastrointestinal stromal tumour	IHC (CD117, C134, DOG1) [17]	Multi-target NGS Panel (KIT, PDGFRA)
Cholangiocarcinoma	ICH for intrahepatic cholangiocarcinoma[18]	-
Pancreatic	-	-
Appendix	-	-
Uterine Ovarian Cervix	IHC (EMA, Ber-EP4, PAX8, CK7) [19] IHC [20] IHC [21]	FISH (EPC1-PHF1) Multi-target NGS panel (BRAC1, BRAC2) Multi-target NGS panel (SMARCA4)
Head and neck squamous cell carcinoma (NOS) Salivary gland (non MASC) Sinonasal adenocarcinoma	IHC (HPV) [22]	Multi-target NGS Panel – (CDKN2A, EGFR, TP53)
Gastro-esophageal junction	Noto recommended in routine practice [23]	-
Prostate cancer	IHC (PSA) [24]	-
Renal cell carcinoma	-	FISH/RT-PCR (TFE3)
Low-grade glioma High grade glioma (inc. glioblastoma multiforme)	 IHC [25]	Multi-target NGS (BRAF), MGMT promotor hypermethylation Multi-target NGS (IDH1, IDH2, ATRX, TERT, H3F3A)

* Cancer types known to harbour NTRK fusions [5 26 1 27 28] - Dashed line indicates where no testing methods currently available of the NHS for that cancer type

For tumours where there is no genomic or histological testing currently in place, new infrastructure may be required to accommodate the testing of these patients. This would be incurred as a cost to the NHS and could include the cost of consumables, the processing of samples, data processing and analysis, interpretation and reporting of results, along with data storage [29].

For cancers where genomic and histological testing is already conducted, additional testing for NTRK could, at least in part, be accommodated within the current infrastructure. For example, clinical advice informed the ERG that in laboratories currently using NGS RNA fusion testing, a new gene could be added to an existing fusion panel to test for NTRK fusions at negligible additional cost. However, in

laboratories that use NGS DNA testing methods, this testing would need to be replaced with NGS RNA assay, making implementing testing for NTRK fusions more difficult. Any implementation of alternative testing methods or an increase in the volume of testing would be incurred as a further cost to the NHS.

The CS has not considered the impact and implications of any additional testing costs, stating that "Testing costs are not included within the model as patients will be tested routinely according to NHS plans". There has been no acknowledgement that screening for NTRK to identify the patients who will be treated with larotrectinib, will be associated with additional cost and that introducing large numbers of tests to NHS pathology services will require practical and infrastructural considerations. The ERG has considered cost implications of additional testing in section 6.5.4.

2.2.2.2 Testing strategies for detection of NTRK fusions

The additional cost incurred by the NHS will depend on the testing approach adopted for the detection of *NTRK*1, 2 and 3, of which several have been suggested in the literature [30]:

- FISH or IHC for detecting NTRK fusions in high NTRK prevalence cancers
- IHC following by confirmatory NGS for detecting NTRK fusions in low to medium NTRK prevalence cancers
- NGS for detecting NTRK fusions in low NTRK prevalence cancers

Each are associated with advantages and disadvantages and assay choice would need to take into account the resources and clinical context. FISH requires a separate assay for each NTRK fusion, which would become expensive and time consuming, but is highly effective when detecting the ETV6-NTRK3 fusion in high prevalence cancers. IHC has a quick turnaround time of approximately one day and is relatively inexpensive. However, it can have low sensitivity, particularly for NTRK3 fusions, approximately 50-70%. NGS is highly accurate with high sensitivity and specificity. However the turnaround time is approximately 2-4 weeks and compared to other testing methods, it is highly expensive.

In section 2.1.1, the ERG estimated the number of patients who require screening to identify one individual with an NTRK fusion or the number needed to screen (NNS). These estimates vary by tumour site, depend on the prevalence of gene rearrangement and will impact the feasibility of implementing screening for NTRK on the NHS. Recommendations by the European Society of Medical oncology (ESMO) note that there are efficacy and cost associated challenges with screening, particularly for low NTRK fusion prevalence tumours [31].

Recently, the ESMO Translational Research and Precision Medicine Working Group conducted a review of the literature on the available methods for the detection of NTRK fusions and made recommendations for the implementation of a rational approach for the detection of NTRK1/2/3 fusion genes in human malignancies (Figure 1) [31].

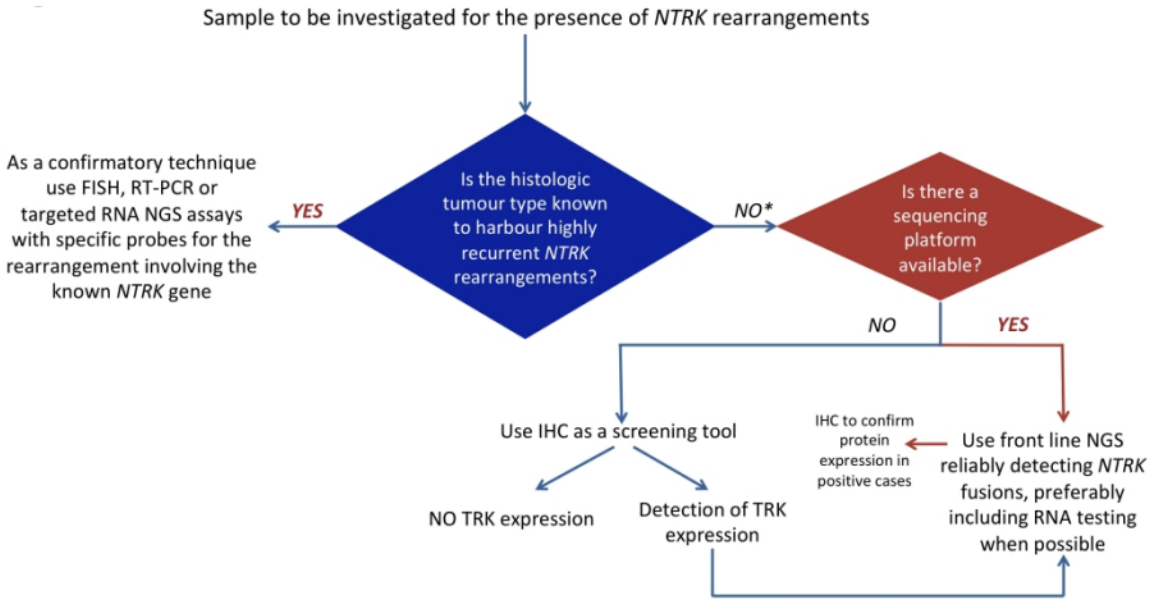


Figure 1. A rational approach for the detection of NTRK1/2/3 fusion genes in human malignancies
Reprinted from C Marchiò et al.

2.2.2.3 Diagnostic accuracy for detection of NTRK fusions

The CS did not consider the diagnostic accuracy of tests for NTRK fusion. The ERG notes that diagnostic accuracy of these tests could have a substantial impact on the efficacy of larotrectinib. No test is perfect; poor sensitivity will mean people with NTRK fusions will not be offered larotrectinib. More critically, poor specificity will mean people without NTRK fusion may be offered larotrectinib, from which they cannot benefit. Because many tumour sites have a NTRK fusion prevalence below 1%, even with a very high specificity the number of such "false positives" may exceed the number of genuine NTRK fusion cases.

The ERG sought to identify the possible diagnostic accuracy of genetic tests, by tumour sites. There appears to have been limited research on the topic, and what there is suggests that the diagnostic accuracy of the various tests to detect NTRK fusions is uncertain, and may vary by test and tumour site [32].

As an illustration, the ERG have considered the impact of diagnostic accuracy on the efficacy of larotrectinib. This assumed a 99% sensitivity and 99% specificity (plausible maximum accuracy; the actual accuracy may be lower) and the numbers eligible for larotrectinib from Table 2. The results are shown in Table 4. This shows that, for tumour sites where NTRK fusion is rare, false positive cases could make up the majority of all patients offered larotrectinib. Even if the response rate to larotrectinib were 80% in patients with a genuine NTRK fusion, these false positive patients would mean the overall response rate would drop to only [redacted].

If larotrectinib were offered only in cases with higher rates of NTRK fusion, or to children (i.e. for MASC, GIST, thyroid cancer, secretory breast cancer and all childhood sarcomas) there would be very few false positive cases, and the overall response rate would rise to [redacted].

Table 4 Impact of imperfect diagnostic testing on larotrectinib efficacy

Cancer type	Eligible for testing	Eligible and has NTRK fusion	True positives (has NTRK fusion detected)	False positives (no fusion but test is positive)	% of treated who are false positive	% of treated who respond (based on 80% response)
NSCLC cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Salivary cancer patients (non-MASC)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
MASC cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Melanoma cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Colorectal cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Appendix cancer patients (assumed same as pancreatic)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
STS adults (GIST) cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
STS adults (non-GIST) cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Bone Sarcoma cancer patients (assumed same as STS adults)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
STS paediatrics cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
STS infantile sarcoma cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Secretory breast cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cholangiocarcinoma cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CNS cancer patients (assumed same as brain)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CNS paediatric cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pancreas cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Thyroid cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total – all tumour sites			[redacted]	[redacted]	[redacted]	[redacted]
Total – sites with high NTRK fusion rate or paediatric cases			[redacted]	[redacted]	[redacted]	[redacted]

3 Critique of company's definition of decision problem

3.1 Population

The population in the NICE final scope is "People with NTRK fusion-positive advanced solid tumours who have:

- either progressed on or not responded to prior therapies
- are unfit for chemotherapy
- for whom no curative therapy exists."

The population addressed in the CS is adult and paediatric patients with solid tumours who have a tumour with a Neurotrophic Tyrosine Receptor Kinase (NTRK) gene fusion, and a disease that is locally advanced, metastatic or where surgical resection is likely to result in severe morbidity and no satisfactory treatment options. This is stated to be in line with the anticipated marketing authorisation, however, the company did not provide a definition of 'satisfactory treatment option'. In the clarification response the company stated that [redacted]

The clinical effectiveness evidence presented in the CS is informed by three, single arm clinical studies, all of which are still on-going.

One was a phase I multicentre, dose escalating clinical trial in adult patients with advanced solid tumours (LOXO-TRK-14001, NCT02122913). Patients from a dose expansion phase, documented with the NTRK fusion were included in the pooled analysis.

The second was a phase I/II multicentre, open-label clinical trial, in children with advanced solid or primary CNS tumours. (LOXO-TRK-15003, NCT02637687, SCOUT). Those documented with the NTRK fusion were included in the pooled analysis.

The third was a phase II, international, multicentre, open label "basket" trial in adults and adolescents over the age of 12, with recurrent, advanced solid tumours with a documented NTRK fusion (LOXO-TRK-15002, NCT02576431, NAVIGATE).

The pooled analysis consisted of data from the July 30th 2018 cut-off and included 102 patients across 15 distinct tissue histologies, for six of which - soft tissue sarcoma, salivary gland cancer, primary CNS tumours, thyroid cancer, lung cancer and bone sarcoma - there were multiple subtypes. Clinical advice to the ERG confirmed that theoretically NTRK fusion may be present in any solid tumour type.

The ERG identified thirty tumour sites in the literature [33] [34] that harbour NTRK fusion. Therefore, there are unrepresented tumour sites and the clinical evidence does not cover all those who would be eligible to receive larotrectinib in clinical practice.

Around one [redacted] ([redacted]/102 ([redacted]), (28/93 (30%) if excluding CNS patients)) of the patients in the pooled analysis set are children or adolescents, who may have more favourable outcomes than adults. Overall, 31/102 (30%) of patients (31/93 (33%) excluding CNS patients)) have very high (over 80%) NTRK fusion prevalence cancers, these are the paediatric cancers infantile fibrosarcoma and congenital mesoblastic nephroma and the salivary gland cancer, MASC. The proportion of patients with high NTRK fusion prevalence cancers in the clinical trial evidence may be greater than that of the total eligible population. However, in clinical practice these patients be more easily identified for treatment with larotrectinib, given the practical and cost-associated challenges with screening low prevalence NTRK fusion tumours. Therefore, they may be more likely to receive treatment[31].

The number of patients with NTRK2 fusions is low 10/102 (9.8%), primarily in patients with a primary CNS tumour (7/10). It is unclear how this reflects the distribution of NTRK2 fusions within the eligible populations and the ERG note that it may be underrepresented in the trial populations. Clinical advice to the ERG suggested that theoretically there is no reason that only one type of NTRK fusion should be present within any given tumour site.

The ERG also note that [redacted]102 ([redacted]) of patients in the clinical trial evidence have an Eastern Cooperative Oncology Group (ECOG) performance status of 1 or 0. The NICE scope covers patients who are “unfit for chemotherapy” and are therefore likely to have a poor performance status. There is uncertainty around the evidence for those patients with an ECOG performance status of 2 or above, who may be given larotrectinib in clinical practice.

Overall, the trial population within the CS falls into that specified in the NICE scope. Given the uncertainty around the identification of eligible patients in clinical practice, (see section 2.2.2.1), and how this will impact on those receiving the treatment, the ERG cannot say definitively the extent to which the trial populations reflect those who will receive larotrectinib in clinical practice. The ERG have concerns surrounding unrepresented tumour sites within the clinical evidence, the high proportion of paediatric and adolescent patients, the potential under-representation of patients with NTRK2 fusion and the uncertainty of the evidence in patients who have an ECOG performance status ≥ 2.

3.2 Intervention

The intervention is larotrectinib in line with the NICE scope. The dosing schedule is not specified in the NICE scope. In the pooled clinical evidence informing the submission, broadly larotrectinib was given to patients at either 100mg or 150mg orally b.d. ([redacted]/68 adults received 100mg b.d dosing of larotrectinib [redacted]68 received 150mg b.d. At least [redacted]34 paediatric patients received 100mg/m^2 b.d. dosing of larotrectinib). The phase I, LOXO-TRK-14001 and the phase I/II, LOXO-TRK-15003 both included a dose escalating phase. In the phase II, LOXO-TRK-15002, NAVIGATE trial a maximally tolerated dose of larotrectinib was not defined, however a dose of 100 mg twice daily was selected for adults and children who had a body-surface area of at least 1m^2.[35]

3.3 Comparators

The comparators specified in the NICE final scope are "Established management without larotrectinib". There are no current treatment options available that specifically target NTRK fusion cancers.

None of the included trials had a control arm, and so the submission presented no direct comparisons with larotrectinib. As an alternative, the company conducted a series of systematic literature reviews in tumour sites / locations known to harbour NTRK fusions and identified standards of care after patients have exhausted all satisfactory treatment option specific to tumour sites, which was typically best supportive care. In the absence of any data after the final line of approved active treatment, the company used a proxy such as the last line of active treatment. The company prioritised the most recent NICE technology appraisal (TA) sources, when NICE TAs were not available the company used best supportive care (or placebo) arms in trials identified from the systematic literature reviews. Identified comparators that subsequently include a mix best supportive care and active comparators, were weighted by patient enrolment per tumour location in the clinical trials.

The CS also reported sensitivity analyses using non-responding patients as a proxy for outcomes on best supportive care, and by comparing time to progression on previous line of therapy to progression-free survival on larotrectinib.

The ERG consider these approaches to be reasonable given the lack of suitable comparator data within the larotrectinib trials, but still inferior to the use of actual patient data on patients receiving best supportive care. The ERG present a full critique of the identification of comparator data and the indirect comparison analyses in section 4.3 and 4.4

3.4 Outcomes

The outcomes specified in the CS match the NICE final scope and include overall survival (OS), progression free survival (PFS), and overall response rate (ORR), duration of response (DoR), adverse effects and health-related quality of life. The pooled analysis of clinical trial data specifies ORR as the primary endpoint by independent review committee assessments. Secondary outcomes are ORR by investigator assessment, DoR, PFS, OS and safety. Other reported outcomes are time to response, best response, time on treatment and disease control rate.

4 Clinical Effectiveness

This section contains a critique of the methods of the systematic review of clinical effectiveness and safety data on larotrectinib, followed by a description and critique of the included studies.

4.1 Critique of the company's systematic review methods

The company conducted a systematic review to identify studies on the effectiveness and safety of larotrectinib for treating NTRK fusion-positive advanced solid tumours. Further reviews were also conducted to identify comparator treatments in tumour sites known to harbour NTRK gene fusions, reflective of the patients investigated in the larotrectinib studies (see Section 4.3).

4.1.1 Searches

Appendix D of the company submission contained a detailed description of the searches undertaken to identify clinical evidence on larotrectinib. The following databases were searched on 11th March 2019: MEDLINE, MEDLINE in process and EMBASE via Proquest Dialog, and the Cochrane Library via Wiley. The company confirmed in the responses for clarification that CENTRAL and CDSR were the only two database searched via the Cochrane Library. The following Conference proceedings (2016-2018) were searched: American Society of Clinical Oncology (ASCO) and the European Society of Medical Oncology (ESMO). In addition reference checking of included studies and reviews (published in 2018 and 2019) was undertaken to identify any further relevant studies.

The sources searched to identify both published and unpublished studies were appropriate. The search strategies were examined in detail by the ERG and several issues were identified which may have affected retrieval of relevant studies. The search strategy presented for MEDLINE and EMBASE contained search terms for NTRK fusion-positive tumours combined with terms for the intervention larotrectinib or terms for any other treatments. To ensure comprehensive retrieval of all possible interventions for NTRK fusion positive tumours it would have been more effective to search using terms for the population only.

Additional intervention terms such as tissue agnostic, tumour agnostic, histology agnostic, histology independent and TRK inhibitors could have been added as alternative intervention terms to improve the comprehensiveness of the search. Truncation was missing or in the wrong place for some of the terms. Search line S1 missed alternative terms for NTRK and TRK such as neurotrophic tropomyosin receptor kinase, tropomyosin receptor kinase and tyrosine receptor kinase.

The search strategy used for the Cochrane Library contained only one search term *TRK*, with left and right hand truncation. The ERG checked this and found that no additional studies would be found by including the additional terms for the population used in the MEDLINE strategy reported in the

company submission. However, the MEDLINE strategy was found to have missed some alternative terms for NTRK and TRK (neurotrophic tropomyosin receptor kinase, tropomyosin receptor kinase and tyrosine receptor kinase). Adding these further terms identified by the ERG may have improved retrieval of relevant studies from the Cochrane Library.

4.1.2 **Inclusion criteria**

The larotrectinib review eligibility criteria were presented in Appendix D of the CS, Table 64. The population criterion was people with NTRK fusion-positive solid tumours. This is broader than the criteria specified in the NICE decision problem: patients with advanced tumours which have either progressed or not responded to prior therapies or patients unfit for chemotherapy or for whom no curative therapy exists. The comparator criteria were similarly broad, being any treatment apart from surgery, although the PRISMA diagram (Figure 37 in Appendix D) indicated that experimental treatments (i.e. not standard care) were subsequently excluded at the data extraction phase. Although these broader criteria may have led to the unnecessary screening and data extraction of some studies, it should not have adversely affected the identification of relevant studies. The outcomes included were as per the NICE scope, with the exception of duration of response which was omitted. The study design criteria were appropriately broad, given the likely absence of RCT evidence. The screening methods used were appropriate for minimising the possibility of reviewer errors and biases affecting the final list of studies included.

4.1.3 **Critique of data extraction**

For the review of larotrectinib evidence, data extraction methods were not reported in the CS. It seems likely that one reviewer extracted data which were then checked by a second reviewer, as was described in the review of comparators (Appendix D p57).

4.1.4 **Quality assessment**

The quality assessment of the included studies was reported in the section B.2.5 of the CS with further details given in Table 80 of Appendix D. The Downs and Black checklist was used in which 27 questions were answered Yes, No, Unclear or Not applicable, with studies given a summary score.

Although scoring systems are not recommended for quality assessments in systematic reviews - primarily because high scoring studies may still have important deficiencies which may be overlooked when using the summary score – the use of the Downs and Black checklist in this appraisal appears understandable, given the sparse guidance available on how to quality assess single arms studies or basket trials. Nevertheless, the Downs and Black checklist has important limitations in the context of this appraisal: 10 questions are based on quality of reporting rather than on the quality of study

methods or conduct and only one of the six questions in the confounding section (see Table 80, Appendix D of the CS) appears to be relevant to single-arm basket trials.

A key aspect of the quality of a basket trial will be its external validity, or applicability to the review question, but the Downs and Black Checklist has only 3 questions on external validity. On p67 the CS stated that "there was no information on patients that were asked to participate in the study but declined or were screening failures". The ERG requested details about this (in the form of CONSORT diagrams) but these have not been supplied.

As with the data extraction stage, no details were provided about how many researchers were involved in the quality assessment process so the possibility of errors or bias affecting the assessments cannot be ruled out (although it seems likely that two researchers were involved, as was described for the review of comparator evidence).

4.1.5 **Evidence synthesis**

Data from the three larotrectinib studies cited in the submission were pooled by simple addition of the individual trial results. Given the target-specific mode of action of larotrectinib together with the included studies having single-arms, small sample sizes, and similar eligibility criteria this approach to data pooling is understandable. However, the heterogeneous populations included in the datasets raises concerns about how appropriate it is to pool all the trial data together. The ERG's clinical advisers thought that tumour type may be likely to influence PFS and OS following treatment with larotrectinib. It may also be inappropriate to pool data from adults and children together since there may be considerable variation in time to progression and in subsequent overall survival. For example, a key issue could be variation across tumour sites in the possibilities for potentially curative surgery following successful treatment with larotrectinib. Further heterogeneity was evident as a consequence of two of the three studies having dose-finding phases. This meant that some patients contributed to the pooled datasets who were on doses different to those anticipated in the license (see Section 4.2).

The CS stated that a pooled sample size of 55 patients was sufficient to provide 80% power to achieve a lower boundary of the 2-sided 95% exact binomial CI about the estimated ORR exceeding 30% (ruling out a lower limit of 30% for ORR was considered clinically meaningful). The submission focussed on presenting effectiveness results for 102 patients relating to the most recent data cut-off of 30 July 2018. These patients came from two such datasets: an 'extended primary analysis set' called ePAS2 (n=93), which excluded patients with primary CNS tumours, and a 'supplementary analysis set' called SAS3 (n=9) which included only patients with primary CNS tumours. The ERG requested clarification about why patients were excluded from these analysis sets but included in the larger safety analysis set (n=137). The company stated that 28 patients were excluded because they were

recruited later than 19th February 2018 and so would not have the potential of six or more months of follow-up by the July 2018 data cut; the accompanying figure revealed that for these 28 patients there was insufficient follow up to permit IRC assessment. Seven patients were excluded for having a solid tumour with NTRK fusion but no measurable lesion.

Given the small numbers of patients in the analyses (n=102) and that larotrectinib responses are usually (though not always) achieved by month 2, the inclusion of these 35 patients might have helped to reduce response rate uncertainty and concerns regarding whether there was any variation in response by subgroups. However, this would have meant using investigator data, rather than IRC data which were not yet available. Nevertheless, this approach was used - and was in fact unavoidable - for the SAS3 (primary CNS tumours) dataset, since only investigator assessments were available (as disease was not independently assessed). Importantly, the inclusion of the 35 excluded patients would also have added more data to the PFS and OS analyses (although the data would be immature). The ERG also requested data on the flow of participants from screening to inclusion in the analyses including the number of patients screened for eligibility, number ineligible/excluded (with reasons), etc. but these have not been supplied.

4.2 Critique of trials of the technology of interest, their analysis and interpretation

The submission included three ongoing single-arm, open-label trials of larotrectinib which are summarised in Table 5 (adapted from Table 5 of the CS):

- LOXO-TRK-14001 (n=10): a phase I study which contributed 8 patients the pooled analysis
- LOXO-TRK-15003 a.k.a. SCOUT (n=45): a phase I/II study which contributed 32 children to the pooled analysis;
- LOXO-TRK-15002 a.k.a. NAVIGATE (n=82): a phase II basket trial which contributed 62 adults to the pooled analysis

4.2.1 Design and analysis of basket trials

The largest trial, NAVIGATE, has a basket trial design. When studying targeted agents investigators need to know whether the drug works uniformly in all cancer sites with the mutation of interest or whether treatment effectiveness is site dependent. The basket design, in which patients are recruited to gain knowledge of a drug's efficacy in distinct cancer sites - or baskets - is a design strategy to address this issue.[36] Basket trials evaluate therapies which have a mechanism of action based on targeting a specific genomic alteration, irrespective of tumour histology. Typically, two-stage studies are designed to recruit a certain number of patients to each 'basket', in this case a tumour site, and if a pre-specified proportion of patients in a particular basket respond, then recruitment is expanded

within this disease area. If too few responses are observed within a basket then recruitment is stopped due to low promise of efficacy.

Larotrectinib is described as being 'tumour agnostic', with patients being eligible for trials based on the specific genomic alteration in question. However, despite being 'tumour agnostic' in *target*, basket trial therapies might show heterogeneity of effectiveness across different tumour types in terms of treatment response and in the development of resistance or loss of response and its subsequent impact on PFS and OS. An example is the basket trial of vemerafinib in patients with BRAF V600E mutations: vemerafinib was active in NSCLC and other histologies, but not in colorectal cancer.[37] The authors of the vemerafinib trial concluded that "the histologic context is an important determinant of response in BRAF V600–mutated cancers".

Heterogeneity of response across baskets is an important issue in the design and analysis of conventional basket trials, and care must be taken to accommodate the potentially large variation and imprecision in response rate estimates introduced by very small sample sizes. The approach taken by the company was to assume equal efficacy across all baskets and to generate a pooled response estimate, but in doing so reject the potential for heterogeneity of response across baskets.

As the included trials were not designed or sufficiently powered to test the assumption of heterogeneity of response across subgroups, the ERG consider it inappropriate to assume a common response rate independent of tumour histology. In particular, poor response in patients with primary CNS tumours (1/9 patients) suggests that this assumption is unlikely to hold across all tumour sites.

In light of this the ERG requested subgroup results from the company based on tumour site, fusion type, fusion partner, age, and response status. This would permit a more detailed analysis of potential heterogeneity. However the company declined to provide such data, on the grounds that:

1) [redacted]
2) [redacted]

The ERG disagrees with this justification, and considers that at least some analyses by subgroups are justified. As detailed data by subgroup were not provided a more limited analysis of heterogeneity was performed (see section4.6).

4.2.2 **Details of the included trials**

All three included trials are multi-site, ongoing studies which recruited patients with locally advanced or metastatic solid tumours, or primary central nervous system (CNS) tumours who were previously treated with standard therapy (if it was available or possible, Table 5). Both SCOUT and LOXO-TRK-14001 had dose-finding phases, so some patients included in the pooled analysis did not receive the proposed licenced dose of 100mg twice daily (for adults) or 100mg/m^2 twice daily (for children). Two adults received 150mg and Table 5 of the CS indicated that 9 children did not receive 100mg/m^2. Two of the five case studies presented in the CS related to patients who had starting doses of 150mg (or equivalent, in the child case) so are not the best representation of how larotrectinib will be administered in the NHS.

Table 5 Summary of the larotrectinib trials methods (adapted from Table 5 of the CS)

	Trial acronym and number		
	NAVIGATE, LOXO-TRK-15002	**LOXO-TRK-14001**	**SCOUT, LOXO-TRK-15003**
Trial design	Phase II, multicentre, open-label basket study	Multicentre, open-label, phase I, dose-escalation and dose expansion study	Multicentre, open-label, phase I/II study
Clinicaltrials.gov	NCT02576431	NCT02122913	NCT02637687
Location	35 sites: Asia and Europe	8 sites: U.S.	26 sites: Australia, North America, Europe
Duration	October 15, 2015 – ongoing	May 1, 2014 – ongoing	December 22, 2015 – ongoing
Eligibility criteria	*Inclusion:* -Age ≥12 years -Locally advanced or metastatic solid tumour with documented NTRK gene fusion that could be assessed according to RECIST, version 1.1 -Previously treated with standard therapy (if available or possible) -ECOG PS 0–3 -adequate organ function -life expectancy of ≥3 months -Patients with primary CNS tumours or metastasis who were neurologically stable *Exclusion:* -Current treatment with a strong CYP3A4 inhibitor or inducer - An investigational or anticancer therapy within 2 weeks, or major surgery within 4 weeks, prior to enrolment -previous treatment with kinase inhibitors -clinically significant cardiovascular disease or history of prolonged QT interval corrected for heart rate (QTc) - Symptomatic or unstable brain metastases	*Inclusion:* -Age ≥18 years -Locally advanced or metastatic solid tumour with documented NTRK gene fusion for expansion phase of study -Previously treated with standard therapy (if available or possible) -ECOG PS 0–2 -adequate organ function -life expectancy of ≥3 months -Patients with primary CNS tumours or metastasis who were neurologically stable, and did not require steroid management of CNS symptoms within 2 weeks before entry *Exclusion:* -Current treatment with a strong CYP3A4 inhibitor or inducer - An investigational or anticancer therapy within 2 weeks, or major surgery within 4 weeks, prior to enrolment - clinically significant cardiovascular disease or history of prolonged QT interval corrected for heart rate (QTc)	*Inclusion:* -Age 1 month–21 years; -Locally advanced or metastatic solid tumour or primary CNS tumour or patients with locally advanced IFS who required disfiguring surgery or limb amputation to achieve surgical CR Measurable disease (per RECIST v1.1, RANO criteria, or International Neuroblastoma Response Criteria) with documented NTRK gene fusion for expansion phase / phase II -Previously treated with standard therapy (if available or possible) - Karnofsky (≥16 years) or Lansky (<16 years) PS of ≥50 *Exclusion:* -Clinically significant cardiovascular disease or corrected QT interval >480 ms -an active uncontrolled systemic infection -any conditions affecting oral absorption Current treatment with a strong CYP3A4 inhibitor or inducer

	Trial acronym and number		
	NAVIGATE, LOXO-TRK-15002	**LOXO-TRK-14001**	**SCOUT, LOXO-TRK-15003**
	-any conditions affecting oral absorption	-any conditions affecting oral absorption	-receipt of an investigational or anticancer therapy within 2 weeks, or major surgery within 4 weeks, prior to enrolment
Trial drugs and method of administration	Oral larotrectinib 100 mg b.d. in 28-day cycles. Larotrectinib was administered as capsules unless patients could not swallow capsules, in which case a liquid formulation was available.	**Dose escalation:** Oral larotrectinib, once- or twice-daily, on a continuous 28-day schedule, in increasing dose levels according to a standard 3+3 dose escalation scheme. Dose levels: 50 mg q.d. / 100 mg q.d. / 200 mg q.d. / 100 mg b.d. / 150 mg b.d. / 200 mg b.d. **Expansion Phase** Oral larotrectinib 100 mg b.d.	**Dose escalation:** Oral larotrectinib (capsule or liquid) Cohort 1: Doses ranging from 17%–96% of the BSA-adjusted recommended adult phase 2 dose of 100 mg b.d. Cohort 2: Doses ranging from 30%–208% of the BSA-adjusted adult dose of 150 mg b.d. Cohort 3: 100 mg/m² b.d. Dosing was continuous for 28-day cycles. **Phase II:** Oral (capsule or liquid). 100 mg/m2 b.d., not to exceed 100 mg b.d.
	Larotrectinib was administered until disease progression, the occurrence of unacceptable toxicity, or the withdrawal of patient consent.		
Permitted and disallowed Concomitant medication	**Permitted:** Palliative radiotherapy to specific sites of disease; Standard supportive medications; standard of care medications received for the previous 28 days at stable doses; glucocorticoids for primary CNS tumour patients **Disallowed:** Other anti-tumour approved or investigational agents that were being used with the intent to effect tumour shrinkage (e.g. chemotherapy); known strong inhibitors or inducers of CYP3A4; any other investigational agents		
Outcomes	ORR (CR+PR), best OR, DoR, PFS, OS, safety, HR-QoL	Safety, ORR (CR+PR), DoR	ORR (CR+PR), best OR, DoR, PFS, OS, safety, HR-QoL

AEs adverse events, b.d.twice-daily; CBR clinical benefit rate; CNS central nervous system; CR complete response; CTCAE Common Toxicity Criteria Adverse Events; DLT Dose-limiting toxicity; DOR duration of response; ECOG Eastern Cooperative Oncology Group; HR-QoL health-related quality of life; MTD maximum tolerated dose; OR Objective response, ORR overall response rate; PR partial response; PS performance status; q.d. once-daily; RANO Response Assessment in Neuro-Oncology; SD stable disease

Table 8 in the CS presented baseline characteristics for each study and for the pooled dataset (n=102). An updated table was presented as part of a clarification response to a question regarding the company's rationale for excluding patients from the efficacy evaluable datasets, an adapted version of this is presented below (Table 6). Fifteen different cancers were represented with the most common being soft tissue sarcoma (n=20), salivary gland tumour (n=17), infantile fibrosarcoma (n=13), thyroid cancer (n=10) and primary CNS tumours (n=9). The median age of patients was 41 years (range 0.1 to 78) for non-CNS solid tumours and 12 years (range 2 to 79) for primary CNS tumours. Most ePAS2 patients had NTRK 1 (44%) or NTRK 3 (48%) gene fusions whereas most SAS3 patients had NTRK 2 fusions (78%). Twenty-seven different NTRK fusion partners were found across the study population.

Most patients had an ECOG performance status of 0 or 1 (89%) and all stages of diagnosis (I-IV) were adequately represented, though this was unknown in 18% (17/93) of patients in ePAS2 and 44% (4/9) patients in SAS3.

The CS reported that for 23% of non-CNS tumour patients larotrectinib was the initial systemic therapy as there was no standard of care. The ERG's clinical advisors were of the opinion that the larotrectinib trial populations were likely to be generalisable to patients seen in NHS settings.

Table 6 Baseline characteristics for the efficacy evaluable patients: pooled analysis datasets and individual larotrectinib studies (adapted from Table 8 of the clarification response)

Baseline Characteristic	ePAS2 dataset n=93	SAS3 dataset n=9	NAVIGATE LOXO-TRK-15002 N=62	LOXO-TRK-14001 N=8	SCOUT LOXO-TRK-15003 N=32
Median age, years	41.0	12.0	[redacted]	[redacted]	[redacted]
Mean age, years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sex, n (%): Male	49 (53)	5 (56)	[redacted]	[redacted]	[redacted]
Female	44 (47)	4 (44)	[redacted]	[redacted]	[redacted]
ECOG PS, n (%): 0	42 (45)	5 (56)	[redacted]	[redacted]	[redacted]
1	41 (44)	3 (33)	[redacted]	[redacted]	[redacted]
2	10 (11)	1 (11)	[redacted]	[redacted]	[redacted]
Tumour type, n (%)			[redacted]	[redacted]	[redacted]
NSCLC	7 (8)	-	[redacted]	[redacted]	[redacted]
IFS	13 (14)	-	[redacted]	[redacted]	[redacted]
STS	20 (22)	-	[redacted]	[redacted]	[redacted]
Colon	6 (6)	-	[redacted]	[redacted]	[redacted]

Baseline Characteristic	ePAS2 dataset n=93	SAS3 dataset n=9	NAVIGATE LOXO-TRK-15002 N=62	LOXO-TRK-14001 N=8	SCOUT LOXO-TRK-15003 N=32
Salivary gland	17 (18)	-	[redacted]	[redacted]	[redacted]
Breast	1 (1)	-	[redacted]	[redacted]	[redacted]
Pancreas	1 (1)	-	[redacted]	[redacted]	[redacted]
Thyroid	10 (11)	-	[redacted]	[redacted]	[redacted]
Bone sarcoma	2 (2)	-	[redacted]	[redacted]	[redacted]
Cholangiocarcinoma	2 (2)	-	[redacted]	[redacted]	[redacted]
GIST	5 (5)	-	[redacted]	[redacted]	[redacted]
Melanoma	7 (8)	-	[redacted]	[redacted]	[redacted]
Appendix	1 (1)	-	[redacted]	[redacted]	[redacted]
Primary CNS	-	9 (100)	[redacted]	[redacted]	[redacted]
Congenital mesoblastic nephroma	1 (1)	-	[redacted]	[redacted]	[redacted]
Stage at diagnosis, n (%)			[redacted]	[redacted]	[redacted]
I	10 (11)	1 (11)	[redacted]	[redacted]	[redacted]
II	16 (17)	0	[redacted]	[redacted]	[redacted]
III	25 (27)	2 (22)	[redacted]	[redacted]	[redacted]
IV	25 (27)	2 (22)	[redacted]	[redacted]	[redacted]
Not reported/Unknown	17 (18)	4 (44)	[redacted]	[redacted]	[redacted]
Disease extent at enrollment n (%)			[redacted]	[redacted]	[redacted]
Locally advanced	16 (17)	4 (44)	[redacted]	[redacted]	[redacted]
Metastatic	77 (83)	0	[redacted]	[redacted]	[redacted]
Other / not reported	-	5 (56)	[redacted]	[redacted]	[redacted]
Prior cancer therapy - Yes, n (%)	90 (97)	9 (100)	[redacted]	[redacted]	[redacted]
Surgery	78 (84)	5 (56)	[redacted]	[redacted]	[redacted]
Radiotherapy	45 (48)	5 (56)	[redacted]	[redacted]	[redacted]
Systemic therapy	72 (77)	9 (100)	[redacted]	[redacted]	[redacted]
0 prior systemic	21 (23)	0	[redacted]	[redacted]	[redacted]
1-2	46 (49)	8 (89)	[redacted]	[redacted]	[redacted]
≥3	26 (28)	1 (11)	[redacted]	[redacted]	[redacted]
Mean no. prior systemic	1.8±1.8	1.7±1.3	[redacted]	[redacted]	[redacted]
Median no. prior systemic	1.0 (0-10)	1.0 (1-5)	[redacted]	[redacted]	[redacted]

Baseline Characteristic	ePAS2 dataset n=93	SAS3 dataset n=9	NAVIGATE LOXO-TRK-15002 N=62	LOXO-TRK-14001 N=8	SCOUT LOXO-TRK-15003 N=32
***NTRK* gene fusion, n (%)**			[redacted]	[redacted]	[redacted]
NTRK1	41 (44)	1 (11)	[redacted]	[redacted]	[redacted]
NTRK2	3 (3)	7 (78)	[redacted]	[redacted]	[redacted]
NTRK3	45 (48)	1 (11)	[redacted]	[redacted]	[redacted]
Inferred NTRK3	4 (4)	-	[redacted]	[redacted]	[redacted]

4.2.3 Summary of the pooled dataset effectiveness results

4.2.3.1 Overall response rate

The primary outcome was ORR (i.e. CR or PR) as determined by Independent Review Committee (IRC) assessments, based on RECIST (Response Evaluation Criteria in Solid Tumours) for non-CNS solid tumours. For the ePAS2 analysis set the ORR was reported as being 72% (67/93) which included 16 (17%) patients with a complete response. However, one of these was a "surgical CR" which was actually a PR to larotrectinib. The ORR using investigator assessment was 8% higher, driven primarily by more patients being judged as having a partial response. SAS3 patients did not have independent assessments of disease; the investigator-assessed ORR was notably lower than for ePAS2, being one out of nine patients (11%). Whereas the ePAS2 population included a tiny minority of NTRK-2 patients (3%), 7 out of the 9 SAS3 patients (78%) were NTRK 2.

Possible reasons for the difference between the ePAS2 and SAS3 ORR results could be:

- Primary CNS patients may be less likely to respond to larotrectinib than other patients
- NTRK2 patients may be less likely to respond to larotrectinib than other patients
- NTRK2 patients may be more likely to have false-positive NTRK test results (such patients cannot respond)
- The very low ORR seen in the SAS3 population (n=9) may have been a chance result

More broadly, some of the non-responders may not have had NTRK fusions i.e. they had false-positive NTRK fusion test results, or the fusion may not be expressed at the protein level. The Drilon 2018 paper[35] on larotrectinib, which reports pooled trial data at a July 2017 cut-off (n=55, called the original PAS in the CS), noted this to be the case for three non-responding patients who had samples available for re-testing for NTRK fusions. No central assay methods were used in the three studies. Data on false positive test results and their consequences were not reported in the CS.

Heterogeneity in response rate

Given the difference in ORR between the EPAS2 and SAS3 datasets, and to investigate the probability of further heterogeneity in outcomes, the ERG requested subgroup results for both datasets for the following categories: tumour site, age (adults vs children), response status (responders vs non-responders). Results for the following outcomes were requested: PFS and OS (median, 6 month rate and 12 month rate), time on treatment and duration of response. The company responded saying they did "not believe that providing subgroup data was justified or helpful since there is no evidence of heterogeneity in treatment effect for the subgroups requested and because patient numbers are already small and further post-hoc 'slicing and dicing' of the data will only serve to increase uncertainty".

The ERG disagrees with the company's view that there is no evidence of treatment effect heterogeneity, which is why the subgroup results were requested. Results were not provided for the outcomes requested by the ERG. The company instead provided limited subgroup results data for the ePAS2 dataset but for ORR, some of which was repetition of results already provided in the original submission. The company did not provide subgroup data for the SAS3 (primary CNS) cohort neither in the CS nor in response to the ERG's request. The ERG conducted analyses to explore the likelihood of differences in ORR and in PFS across tumour sites (see section 4.6.2).

Updated response data

Given the data cut-off point used in the CS was some time ago - July 2018 - the ERG also requested more up to date results. In response, the company provided a conference abstract which had a data cut-off date of 19 February 2019 and which reported results for a total of 159 patients (153/159 patients evaluable for efficacy). The relevant results presented were an ORR of 79% (95% CI 72% to 85%) and [redacted]. No subgroup results were reported and it was unclear whether or not the cohort of 159 patients included primary CNS patients. Moreover, the results reported were based on investigator assessments – as noted earlier, the availability of only investigator assessment results was the reason for excluding 28 patients from the main analysis set used in the CS (i.e. no IRC assessment available). These updated results were therefore of little value in terms of resolving uncertainty surrounding the effectiveness of larotrectinib. Updated results were also presented for a smaller 'primary' cohort of 55 patients used for the FDA assessment: the [redacted]

4.2.3.2 Time to response and duration of response

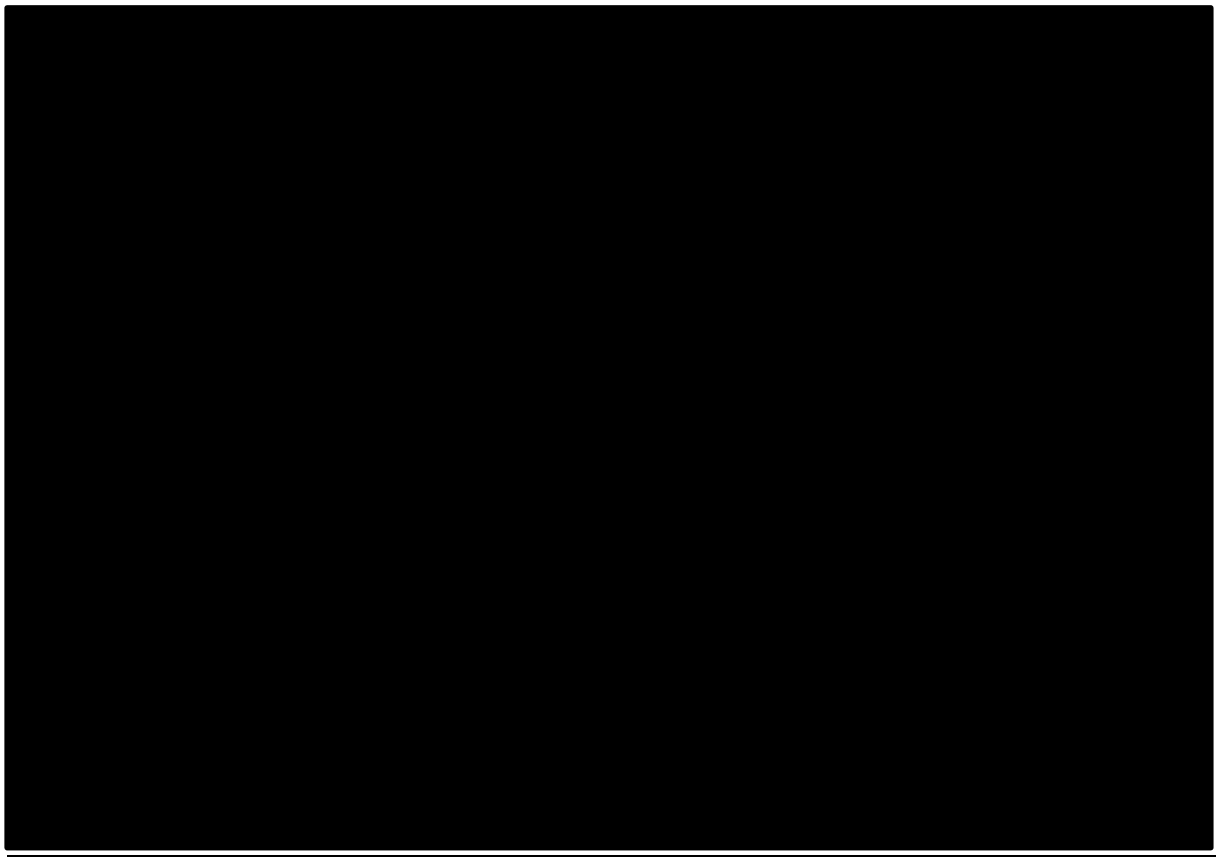
Most (82%) of the 67 responding ePAS2 dataset patients (based on IRC assessments), did so by the month 2 assessment, with a median time to response of 1.8 months. All partial responses had occurred by month [redacted]. Of these 67 patients, 50 (75%) were still in response (at the last data cut) and 17 had progressed disease; the median duration of response had not been reached (after a median follow up of 12.7 months). For the SAS3 patients Figure 13 of the CS showed that the one responder responded by month 2 and of the remaining 8 non-responders all but one had been treated for more than 2 months with 4 having been treated for more than 6 months. Based on the median time to response of **1.8** months in the ePAS2 dataset, and the fact that all partial responses had occurred by month [redacted], it appears very unlikely that many of the SAS3 non-responders will go on to achieve a response.

The ERG notes that this raises questions about the appropriate duration of treatment with larotrectinib when a response is not observed (i.e. a stopping rule); for example, whether it should be used for a maximum of [redacted] months. This was not considered in the CS, and could impact on both clinical and cost effectiveness.

4.2.3.3 Progression free survival and overall survival

The CS presented results for patients who had measurable disease and at least six months of follow up which meant that PFS and OS data were missing for 35 patients who had received at least one dose of larotrectinib. The ERG requested results which included these patients. The company provided results based on investigator assessments – see [redacted] for the PFS Kaplan-Meier (KM) plot. The median PFS in the [redacted] The median follow-up for PFS was [redacted] months. [redacted] had progressed disease by the 30 July data cut-off. For ePAS2 the median PFS was [redacted] with a 6-month PFS rate of [redacted] and a 1-year rate of [redacted].

[redacted]



The K-M plot for OS is presented in [redacted]. [redacted] had died by the 30 July 2018 data cut-off. The median follow-up for overall survival was [redacted] months and the median duration of OS [redacted]. The ERG's clinical advisers thought that tumour type may be likely to influence PFS and OS following treatment with larotrectinib but, as previously discussed, the ERG's request for tumour type and other subgroup data on PFS and OS was not fulfilled.

[redacted]

Post-response and post-progression interventions

Figure 12 (p80 of the CS) showed that [redacted] patients had received surgery after achieving a partial response indicating that for these cases larotrectinib was acting as a bridge to surgery. No further details were provided in the CS and since the surgery is likely to affect PFS and OS (and hence cost-effectiveness) the ERG asked the company which types of tumour were resected: [redacted]. The specificity of these data suggest that heterogeneity in PFS and OS across certain tumours sites is likely.

Intrinsic or acquired resistance has been noted as a major limitation of targeted anticancer therapies.[38] Since post-progression treatments might also affect estimates of OS the ERG requested details of treatments received after disease progression or the development of resistance to larotrectinib. The company stated that of the 93 ePAS2 patients, 34 (37%) had progressed - with mutations identified as a mechanism for resistance in [redacted] patients. [redacted] of the 93 patients continued to receive larotrectinib post-progression with the duration of treatment ranging from [redacted] days ([redacted] patients continuing to receive treatment). The median duration of post-progression treatment was [redacted] months and the mean was [redacted]. [redacted] post-progression patients received LOXO-195, an experimental

therapy manufactured by the company for patients who become resistant to TRK inhibitors. In a study of LOXO-195, in which all the patients had prior exposure to a TRK inhibitor (larotrectinib in 21 cases, entrectinib in 9, and PLX7486 in 1), an ORR of 34% (10/29) was seen.[39] The company stated that if patients go on to receive other interventions which are not currently available or recommended for use in the NHS after larotrectinib, they would not expect to adjust for this in any cost-effectiveness analysis and that such interventions were not a consideration in overall survival analysis. The company added that this was because it is not unreasonable to assume that some patients would go on to receive further innovations as part of a clinical trial, compassionate access to medicines not yet licensed, or drugs approved via a system of individual funding requests.

Whilst the ERG acknowledges that some resistant patients would receive experimental treatments via these pathways, the number of patients receiving LOXO-195 (as opposed to other experimental treatments) suggests that the impact of experimental LOXO-195 on survival after progression should not be ignored. The importance of this issue was also evident following the ERG's clinical advisers' opinion that it is likely that all patients would eventually develop resistance to larotrectinib meaning that in the future many patients may become eligible to receive LOXO-195, should it become licensed.

4.2.3.4 Health-related quality of life

The CS noted (on p85) that the HRQoL results were not statistically significant and should be interpreted cautiously. Of the 137 patients who received ≥1 dose of larotrectinib only [redacted] adults and [redacted]children ([redacted]) had HRQoL data both at baseline and for at least one post-baseline time point. It is unclear how representative these patients are of the 137 patients who received ≥1 dose of larotrectinib. Results were presented for the following outcome measures: EORTC QLQ-C30 (adults), EQ-5D-5L (adults) and PedsQL (both for children<2 years and for children ≥2 years). A further limitation of these data is that they are based on subjective, self-assessed outcomes in an open-label study. This means it is very difficult to distinguish the extent of the treatment effect from any patient expectation effects. Bias will also have been introduced by reporting results for the 'best change' from baseline, rather than changes from baseline at specific time points. However, the utility health state values for larotrectinib were derived using all available data (p171 of the CS).

No results data were presented for the SAS3 (primary CNS tumours) cohort.

4.3 Critique of systematic review for indirect comparison analyses

Only single arm data were available on the effectiveness and safety of larotrectinib (i.e. an absence of RCTs) and the company did not have access to individual patient data (IPD) from patients receiving relevant comparator treatments. Therefore, to compare the outcomes of patients receiving larotrectinib

to those receiving other relevant therapies, or best supportive care the CS used summary data from previous STAs, and published trials. These were identified via a series of systematic reviews of interventions for tumour types known to harbour NTRK gene fusions.

As the CS only summaries these reviews the ERG requested further details on how relevant comparator data were identified, including documentation and reasons for selection decisions and the full report describing the systematic reviews (cited as reference 77 in the CS). This is reviewed here.

4.3.1 **Searches**

The following databases were searched: MEDLINE (via PubMed), EMBASE, and the Cochrane Library. The interface/provider was not reported for EMBASE or the Cochrane Library. All searches took place between May – August 2018 and were updated during January - March 2019.

In addition to the database searches, a comprehensive set of sources were searched for unpublished, grey literature, including five clinical trials registers and the following conference proceedings: American Society of Clinical Oncology (ASCO) and the European Society of Medical Oncology (ESMO) and the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) International and European Annual Meetings. Searches of several key international HTA websites were also undertaken, including the NICE website. In some of the search strategies for EMBASE, retrieval of conference abstracts was limited to those published during 2015-2019.

Fourteen sets of database searches were undertaken in total, one for each of the tumour sites/locations. Most of the searches followed a similar structure with terms for the population combined with terms for the interventions and comparators with a limit to RCTs. However, some of the searches were broader, including terms for the population only, to identify all possible studies of the various interventions and comparators used to treat the specific condition. For some of the searches terms for the population were limited to advanced forms of the particular cancer/tumour.

In general, the searches were conducted appropriately, search lines were combined correctly and no major errors were found by the ERG. However certain restrictive elements were found in the search strategies which may have impacted on the comprehensiveness of the search. The search terms used in several of the strategies to restrict retrieval to RCTs were fairly limited (three terms only). This could have led to relevant RCTs being missed. Validated, sensitive RCT search filters are available for PubMed and EMBASE however these appear not to have been used in some of the strategies where RCTs only were required. In addition, the searches of the Cochrane Library which includes the clinical trials database CENTRAL and the Cochrane Database of Systematic Reviews (CDSR), were also restricted to RCTs in some of the strategies. This is unnecessary as CENTRAL only contains

controlled clinical trials and CDSR only contains systematic reviews. Applying this limit to RCTs in the searches of the Cochrane Library may have caused relevant trials or reviews to be missed by the searches.

4.3.2 Selection of TAs and trials

The company reported that their data selection approach prioritised the most recent NICE technology appraisal (TA) sources. When several sources were available the appropriateness of the identified NICE TAs were judged based on:

- The trial comparator arm used as a proxy of BSC accepted by the ERG/NICE
- The extent to which the publication matched the treatment criteria in the larotrectinib trial protocol
- The source with the most advanced patients (e.g. last line of systemic therapy)
- The date of publication, prioritising more recent publications, and those reporting outcomes that inform the economic model inputs

When NICE TAs were not available the company used best supportive care (or placebo) arms in trials identified from the systematic review, or identified via further searching if required. For each tumour type covered by the larotrectinib studies the company provided justifications for the data sources selected. The company's systemic review methods for identifying studies were appropriate and their approach and justifications for selecting specific comparator data appeared reasonable although, given the extensive list of tumour sites, it was not practicable for the ERG the check the appropriateness of specific selection decisions. Tables 25 and 26 of the CS provided summaries of the comparator treatment efficacy and safety for each tumour type.

4.4 Critique of the indirect comparison analyses

4.4.1 Comparison to previous NICE TAs and trials

The PFS and OS comparator data used in the model were reported in Appendix M of the CS. However, the company's submission did not present how these data compared with the corresponding data for larotrectinib, since the company did not provide PFS or OS results by tumour type for the larotrectinib cohort. Similarly, comparisons could not be made of baseline characteristics - such as ECOG status or disease stage - which may have an important impact on PFS and OS.

For each tumour site Kaplan-Meier data were extracted from the selected data source and digitised. Parametric survival curves were fitted to estimate comparator PFS and OS outcomes over time for each tumour site. The CS stated that due to the absence of data for certain tumour sites, tumours such as colorectal and appendix cancer were grouped together. The ERG asked the company to clarify the

clinical rationale for the groupings made. The company stated that, following clinical advice, infantile fibrosarcoma (IFS, n=13) should be grouped with the paediatric STS patients because is a type of soft tissue sarcoma (STS) and because no relevant comparator data were identified in the review for IFS patients with advanced or metastatic disease (combined group size n=[redacted]). A similar rationale was provided for grouping bone sarcoma (n=2) with STS adults (non-GIST) with a combined size of n=[redacted], and appendix tumours (n=1) with colorectal cancers (combined size n=7). The ERG notes that the validity of these assumptions and groupings is unclear, given the lack of clinical evidence, but accepts it is the only plausible approach, given the stated lack of available evidence for these subgroups. The ERG notes that, given both the historical and broad nature of the comparator datasets (many patients will not have had NTRK gene fusions), it is likely that no patients would have received targeted experimental therapy – such as LOXO-195 – a bias which would favour larotrectinib.

To produce a summary survival curve the company combined comparator data across all tumour types, assuming a distribution of tumour types matching that in the ePAS2 cohort. The summary PFS and OS curves across all tumour sites are reproduced here (Figure 2). The company presented this curve for illustration only; it was not used in the economic analyses, nor to compare the clinical effectiveness of comparators to larotrectinib. The CS reported no formal comparison of these summary survival curves for the comparators with the larotrectinib data, either across all tumour sites or within tumour sites. Comparator curves for each individual tumour site were used separately in the economic modelling. The ERG notes that the validity of pooling across multiple tumour sites is highly uncertain. Different tumour sites have varying expected survival, so combining them in an average survival curve may not be meaningful.

Although the company's approach to selecting suitable comparator data and survival extrapolation was reasonable, the approach may still be sensitive to the choice of data, or the analysis method used. Patients in other trials may not have had NTRK fusion, and so might have a different prognosis to patients eligible for larotrectinib. The ERG considers that an analysis accounting for potential heterogeneity across tumour sites would have been more appropriate.

The ERG considers that the validity of the comparator data is uncertain, as it is drawn from arbitrarily selected past TAs, with numerous assumptions where past TAs do not exist. Given this uncertainty, the ERG did not attempt any detailed further analyses of these comparator data, either across tumour sites or within site. Some analyses of these data are considered in Section 4.6.

Figure 2 Pooled PFS and OS curves for the comparator data from past NICE TAs

4.4.2 **Responder/non-responder analysis**

The company considered an alternative approach to constructing a comparator dataset which was used in exploratory scenario analyses in the cost-effectiveness modelling. This approach used data from non-responders in the larotrectinib studies as a proxy for patients who did not receive larotrectinib. This had the advantage of using data from patients who definitely met the larotrectinib trial eligibility criteria, and so had NTRK fusions. However, the company identified key limitations of this approach as being the small number of patients available (n=30), the inability to balance groups in tumour sites, and the possibility of non-responders being inherently different to responders.

To investigate how this might impact on the relative effectiveness of larotrectinib the ERG compared the Kaplan-Meier data for responders and non-responders (provided in the economic model) to the ERG's version of the likely survival curves for the comparator data (Section 4.4.1). This is shown for PFS in and for OS in .

These figures show that PFS appears to be worse for non-responders than in the comparator data. This suggests that non-responders are not representative of people who do not receive larotrectinib, and may have a poorer prognosis. For OS the opposite is true: non-responders have better survival than the comparator data. This may be because non-responders are likely to receive further experimental treatment (e.g. LOXO-195), and so may have improved survival compared to best supportive care.

Therefore it is likely that using non-responders as a proxy for patients not receiving larotrectinib may lead to biased results. This may overestimate the effect of larotrectinib for PFS, but underestimate it for OS.

[redacted]

[redacted]

4.4.3 Growth Modulation Index (GMI) analysis

A third approach for investing the comparative effectiveness of larotrectinib, used as a sensitivity analysis in the CS, involved patients acting as their own control by using outcome data (ORR and time to progression) from the patient's previous line of therapy. As data on previous therapy was not available for all patients, the GMI analysis was limited to 53 patients. As the company pointed out, this analysis is restricted to changes in PFS from previous therapy to larotrectinib; by definition, OS data for the previous line of therapy cannot be used.

The IRC analysis (Table 24, page 96 of CS) found that [redacted] had lower PFS on larotrectinib (GMI < 1); [redacted] had PFS ratios between 1 and 1.33, and [redacted] had PFS at least 1.33 times better than previous line of therapy. 1.33 was a recommended cut-off for suggesting treatment benefit.

As data were not presented by tumour site the ERG requested a breakdown of GMI by site. This was not produced, but a figure presenting the results was supplied by the company. This is reproduced here as [redacted]. The ERG notes that this plot suggests a wide variation in GMI results, with

potentially substantial heterogeneity by tumour site. [redacted].

The ERG notes that GMI analysis may be unreliable in general, as it is based on a patient's previous, unsuccessful, line of therapy, which may not represent survival on best supportive care. The threshold of 1.33 to determine treatment benefit is also arbitrary.

[redacted]

4.5 Adverse events

Data on adverse events were derived from the safety analysis cohort of 137 patients who had received at least one dose of larotrectinib and were reported on pages 107-113 of the CS. The median time on treatment was [redacted] months.

4.5.1 Adverse events of any cause

The most frequent adverse events of any cause were an increase in ALT (34%) and AST (31%) (both enzymes, used to assess liver function), fatigue (31%), dizziness (29%), constipation (29%), cough (28%), nausea (28%), diarrhoea (27%) and anaemia (25%). Sixty-one patients (45%) had a grade 3 or grade 4 adverse event, with the most frequent being [redacted] Adverse events judged to be related to larotrectinib

Most patients (83%) had at least one treatment-emergent adverse event (TEAE) thought to be related to larotrectinib (Table 27 of the CS). Fourteen patients (10%) experienced a grade 3 or 4 TEAE thought to be related to larotrectinib; [redacted] in the conference abstract the company supplied to the ERG in response to a request for more up-to-date results.

Five patients (4%) had a TEAE which led to discontinuation of larotrectinib, one of which was thought to be related to larotrectinib. Adverse events that led to a dose reduction occurred in [redacted] patients.

Considering the stage of disease at which patients will receive larotrectinib, and their limited treatment alternatives, the ERG considers its safety profile to appear acceptable.

4.5.2 Adverse events in comparator trials

Table 26 of the CS presented data on SAEs and treatment-related SAEs for 10 different tumour types, by line of treatment. However, comparison with larotrectinib SEA data was difficult since no data were presented on grade 3 or 4 SAEs for comparator therapies.

4.6 Additional work on clinical effectiveness undertaken by the ERG

4.6.1 Modelling response rates across tumour sites

This section considers a Bayesian Hierarchical Modelling (BHM) framework to estimate the overall response rate which accounts for and explores the potential heterogeneity in effects across tumour sites, following the model of Thall et al.[40] The method estimates posterior probabilities of response for each tumour site, a pooled posterior probability of response across all tumour types, accounting for the potential lack of uniformity of effect across tumours, and a predictive distribution, predicting the possible response in an as yet untested tumour type.

For the response outcome, data available for each of the tumour types in the integrated efficacy analysis population, (ePAS2), are the number of responders, x_j, out of the total number of patients, n_j for tumour site j, which are assumed to follow a binomial likelihood

$$x_j \sim \text{Binomial}(n_j, p_j) \quad (1)$$

where p_j is the probability of response for tumour site j, with $j = 1,...,G$, and G is the total number of tumour sites. We model the log-odds of response in tumour site j, θ_j, on the log-odds scale:

$\text{logit}(\theta_j) = p_j$. The BHM assumes that for each of the G tumour types, the log-odds of response, θ_j, are exchangeable and follow a Normal distribution

$$\theta_j \sim \text{Normal}\left(\mu, \sigma^2\right) \quad (2)$$

where σ is the standard deviation quantifying the between-tumour heterogeneity and μ is the pooled mean effect across all sites. Prior distributions must be selected for μ and σ and are likely to have some influence on the posterior estimates,[40, 41] particularly when a small number of sites and patients per site are included.

The prior distribution for μ was centred around a probability of 0.3 (a log-odds of -1.3863), which was considered as promising response rate in the CS, with a variance of 10. The prior for the between-site heterogeneity variance was set as a uniform distribution from 0 to 5 (following Cunanan).[41]

We also calculated the probabilities that the response rate for each site is at least 30% or at least 10%.

Primary CNS tumours were not included in the main BHM analysis, given that these may represent a substantially different patient population. They were included in a sensitivity analysis.

Response data were extracted from the CS Appendix E, Table 82. The number of patients and responses by tumour type obtained are given in Table 7.

Table 7 Number of responders by tumour type (adapted from CS, Appendix E, Table 82)

Tumor Type	N	Responders
Overall	93	67
Soft tissue sarcoma	20	16
Salivary gland	17	15
Infantile fibrosarcoma	13	12
Thyroid	10	7
Lung	7	5
Melanoma	7	3
Colon	6	2
GIST	5	5
Bone sarcoma	2	1
Cholangiocarcinoma	2	0
Appendix	1	0

Tumor Type	N	Responders
Breast	1	0
Congenital mesoblastic nephroma	1	1
Pancreas	1	0

4.6.1.1 Results

Figure 3 shows the predicted distributions of response by tumour site from the BHM. Tumour sites are categorised by frequency of NTRK fusion, to aid interpretation.

Figure 4 presents the same results in a forest plot. The predicted probabilities of response rates exceeding 10% and 30% are given in Table 8.

Figure 3 Predicted response rate distributions from BHM

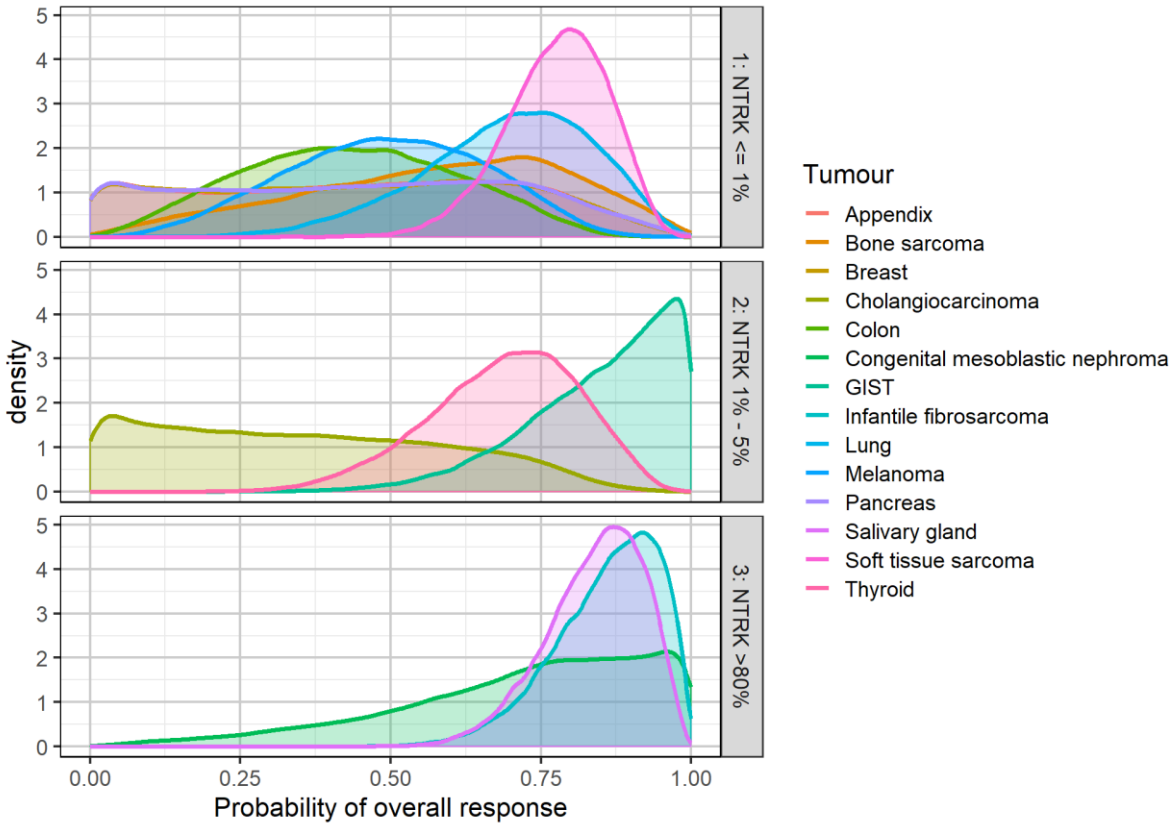
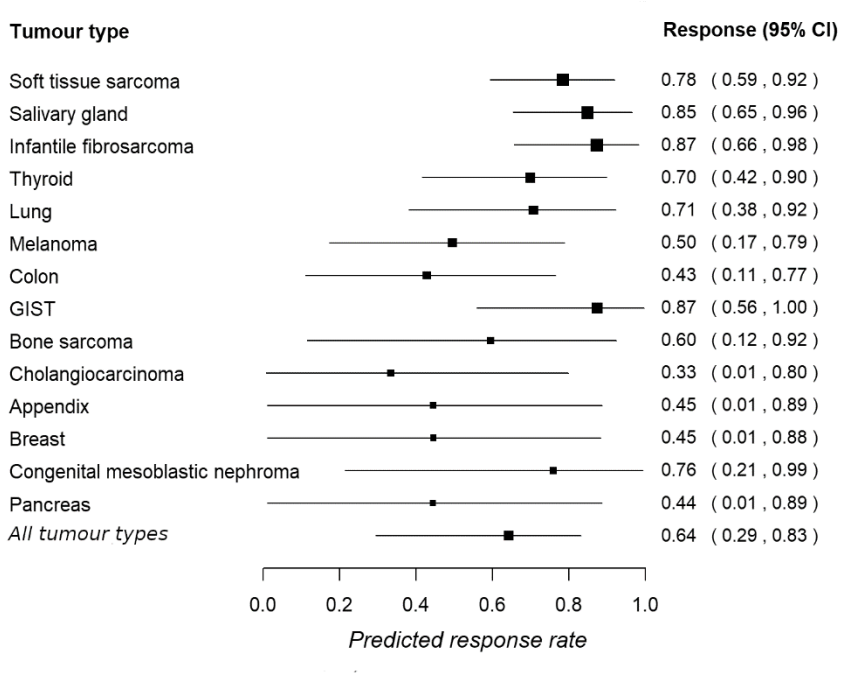


Table 8 Probabilities that response rate exceeds 10% or 30% from BHM

Tumour	Probability response rate exceeds:	
	30%	10%

Soft tissue sarcoma	100	100
Salivary gland	100	100
Infantile fibrosarcoma	100	100
Thyroid	99.8	100
Lung	99.3	100
Melanoma	86.7	99.6
Colon	74.8	98
GIST	100	100
Bone sarcoma	86.4	98
Cholangiocarcinoma	54.2	81.9
Appendix	66.2	86.8
Breast	66.1	87
Congenital mesoblastic nephroma	95	99.4
Pancreas	65.7	86.7

Figure 4 Predicted mean response rates from BHM



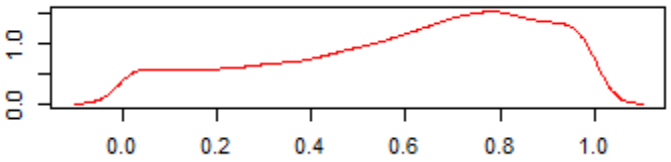
The model gave an estimated overall response rate (ORR) across all tumour sites of 64% (95% CrI 29 to 83). This is lower than the estimated response rate of 72% presented in the CS, because the BHM accounts for across-site heterogeneity. The ERG considers that this 64% ORR represents a more realistic estimate of the response to larotrectinib across all patients than the CS analysis.

The BHM model shows clear evidence of heterogeneity in response across tumour sites (heterogeneity estimate: 1.58, 95% CrI 0.38 to 3.64). Some sites, most notably IFS, salivary gland, STS and GIST, have very high response rates, and a 100% probability that response rates exceed

30%. Other sites, particularly those with limited data (including colon, pancreas, appendix and breast), have predicted response rates below 50%, and substantial probability that the true response rate is lower than 30%, or even 10%.

The predicted response distribution for an as-yet unevaluated tumour site is shown in Figure 5. This shows a broad range of possible response rates in a new tumour site, with some probability that the response is below 30% (18% chance), or even below 10% (7% chance).

Figure 5 Predicted response for an unevaluated tumour site



4.6.1.2 Primary CNS tumours

Adding Primary CNS tumours (1 response from 9 patients) to the BHM has some impact on the results. The estimated ORR for Primary CNS tumours is 17% (95% CrI 2 to 48), with a 73% probability of having a response below 30% and an 18% probability that the response rate is below 10%. The overall estimated ORR drops to 57% (95% CrI 23 to 80) when primary CNS tumours are included.

4.6.1.3 Regression models

Given the evidence from the BHM we investigated further whether response rate was associated with frequency of NTRK fusion. To do this a logistic regression of response against the logarithm of fusion frequency (from Table 2) was carried out. Also included in the model were whether the tumour site occurred in adults only or in children, and the logarithm of median survival without treatment [from the CS analysis, assuming exponential survival distributions for all tumour sites, data in the supplied economic model]. Random intercepts by tumour sites were included to account for heterogeneity across tumour sites.

The model found strong evidence that response rate increases with frequency of NTRK fusion (model coefficient: 0.36 per log(percent), p = 0.0084); a possibility that response rates are higher in children (coefficient 1.30, p = 0.082); and no evidence that expected survival time has any impact (coefficient: -0.29, p = 0.383). There was zero residual heterogeneity, suggesting that NTRK fusion rate and differences between adults and children might explain all the observed heterogeneity in response rate.

4.6.2 Exploring heterogeneity in time-to-event outcomes across tumour sites

Heterogeneity in time to event outcomes (PFS, OS) could potentially be explored using the BHM in a similar way.[40] The ERG had intended to conduct these analyses; however, the company did not supply PFS and OS data by tumour site, so this analysis could not be carried out.

In order to investigate possible heterogeneity in PFS and OS an alternative approach was taken. The company submission included full Kaplan-Meir curve data on responders and non-responders as part of the analysis of responders/non-responders in the CS. The ERG reconstructed full survival data from the supplied data using the Guyot method.[42] This gave full PFS and OS data according to response.

From this data we linked survival times to tumour site by sampling from the reconstructed data, without replacement, so that responders and non-responders by tumour site matched the data in Table 7. This produced a complete, possible, data set of PFS and OS by tumour site.

4.6.2.1 Progression-free survival

The Kaplan Meier curve for PFS, stratified by tumour site for one such sample data set is presented in Figure 6. Although this is only a simulated sample, it suggests substantial heterogeneity in progression-free survival, including median survival and long-term survival proportions.

Figure 6 Sampled PFS Kaplan-Meier curve, stratified by tumour site

For this simulated sample an exponential survival model was fitted separately for each tumour site to estimate the median survival time. An exponential model was chosen for simplicity, given the sparse data, and as the CS found it to be a reasonable fit to the complete data.

This process of sampling without replacement from the reconstructed survival data, and fitting exponential models to the sample was repeated 1000 times, to generate a bootstrap sample of

plausible median survival times for each tumour site. These bootstrapped distributions of median PFS are shown in Figure 7.

Figure 7 Distributions of plausible median PFS by tumour site

[redacted]

Although these results are highly uncertain, with wide distribution ranges, there is clear heterogeneity.

[redacted]

The company submission estimated median PFS and OS by tumour site for patients not receiving larotrectinib by fitting exponential models to data from past STA assessments. Taking the difference between these estimated median survival times and those estimated for larotrectinib patients in Figure 7 we can predict the possible benefit of larotrectinib by tumour site. This is shown for PFS in

[redacted]

Figure 8.

[redacted]

Figure 8 Potential improvement in PFS for larotrectinib versus comparator, by tumour site

[redacted]

4.6.2.2 Overall survival

The ERG performed similar analyses for OS. These are presented here for completeness, but the small number of deaths (particularly in responders) makes those analyses more difficult to interpret.

The predicted median OS times by tumour site are given in

[redacted]

Figure 9 and the predicted improvement in OS compared to patients not receiving larotrectinib (calculated using the same approach as for PFS) is shown in Figure 10.

The results are much more dispersed than for OS due to the very few deaths in the larotrectinib trials. The pattern of results is broadly similar to those for PFS.

[redacted]

Figure 9 Distributions of plausible median OS by tumour site

[redacted]

Figure 10 Potential improvement in OS for larotrectinib versus comparator, by tumour site

4.7 Conclusions of the clinical effectiveness section

4.7.1 Systematic review and trials of larotrectinib

Most of the systematic review methods used by the company to identify larotrectinib studies were appropriate and the recruited patients appear to be representative of patients who might be eligible for larotrectinib in the NHS.

Overall the data on larotrectinib is limited, with only three trials, all of which are phase I or II, and only 93 patients in the main ePAS2 analysis set. The data are further limited by being spread across both adults and children, and multiple tumour sites: the largest number of patients in any one site is 21 (soft tissue sarcomas), and the largest number of adult patients in any one site was 17 (salivary gland).

The data may have sufficient statistical power to detect an overall response rate (ORR) of at least 30%; however, the larotrectinib trials were not designed, nor sufficiently powered, to detect response in any individual tumour site, or to test the assumption of heterogeneity of response across subgroups. All three trials were single-arm, limiting the extent to which the effectiveness of larotrectinib could be compared to best supportive care, or alternative therapies.

All analyses in the CS assumed that larotrectinib had the same efficacy in all tumour sites or patient subgroups. The company justified this on the basis that larotrectinib is intended to be site-agnostic, there was no evidence of heterogeneity and the data were too limited to make subgroup analysis reasonable. The ERG disagrees with this justification, considers it inappropriate to assume a common response rate independent of tumour histology, and therefore finds this to be a limitation of the submission. In particular, the ERG notes that lack of data is not sufficient grounds to ignore analyses of heterogeneity, still less grounds to assume homogeneity. Furthermore, the division between the submission's efficacy datasets ('ePAS2', and 'SAS') appears quite arbitrary and was not clearly justified in the submission. It was also unclear why the safety analysis set (i.e. the modified intention-to-treat dataset) was not used for analysing PFS and OS, although the two datasets yielded very similar results (albeit with immature data).

The company declined to provide PFS and OS results for subgroups which the ERG considered might be important for investigating possible treatment effect heterogeneity. From the data which *were* provided this heterogeneity was most clearly evident when comparing the ORR for patients with primary CNS tumours with the ORR for the ePAS2 cohort. Moreover, the observation that [redacted] had surgery following a partial response and that these patients had either [redacted] also suggests that heterogeneity in PFS and OS across tumours sites is likely.

[redacted] post-progression patients received an experimental therapy called 'LOXO-195' which was developed for treating patients who become resistant to TRK inhibitors, and which is produced by the manufacturer of larotrectinib. Whilst the ERG acknowledges that some resistant patients would receive experimental treatments in the NHS, the number of patients specifically receiving LOXO-195 (as opposed to other experimental treatments) suggests that the impact of LOXO-195 on survival after progression following treatment with larotrectinib should not be ignored (as was suggested by the company). Furthermore, [redacted] patients continued to receive larotrectinib post-progression, which would not happen in clinical practice.

4.7.2 **Indirect comparisons**

All larotrectinib trials were single arm, so direct comparison of larotrectinib to other therapies (or best supportive care) was not possible. The CS considered three indirect comparison approaches, all of

which had limitations, as acknowledged in the CS. Indirect comparisons were used to inform the cost-effectiveness analyses, but clinical effectiveness comparisons were not reported.

Although the company undertook a comprehensive systematic review of comparator therapies, the submission did not present how the comparator baseline characteristics or outcome data for individual tumour sites compared with the corresponding data for larotrectinib, so it was not possible to evaluate the appropriateness of the comparator data sets selected (in the systematic review). Given both the historical and broad nature of the comparator datasets (many patients will not have had NTRK gene fusions), it is unclear whether these patients can be considered comparable with those in the larotrectinib trials. In particular, it is likely that no or very few patients who progressed would have received targeted experimental therapy – such as LOXO-195 – a bias which would favour larotrectinib.

The use of non-responders as a proxy for people not receiving larotrectinib has the advantage of using observed data on patients with NTRK fusion. However, non-responders may have a different prognosis from people not receiving larotrectinib. The ERG analysis found that PFS among no-responders appeared worse than in the comparison based on past NICE TAs, suggesting a worse prognosis. By contrast OS was better for non-responders, which may be a consequence of non-responders being recruited into the trial of LOXO-195. Hence the responder/non-responder analysis may give biased results.

The ERG considers the GMI analysis based on past line of therapy to be of doubtful value. It was based on an incomplete data set, with an arbitrary cut-off for determining efficacy. Survival on previous lines of therapy may not be a good proxy for later survival (e.g. on best supportive care), because patients transfer to larotrectinib precisely because the previous therapy was ineffective.

4.7.3 **Analyses conducted by the ERG**

As stated above the ERG did not agree with the company that heterogeneity across tumour sites or subgroups should not be analysed. The ERG performed a Bayesian analysis to analyse ORR, which accounted for possible heterogeneity by tumour site. The analysis concluded that the best estimate of ORR was 64% across all tumour sites; lower than the 72% reported in the CS. The ERG therefore considers that the CS overestimates the response to larotrectinib. If primary CNS tumours were included the ORR dropped further, to 57%

The ERG's analysis found good evidence of heterogeneity in ORR across sites. Specifically, tumour sites with high NTRK fusion prevalence (MASC, IFS, GIST, Thyroid) all had high ORR, with near-zero probability that ORR was below 30%. Other sites, particularly some with low NTRK fusion

prevalence (e.g. Appendix, Breast, Melanoma, Pancreas), had substantial probability that the true ORR was less than 30%.

The ERG requested data on PFS and OS by tumour site, but this was not supplied. Analysis by the ERG suggested that the heterogeneity in ORR by tumour site could lead to substantial heterogeneity in PFS and OS by tumour site. This heterogeneity was sufficient to suggest that larotrectinib may not be more effective than best supportive care in some tumour sites. The ERG notes that this analysis was speculative only, as the required data were not available.

4.7.4 **Identifying NTRK fusions**

The CS did not discuss genetic testing to identify patients with NTRK fusions who would be eligible for larotrectinib, on the grounds that including costs of diagnostic tests in the economic model was not required. The ERG disagrees with this (see 5.2.8.5). The ERG also notes that, because only patients with NTRK fusion can benefit from larotrectinib, considering the clinical impact of genetic testing is also important. The ERG notes that genetic testing specifically for NTRK fusion may differ from more general genetic/genomic screening; for example, by requiring extra gene panels.

The impact of genetic testing depends on the prevalence of NTRK fusions. Tumours where NTRK fusion is common, are also generally cases with small numbers of patients (e.g. MASC, IFS). In these tumour types genetic testing specifically for NTRK fusion is unlikely to place a burden on the health service (and may already be in place for some tumour types). Conversely, in tumour sites with low rates of NTRK fusion the numbers needed to screen to identify each NTRK fusion cancer may be considerable (see Table 2), which may put a considerable extra screening burden on the NHS, particularly in cancers which are common and genetic screening is not widely used at present.

The ERG also considers that the diagnostic accuracy of genetic testing to identify NTRK fusions should be considered. Even a near-perfect test will still lead to some false results, and there appears to be little current research, and considerable uncertainty, on the diagnostic accuracy of NTRK fusion testing. This may not be of concern for tumour sites with a high NTRK fusion prevalence, where errors will be small in number. However, for tumour sites with low NTRK fusion prevalence the number of false positives (people who test positive for fusion despite not having one) may substantially outnumber those with genuine NTRK fusions (see Table 4). Hence, even if larotrectinib is effective in people with NTRK fusions the observed response rate could be low because of the large number of false positive cases where larotrectinib cannot work.

4.7.5 **Summary**

The ERG notes several broad concerns with the clinical effectiveness data and analyses. Overall there was very limited data on larotrectinib, with no direct comparison to other treatments or best supportive care. This lack of data was made more critical by the diversity of patients in tumour site and age.

Most critically, the ERG's analyses found evidence that ORR varied by tumour site, and therefore the ERG rejects the company's assertion that the effectiveness of larotrectinib can be assumed to be the same for all patients. Because heterogeneity was not considered in any of the analyses in the CS the ERG considers the findings in the CS to be unsound, and may not adequately reflect the efficacy of larotrectinib across different tumour sites, particularly in terms of progression-free and overall survival.

The ERG found that evidence for the effectiveness of larotrectinib was strongest in tumour sites with higher NTRK fusion prevalence, particularly salivary gland and IFS, and also possibly thyroid cancer and GIST. All have estimated response rates of 70% or more. Because NTRK fusion is common in these tumours using genetic screening may be practical, and would have few incorrect results. They may also represent the bulk of detectable tumours in any year.

By contrast there were several tumour sites, particularly those with low NTRK fusion prevalence, where it remains unclear whether larotrectinib is effective, and there is a reasonable probability that the ORR is below 30%. This may be due to chance, because there were few patients in these tumour sites; it may represent genuine heterogeneity in efficacy; or it may be that some patients in the NAVIGATE trial were false positives, and did not have NTRK fusion. In tumour sites where NTRK fusion is rare the screening burden is high, with large numbers needed to screen to detect each genuine NTRK fusion cancer. The potentially large number of false positive genetic tests may also reduce the effectiveness of larotrectinib in these tumour types.

5 Cost Effectiveness

5.1 ERG comment on company's review of cost-effectiveness evidence

The company submission contained brief details of two searches undertaken to identify cost-effectiveness analyses in Section B.3.1, p. 135. The first search was to identify cost-effectiveness analyses of treatments for patients with TRK-fusion cancer, with search strategies and sources reported in Appendix G.1. The second search was to identify cost-effectiveness evidence of treatments for patients with solid tumours that are known to harbour NTRK gene fusions. Search sources were reported in Appendix G.2, however search strategies were not provided. The company provided the search strategies in Appendix 2 of their response to the points for clarification raised by the ERG.

5.1.1 Searches for cost-effective analyses of treatments for patients with TRK-Fusion cancer

The following databases were searched on 5th May 2019: MEDLINE, MEDLINE in process, EMBASE, EconLit and Northern Lights Life Sciences Conference Abstracts (all searched via Proquest Dialog), and the Cochrane Library via Wiley. The company reported that searches of the Database of Abstracts of Reviews of Effects (DARE), the NHS Economic Evaluations Database (NHS EED) and the Health Technology Assessment (HTA) database were undertaken via the Cochrane Library, however these databases were removed from the Cochrane Library in August 2018. The two databases available via the Cochrane Library at the time that the searches were carried out were The Cochrane Controlled Register of Trials (CENTRAL) and the Cochrane Database of Systematic Reviews (CDSR).

In general, the methods used and sources searched to identify both published and unpublished studies were appropriate, and the reporting of the searches was mostly clear.

The search strategy for MEDLINE, EMBASE, EconLit and Northern Lights Conference Abstracts and the strategy for the Cochrane Library consisted of a set of terms relating to cost-effectiveness combined with terms for the population, TRK fusion-positive tumours. A reasonable variety of terms and subject headings relating to cost-effectiveness were included in the strategy. However only one term was included for TRK fusion, the abbreviated term *TRK* with left and right hand truncation. Although this may have identified relevant studies, a comprehensive approach would have been more appropriate here, including all possible terms for TRK fusion such as neurotrophic tropomyosin receptor kinase, tropomyosin receptor kinase, tyrosine receptor kinase and variations of these terms, to reflect the variety of ways they are described in the literature.

5.1.2 Searches for cost-effectiveness evidence of treatments for patients with solid tumours that are known to harbour NTRK gene fusions

The following databases were searched: MEDLINE (via PubMed), EMBASE, and the Cochrane Library. The interface/provider was not reported for EMBASE or the Cochrane Library. All searches took place between May – August 2018 and were updated during January - March 2019. The searches of the Cochrane Library in August 2018 and during January - March 2019 would not have included searches of the Database of Abstracts of Reviews of Effects (DARE), the NHS Economic Evaluations Database (NHS EED), or the Health Technology Assessment (HTA) database, although these databases are listed as being searched on page 325, Appendix G.2 of the submission. These three databases were removed from the Cochrane Library in August 2018. The Cochrane Controlled Register of Trials (CENTRAL) and the Cochrane Database of Systematic Reviews (CDSR) would have been searched via the Cochrane Library. Date limits were applied to some of the search strategies. The limits for human studies and specific publication types were applied to the EMBASE searches. No other limits were applied to the searches. A date limit was applied during study inclusion/exclusion phase and was applied to NSCLC only. It restricted retrieval of articles to those published from 2008 onwards and retrieval of conference abstracts to those from 2017 onward.In addition to the database searches, a comprehensive set of sources were searched for unpublished, grey literature. Five clinical trials registers were searched (Clinicaltrials.gov, ISRCTN register, International Clinical Trials Registry Platform (ICTRP), EU Clinical Trials Register and KlinischePrufungen (PhamNet.Bund, AMIS – Offentilicher Teil)). The following Conference proceedings were searched: American Society of Clinical Oncology (ASCO) and the European Society of Medical Oncology (ESMO) and the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) International and European Annual Meetings. Searches of several key international HTA websites were undertaken. A review of NICE technology appraisals related to oncology treatments based on single arm trial data, and those that have previously considered multiple histologies was also conducted. Further details on how the searches for unpublished, grey literature were carried out, the search terms used, the date of the search or any limits applied, were not reported.

14 sets of database searches were undertaken in total, one for each of the tumour sites/locations. The searches were structured appropriately and matched the inclusion criteria specified in tables 89 and 90 (p.327-31) of the submission. Most of the searches included terms for the population combined with a set of terms to limit to economic evaluations. In some of the search strategies, terms for the population were limited to advanced forms of the particular cancer/tumour (line #1 in economic evaluation strategies for non-small cell lung cancer, colorectal cancer, melanoma, pancreatic cancer, glioma, biliary cancer, gastrointestinal stromal tumours, bone sarcoma and appendix cancer).

All searches were conducted appropriately, search lines were combined correctly and no major errors were found by the ERG. Some of the update searches carried out in January - March 2019 in EMBASE and MEDLINE were limited to articles published in 2018 or 2019. Therefore, any relevant studies added to the databases since the last search but with a publication year pre-2018, may not have been identified by the searches.

5.1.3 Inclusion/exclusion criteria used for study selection

The inclusion/exclusion criteria for the two cost-effectiveness reviews are summarised in Table 88, 89, and 90 (Appendix G.1 and G.2) of the CS, and follow the PICOS framework.

The first review considered all studies on treatments targeting NTRK-positive solid tumours. No restriction was placed on outcomes or type of study design, with the exception of publications focusing solely on screening for NTRK rather than treatment.

For the second review, the company identifies site-specific inclusion criteria for population and comparators. In brief, studies were considered relevant for inclusion it they recruited patients with advanced or metastatic cancer who had failed previous therapies (generally two lines). A broad set of standard costs and HRQoL outcomes, as well as study designs, were considered relevant. Letters, editorials, case studies and non-systematic reviews were excluded. Finally, the review focused on studies conducted in the US, Canada, Japan, Brazil, and five European States (UK, France, Germany, Italy, Spain), published from 2008 onwards (conference abstracts were considered if published after 2017).

Articles were independently assessed by two reviewers against each eligibility criteria. Any uncertainty regarding the inclusion of studies were checked and judged by a third reviewer.

The ERG considers that the inclusion/exclusion criteria for both reviews appear to be appropriate.

5.1.4 Studies included and excluded in the cost effectiveness review

The company did not identify any published cost-effectiveness studies for the first search, i.e. for studies on the treatment of NTRK fusion-positive cancer.

For the second search, the results by tumour site and by endpoint of interest are presented in both tabular format and in narrative in Appendix 2 (company's response to points for clarification). The BMJ Study Checklist for Economic Studies [43] was used to perform the quality assessment of cost-effectiveness studies with results of the quality assessment presented in separate Excel files. In the main CS (p136), the company states that 98 studies were identified across all tumour sites, and that

these *"were informative for assessment of the model structure, and assumptions used in model development"*.

In Appendix 2 (company's response to points for clarification) results are also presented for two other systematic reviews (i. HRQoL and health state utilities, and ii. resource use and cost) in a format similar to that of the cost-effectiveness systematic review by tumour site, but without a quality assessment.

5.1.5 Conclusions of the cost effectiveness review

In the absence of any previously published cost effectiveness studies in patients with *NTRK* fusion–positive solid tumours, the *de novo* analysis in the CS represents the most relevant evidence for the stated decision problem.

5.2 ERG's summary and critique of company's submitted economic evaluation

5.2.1 Model structure

The company presents a *de novo* cost-effectiveness analysis. Table 9 presents an overview of the company's economic evaluation with justifications for key aspects and signposts to the relevant sections of the CS. The ERG has considered the methods applied in the company's economic evaluation in the context of a detailed checklist, reported in Appendix 10.1.

Table 9 Overview of the company's economic evaluation

	Approach	Source / Justification	Location in CS
Model	Cost-effectiveness (cost-utility) analysis using a cohort state transition model with a partitioned survival approach.	A partition survival model is justified based on the basis that this approach is commonly used in oncology modelling to capture the progressive nature of this condition with outcomes requiring an ongoing, time-dependent risk. The company notes the lack of precedence for modelling histology independent treatments, and states that their model methodology is in line with the NICE Reference Case.	Section B.3.2; p139, p141-145
States and events	The model comprises three mutually exclusive health states: 1) Progression-free 2) Progressed 3) Death	Patients start the model in the progression-free heath state and can remain in this state or transition to i) progressed or ii) death. Patients with progressed disease can only remain in this health state or transition to death. Death is an absorbent health state.	Section B.3.2; p139-140; p144

Comparators	Larotrectinib was compared to
	• Base-case: last-line standard of care, with the exact comparator varying by tumour site.
	• Scenario analyses:
	  • Non-responders to Larotrectinib in the integrated efficacy analysis
	  • Previous line of therapy for patients in Larotrectinib integrated efficacy analysis
		The clinical studies are single arm trials, and therefore the company generates a comparator group using 3 approaches. In the base case the comparator is a weighted average of the cost effectiveness estimates of 12 model engines. Each engine assesses cost effectiveness on a separate tumour site. The tumour sites included and tumour site weightings reflect the distribution of tumour sites in the integrated efficacy analysis population (30th July cut off) from the larotrectinib integrated efficacy analysis. Stratification by tumour site is considered by the company to i) account for differences in conventional standards of care, quality of life, costs and resource use across tumour sites, ii) improve transparency by presenting disaggregated results by tumour site, and iii) to allow the use of alternative sources of evidence.
		In the historical comparator, the comparators by tumour site were (base case):
		• NSCLC: BSC
		• Salivary gland cancer: cisplatin + vinorelbine
		• Melanoma: Mixed chemotherapy including dacarbazine, paclitaxel, carboplatin, temozolomide and paclitaxel + carboplatin
		• Colorectal and appendix cancer: BSC
		• Adult soft tissue sarcoma (GIST): BSC
		• Adult soft tissue sarcoma (non GIST): BSC (historical control data)
		• Soft tissue sarcoma (paediatric), infantile fibrosarcoma, and congenital mesoblastic nephroma: irinotecan + vincristine
			Section B.2.6: p92-94
			Section B.3.2; p138, p148-155
			Section B.3.8; p213-217
			Appendix M; p468-575
			Appendix Q: p608

	Approach	Source / Justification	Location in CS
		• Breast cancer: treatment of physician's choice including vinorelbine, gemcitabine, paclitaxel, doxorubicin and docetaxel. • Cholangiosarcoma: gemcitabine + cisplatin • Pancreatic cancer: 5-FU + leucovorin • Gliomas (CNS): lomustine and procarbazine, lomustine & vincristine (PVC) • Thyroid papillary/follicular cancer: BSC The company considered tumour site-specific comparators to reflect current management without larotrectinib, as current practice does not involve specific treatment for TRK-Fusion cancer, and treatment is based on histology and stage of disease. Only one comparator was considered by tumour site (although some are considered as a blended comparator), to Two other alternative approaches to modelling comparator data were explored in scenario analyses. The first approach consisted of using effectiveness data from the non-responders in the larotrectinib integrated efficacy analysis as a proxy for patients not receiving an active treatment. This assumes that non responding patients are not exposed to a treatment effect (as they have no response to treatment). The second approach consisted of comparing the outcomes (in terms of PFS and ORR) of patients in the larotrectinib trials against their outcomes (in terms of time to progression and ORR) while on their most recent previous line of treatment. The alternative comparator approaches only explored different assumptions in terms of PFS and OS for the comparator. Costs and HRQoL for the comparator were the same as base case for these two approaches.	
Natural History	Based on partitioned survival model. Transitions between states were based on survival estimates derived from published data (for the comparator arm). Survival data was modelled separately by tumour site for the comparator arm.	PFS and OS estimates were modelled independently, with the proportion of progressed patients at each cycle, calculated as the difference between the OS and PFS curves.	Section B.3.2; p150 Section B.3.3 p156-166
Treatment effectiveness	Clinical outcomes included PFS and OS.	Treatment effectiveness in the base-case analysis was taken from an uncontrolled comparison between the extrapolated PFS and OS outcomes larotrectinib vs the weighted comparator. The survival data for larotrectinib was sourced from the pooled larotrectinib clinical trial data. For those comparators that were informed by a previous NICE technology appraisal, the company attempted to model survival	Section B.3.3; p159-166 Section B 3.6: Section B.3.8; p213-217

	Approach	Source / Justification	Location in CS
	Larotrectinib OS and PFS were extrapolated from the larotrectinib integrated efficacy analysis. Comparator OS and PFS was extrapolated based on data from published data including previous NICE Technology Appraisals for tumour sites where these were available or identified via systematic reviews.	data as closely to the appraisal committee's preferred approach as possible when generating PFS and OS extrapolated curves. Comparators informed from published data identified in systematic reviews, OS and PFS Kaplan Meier curves were digitised and IPD recreated, to allow for fitting of parametric distributions to extrapolated OS and PFS. For the scenario analysis comparing the larotrectinib treatment arm versus non-responders from the larotrectinib integrated efficacy analysis, the comparator OS and PFS curves were extrapolated from observed survival data for non-responders. The survival models for non-responders were estimated in the full integrated efficacy analysis dataset (30th July data cut) with response status included as a variable. For the scenario analysis where larotrectinib is compared against a previous line of therapy, the comparator average patient's PFS when treated with larotrectinib is compared with the average patient's time-to-treatment progression (TTP) on their prior therapy. The ratio between average TTP for the previous line of treatment and the mean extrapolated PFS for the same patients treated with larotrectinib, the GMI multiplier, was estimated and applied to the OS and PFS curves for larotrectinib to derive comparator OS and PFS curves.	Appendix L; p451-467
HRQoL	Health state utilities were derived from EQ-5D and PedsQL data collected in the larotrectinib clinical trial programme. A different set of health state utilities was applied by tumour site for the comparator, and was informed by previous NICE Technology Appraisals and targeted literature searches. Utility decrements for adverse events were included and tumour site specific. The comparator adverse events disutilities were informed by previous NICE Technology Appraisals and targeted literature searches. Utility decrements for adverse reactions for larotrectinib were assumed to be the maximum disutility for the event across all tumour sites.	Health related quality-of-life was collected in two of the larotrectinib single arm trials: • LOXO-TRK-15002 (patients aged 12 and older) – collected EQ-5D-5L scores at baseline and every 8 weeks during the first year of follow-up; • LOXO-TRK-15003 (patients aged 1 month to 21 years) – collected PedsQL Infant Scale (for infants aged 1-24 months) and the PedsQL Generic Core Scales (for children aged>2 years) scores during pre-treatment screening, and then on the first day of every 28-day cycle, until treatment discontinuation. The EQ-5D-5L scores collected in LOXO-TRK-15002 were mapped to EQ-5D-3L using the crosswalk recommended by NICE. The PedsQL Generic Core Scales scores collected in LOXO-TRK-15003 were mapped to EQ-5D-3L using a published algorithm. Patients with PedsQL Infant Scale scores were excluded from the analysis due to the inexistence of a mapping algorithm that allowed generating EQ-5D-3L estimates. Health state utility values for patients treated with larotrectinib were estimated using the mapped EQ-5D-3L data for patients that had at least one measurement in the original HRQoL instrument, and using a Mixed	Section B.3.4; p169-177 Appendix H; p337-348 Company's response to Points for Clarification; p83

	Approach	Source / Justification	Location in CS
		Model Repeated Measures model that accounts for autocorrelation and repeated measurement of utility values. For the comparator health state utilities, tumour site specific estimates were obtained from previous NICE Technology Appraisals (NICE Appraisal Committee preferred assumptions) for the tumour sites that had appraisals, i.e. NSCLC, melanoma, colorectal, GIST, adult soft tissue sarcoma (nGIST) (also used as proxy for bone sarcoma), breast, CNS/glioma, pancreas and thyroid. Targeted literature searches were conducted to inform the health state utilities for the remaining tumour sites. In the absence of published data on the utility estimates for cholangiocarcinoma, these patients were assigned the weighted average of health state utilities for the other tumour sites. Since no utility data in a paediatric population was identified for STS, it was assumed that utilities were independent of age, and utilities estimated in an adult population where applied for both adult and paediatric STS. The company conducted three scenario analyses, varying the assumptions around the health state utilities for larotrectinib: 1. Utilities estimated from patients aged ≥11 years old 2. Same utilities as for historical comparator 3. Progressed disease utility estimated by applying the ratio between progression free and progressed disease utilities for the weighted comparator to larotrectinib's progression free utility value. Utility decrements were applied for Grade 3 or 4 adverse events that were reported in at least 5% or more of larotrectinib patients or the historical comparator arm. Tumour site specific disutilities were sourced from previous NICE technology appraisals and systematic reviews. Disutilities weighted by frequency of adverse events were applied as a one-off QALY loss at model entry to all patients. As disutility weights were not adjusted for duration of events, it was implicitly assumed that the disutility was incurred for a one year period.	
Adverse events	Adverse events were included if they were grade 3-4 treatment related AEs occurring in ≥5% of subjects in for intervention and comparators.	The 5% threshold rate is a common assumption used in NICE Technology Appraisals.	Section B.3.3 p167-168

	Approach	Source / Justification	Location in CS
		The adverse event rates for larotrectinib were sourced from the larotrectinib clinical trial programme safety population (n=137), while for the comparators they were taken from the respective sources that informed clinical efficacy. The company considered that the use of the 5% criterion by tumour site might bias the rates estimated. The AE rates from the comparator sources were subsequently reweighted using the tumour distribution from the larotrectinib clinical trial programme. Therefore, only AEs with a final weighted rate of >=5% were included in the final model calculations. A scenario analysis was included where the inclusion of AEs was based on unweighted rates.	
Resource use and costs	Resource use and cost categories included: • Drug acquisition • Administration • Health State • AE	Drug acquisition unit costs for the comparator treatments were sourced from electronic market information tool (eMIT) and British National formulary (BNF). Dosage and posology were sourced from previous NICE technological appraisal. The least expensive cost per mg was used to represent the unit cost, and wastage was not considered for the weighted comparator. Expected list prices for larotrectinib presentations were applied to estimate drug acquisition costs. Comparators administered orally and larotrectinib were assumed to have no administration costs. These were instead included for those comparators requiring IV administration. Drug acquisition costs for larotrectinib are applied at the start of the treatment cycle, with no half-cycle correction to account for treatment wastage. All other are stated to be applied a half-cycle correction over a 7 days model cycle. The duration of larotrectinib treatment was assumed to be until progression (and varied on scenario analysis to reflect treatment duration as in the clinical studies). Health care resource use and costs for the comparator was modelled separately by tumour site and sourced from previous NICE Technology Appraisals, when available, or from other published sources. As no sources were identified for cholangiocarcinoma, its costs were based on a weighted average of the costs of other tumour sites. Health state costs for larotrectinib were assumed equal to the weighted average of the comparators costs, using the tumour site distribution in the larotrectinib trial as weights. This assumption was required given the lack of resource use data for larotrectinib. Unit costs were sourced from NHS Reference costs 2017-2018. The unit cost of adverse events (AE) was assumed the same irrespective of tumour site, as per previous NICE Technology Appraisals. Unit costs were sourced from	Section B.3.2; p146 Section B.3.5; p178-191 Appendix I; p349-360 Appendix M; p468-575

	Approach	**Source / Justification**	**Location in CS**
		NHS Reference costs 2017-2018 using HRG codes for the particular AE. For AEs were HRGs were not available, they were assumed equivalent to a similar AE. All unit cost were presented in current value and inflated where necessary to 2017/18 pound sterling.	
Discount rates	Costs and benefits were discounted at 3.5% per annum	In accordance with the NICE reference case.	Section B.3.6; p192
Sensitivity analysis	Probabilistic sensitivity analysis was performed. Deterministic univariate probabilistic analysis was performed on a series of model parameters. A series of scenario analyses was also performed.	In accordance with the NICE reference case.	B.3.6; p196-197 B.3.8; p199-220

The company's cost-effectiveness analysis is based on a partitioned survival cohort model. The model comprises three mutually exclusive health states: progression free, progressed and death. The cycle length was 7 days. Patients enter the model in the progression-free heath state, where they can remain or transition to i) progressed or ii) death. Patients with progressed disease can only remain in the progressed health state or transition to death. The model structure is illustrated in Figure 11.

Figure 11 Model structure (CS, p140)

The model uses PFS and OS data and links these to utilities and costs. PFS and OS are modelled independently due to the partitioned survival modelling approach and directly inform the state membership for the 'Progression free' and 'Death' states over time, respectively. The difference between PFS and OS allows the proportion of patients in the progressed health state to be estimated.

The company performed a series of reviews (see section 5.1) to inform the model structure and parameters for the model. In the absence of studies modelling histology independent treatments, the company chose to use a partitioned survival approach as it is a commonly used modelling approach in oncology.

Larotrectinib's PFS and OS extrapolated curves were derived from observed survival data in the integrated primary analysis (data cut 30th July 2018, n=102), which pooled data from three single arm trials in patients treated with larotrectinib: LOXO-TRK-14001, LOXO-TRK-15002 (NAVIGATE) and LOXO-TRK-15003 (SCOUT). The model does not consider stratification by tumour site or any other subgroup (e.g. children vs adults) of larotrectinib survival data. This is justified in the CS (p140) on the basis of the small number of events (overall and by tumour site).

In the base-case analysis the comparator was generated using 12 different site specific model engines. Costs and QALYs from each of these 12 engines are then pooled and weighted by the distribution of patients across tumour sites in the integrated primary analysis data. These pooled (historical) comparator results are compared against the results for larotrectinib. The company states that modelling the comparator independently by tumour site avoids the need to synthesise data into a single engine, which would imply loss of transparency and require additional assumptions (p143, CS). For each tumour histology, the comparator engine is informed by extrapolated OS and PFS curves derived from published literature specific to that treatment. The company states that, due to the lack of

other NTRK fusion specific treatments in current clinical practice, clinical practice is driven by tumour site and stage of disease (p141, CS). The sources for the comparator evidence are discussed in Section 5.2.4.

The company also explores, in scenario analyses, two alternative approaches to modelling OS and PFS for the comparator, based on data from patients enrolled in the integrated efficacy analysis. These approaches are discussed in Section 5.2.4.

5.2.2 The company's economic evaluation compared with the NICE reference case checklist

Table 10 Comparison of company's economic evaluation with NICE reference case

Attribute	Reference Case	Included in CS	Comment on whether *de novo* evaluation meets requirements of NICE reference case
Comparator(s)	Alternative therapies in the NHS, including those currently regarded as current best practice	Partly	The comparator in the base case analysis is a weighted comparator of the tumour site specific last line of treatment comparators. The tumour sites included in the comparator reflect the distribution of NTRK fusion in the larotrectinib clinical studies. This may not reflect the distribution of NTRK fusions in the relevant population
Type of economic evaluation	Cost-effectiveness analysis	Yes	
Perspective - costs	NHS and PSS	Yes	NHS and PSS costs have been taken into account.
Perspective - benefits	All health effects on individuals	Yes	QALY benefits to treated individuals were considered.
Time horizon	Sufficient to capture differences in costs and outcomes	Yes	The economic model uses a lifetime horizon (40 years for model engines considering adult patients, 80 years for engines considering pooled populations including both adult and paediatric patients).
Synthesis of evidence on outcomes	Systematic review	Yes	Separate searches were performed to inform the treatment effectiveness, HRQoL, and costs for each of the tumour site specific comparators included in the weighted comparator. Searches were supplemented by targeted review of previous NICE Technology Appraisals.
Outcome measure	QALYs	Yes	EQ-5D-5L and PedsQL data was collected in the LOXO-TRK-15002 and LOXO-TRK-15003 trials, respectively. Both datasets were mapped to EQ-5D-3L.
Health states for QALY measurement	Described using a standardised and validated instrument	Yes	Derived from EQ-5D mapped estimates.
Benefit valuation	Time Trade Off or Standard Gamble	Yes	Time Trade Off

Source of preference data	Representative sample of the public	Yes	Societal tariffs from EQ-5D.
Discount rate	3.5% on costs and health benefits	Yes	Costs and benefits have been discounted at 3.5% per annum.
Equity weighting	No special weighting	Yes	No special weighting undertaken.
Sensitivity analysis	Probabilistic sensitivity analysis	Yes	Probabilistic sensitivity analysis was undertaken.

5.2.3 Population

5.2.3.1 Larotrectinib population

The CS states that the population in the economic analysis is in line with the anticipated marketing authorisation for larotrectinib. At the time of submission, such authorisation was yet to be granted by the European Medicines Agency (EMA). On the 25th July 2019, the Committee for Medicinal Products for Human Use (CHMP) adopted a positive opinion, recommending the granting of a conditional marketing authorisation for larotrectinib to be used for the treatment of adult and paediatric patients with NTRK fusion cancers, which are locally advanced, metastatic, or where surgical resection is likely to result in severe morbidity, and who have no other satisfactory treatment options[44]. The CS considers the population in the larotrectinib integrated efficacy analysis, all of which had a documented NTRK fusion and had exhausted all standard treatment (p123-124, CS) and was the key source of evidence for larotrectinib in the economic analysis, to be in line with the decision problem defined by NICE. Table 11 details previous lines of treatment for the patients enrolled in the larotrectinib integrated efficacy analysis. The company states that patients had been considerably pre-treated and that the majority of them had failed previous surgery or radiotherapy. Therefore, these patients were considered to have exhausted alternative and satisfactory treatment options, as required by the NICE Scope and marketing authorisation for larotrectinib. The position in the treatment pathway at which patients would be eligible for treatment with larotrectinib will vary by tumour site, according to clinical advice to the ERG.

Table 11 Previous lines of treatment

Larotrectinib clinical trial population	Value	Source
Previous systemic therapies (mean)	1.8	CS, p. 138
% patients who received >= 1 line	79.4%	
% patients who received >=3 lines	26.5 %	
% patients who failed surgery	81.4%	
% patients who failed radiotherapy	52%	

The key baseline characteristics of the population considered in the model for larotrectinib are summarised in Table 12.

Table 12 Baseline patient characteristics in the model

Patient characteristics	Value	Source
Mean age, years	[redacted]	Weighted average of mean age in the adult ([redacted]) and paediatric ([redacted]) populations of the larotrectinib integrated efficacy analysis, as reported in the model
Female, n (%)	48 (47.1%)	Calculated from Table 8, CS
ECOG, n (%)		
0-1	91 (89.2%)	
2	11(10.8%)	
Average BSA (m^2) – Adults	[redacted]	Table 50, CS
Average BSA (m^2) - Paediatrics	[redacted]	
NTRK fusion type, n (%)		Calculated from Table 8 for ePAS2+SAS3, CS
1	42 (41.2%)	
2	10 (9.8%)	
3	46 (45.1%)	
Primary tumour type, n (%)		Calculated from Table 8 for ePAS2+SAS3, CS
NSCLC	7 (7%)	
IFS	13 (13%)	
STS	20 (20%)	
Colorectal	6 (6) %	
Salivary gland	17 (17%)	
Breast	1 (1%)	
Pancreas	1 (1%)	
Thyroid	10 (10%)	
Bone sarcoma	2 (2%)	
Cholangiocarcinoma	2 (2%)	
GIST	5 (5%)	
Melanoma	7 (7%)	
Appendix	1 (1%)	
Primary CNS	9 (9%)	
Congenital mesoblastic nephroma	1 (1%)	

The patient population in the larotrectinib integrated efficacy analysis has been previously discussed in Section 3.1. In brief, the ERG concluded that while the patient population in the larotrectinib integrated efficacy analysis falls within the population specified in the NICE scope, it may not be representative of the patients who will be eligible to receive treatment with larotrectinib in clinical practice in terms of age, ECOG status, NTRK fusion type distribution and tumour site distribution.

The ERG is not able to explore how uncertainty in terms of the comparability of the larotrectinib trials population to the relevant population impacts on the generalisability of the cost-effectiveness results, as effectiveness data (OS and PFS) for larotrectinib was not reported by patient characteristics (e.g. tumour site, age, NTRK fusion type and mutation isoform.

5.2.3.2 Comparator population

Historical comparator

Since the larotrectinib clinical evidence is sourced from single arm trials, and the treatment effectiveness of larotrectinib is established through a naïve unadjusted comparison, the comparability of the patient population for the comparator is not guaranteed via randomisation. As described in Section 5.2.1, the company's base case analysis considered a pooled (historical) comparator, comprised of 15 different tumour types (grouped into 12 site specific comparators), derived from published literature specific to that tumour type. The company does not describe patient characteristics in the studies used to inform the weighted comparator by tumour site or overall, so it is not possible to ascertain the comparability of the comparator patient populations with the larotrectinib integrated efficacy analysis population with the patient populations. Thus, it is not possible to assess whether there are any differences in patient baseline characteristics between the intervention and larotrectinib that may result in different baseline risks for OS and PFS (e.g. ECOG status, age, disease stage, etc.).

While all patients in the larotrectinib integrated efficacy analysis carried an NTRK fusion mutation, the proportion of patients with NTRK fusions in the comparator studies is unknown, as the NTRK status is not routinely collected for the majority of tumour sites (Section 4.4.1). In section 2.1 the evidence on the prognostic value of the NTRK fusion was discussed. This evidence is generally mixed and sparse. There is a suggestion that the prognosis of patients with NTRK fusions varies between cancer types and that variation may also exist between NTRK fusion types. However, it is unclear whether the NTRK fusion status has independent prognostic value or if this is driven by its association with other prognostic factors (e.g. ECOG). Therefore, it is uncertain the extent to which the larotrectinib and comparator populations differ in terms of NTRK status, and what are the implications of these differences in terms of disease prognosis.

The company implicitly assumes comparability between patient populations (and outcomes) across some tumour sites, by grouping specific histologies together in the model. The company grouped i) IFS, CMN and paediatric STS patients, ii) bone sarcoma and STS adults (non-GIST), iii) appendix tumours with colorectal cancer. The suitability of the grouping approach is discussed in Section 4.4.1.

Non-responder control

The population characteristics for one of the two alternative comparator modelling approaches, the non-responder control (see Section 5.2.4) are summarised in Table 13 (with further details in Table 114, Appendix L).

Table 13 Patient characteristics for the non-responder control (Adapted from Table 114, Appendix L)

Patient characteristics	Value
Mean age, years	[redacted]
Female, n (%)	[redacted]
ECOG, n (%)	
0-1	[redacted]
2	[redacted]
Average BSA (m^2)	[redacted]
NTRK fusion type	NR
Primary tumour type, n (%)	
NSCLC	[redacted]
IFS	[redacted]
STS	[redacted]
Colorectal	[redacted]
Salivary gland	[redacted]
Breast	[redacted]
Pancreas	[redacted]
Thyroid	[redacted]
Bone sarcoma	[redacted]
Cholangiocarcinoma	[redacted]
GIST	[redacted]
Melanoma	[redacted]
Appendix	[redacted]
Primary CNS	[redacted]
Congenital mesoblastic nephroma	[redacted]
NR, not reported	

The key difference between the patient baseline characteristics of the non-responder subgroup and the larotrectinib integrated efficacy analysis (Table 12), appear to be in terms of the distribution of tumour sites. In general, [redacted] appear to be underrepresented in the non-responder control. In Section 5.2.4, the ERG notes that one of the limitations of the non-responder control approach is that it requires the assumption that there no differences other than response status between responders and non-responders that explain the survival outcomes. Therefore, this approach may be biased if the survival outcomes are expected to differ across tumour sites. This is further discussed in Section 5.2.4.

The population characteristics for the subset of patients from the larotrectinib integrated efficacy analysis whose outcomes were used to inform the alternative comparator modelling approach based on the PFS outcomes of larotrectinib patients at the previous line of treatment are not described in the CS.

5.2.4 Intervention and comparators

5.2.4.1 Intervention

The intervention is larotrectinib, an orally administered TRK inhibitor, and is in line with the NICE scope. Details on the intervention are described in Section 3.2.

Since larotrectinib is a treatment specific for patients with NTRK gene fusion-positive solid tumours, initiation of treatment will require establishing whether patients carry a NTRK gene fusion. There are a number of diagnostic strategies that can be used to identify patients with NTRK gene fusions and this may differ across tumour sites (see Section 2.2.2.2). The company did not consider the need for NTRK fusion testing in the cost-effectiveness analysis (see Section 5.2.8.5). The ERG considers this to be a critical omission, a positive NTRK fusion test is essential to identify patients eligible for larotrectinib.

5.2.4.2 Comparator

The final NICE scope defines the comparator as the established management without larotrectinib for patients with NTRK fusion-positive advanced solid tumours who have either progressed on or not responded to prior therapies, are unfit for chemotherapy or for whom no curative therapy exists. In the base-case analysis, the company considers a set of tumour site specific comparator therapies to reflect current practice which is defined by tumour cancer site and disease stage. These tumour site specific comparator therapies are a mixed basket of last line standard of care, as the company considers that the larotrectinib anticipated marketing authorisation restricts the use of larotrectinib to patients who have exhausted all satisfactory treatment options. The basket includes twelve different standard of care

therapies, weighted by the distribution of tumour types (patient enrolment per tumour site) in the larotrectinib integrated efficacy analysis. Table 14 presents the list of comparator treatments by tumour site, alongside the key evidence sources and the company's justification for their selection.

Table 14 List of comparator therapies by tumour site

Tumour Site	Comparator treatment	Data source*	Details	CS Justification
NSCLC	Best Supportive Care	TA374[45]	Placebo arm of Shepherd 2005[46], Phase III RCT in NSCLC after failure of first-line or second-line chemotherapy.	Considered to represent the proposed marketing authorisation for larotrectinib, as it includes patients that progressed following prior chemotherapy. The study was also used in a more recent appraisal to represent standard of care.
Salivary gland	Cisplatin + vinorelbine	Airoldi 2001[47]- survival outcomes and AE rates	Cisplatin + Vinorelbine arm of Phase II RCT in recurrent malignancy of major or minor salivary gland origin	No previous NICE Appraisals identified. Review of ASCO and NCCN guidelines confirmed lack of established practice.
Melanoma	Mixed chemotherapy including dacarbazine, paclitaxel, carboplatin, temozolomide and paclitaxel + carboplatin	TA357[48]	Mixed chemotherapy arm of the Keynote-002[49] Phase II RCT in advanced melanoma progressed after ipilimumab.	Considered representative of larotrectinib patient population according to proposed marketing authorisation.
Colorectal and appendix	Best Supportive Care	TA405[50]	Placebo arm of RECOURSE Phase III RCT[51] in metastatic colorectal cancer after at least two previous lines of standard chemotherapy	Considered representative of larotrectinib patient population according to proposed marketing authorisation. Appendix grouped with CRC as no tumour histology specific relevant source was identified and small number of appendix patients in larotrectinib clinical programme (n=1).
GIST	Best Supportive Care	TA488[52]	Placebo arm of the GRID Phase III RCT[53] in metastatic or unresectable GIST after failure of at least two previous lines	Considered representative of larotrectinib patient population according to proposed marketing authorisation.
Adult STS (non GIST) and bone sarcoma	Best Supportive Care	TA185[54]	Data from unpublished studies and one study[55] in adult patients with advanced STS in the EORTC dataset, who failed ifosfamide or standard chemotherapy as second-line therapy.	Considered representative of larotrectinib patient population according to proposed marketing authorisation. Bone sarcoma grouped with adult STS as no tumour histology specific relevant sources were identified and clinical experts considered outcomes between STS and bone sarcoma after failing previous therapies to be similar.

STS(paediatric), IFS and CMN	Irinotecan and vincristine	Mascarenhas et al, 2010[56]	Irinotecan and vincristine arm from the Mascarenhas et al. study, a Phase II RCT in first relapsed or progressed rhabdomyosarcoma after previous therapies failure.	Considered representative of larotrectinib patient population according to proposed marketing authorisation. IFS and CMN were grouped with STS as no tumour histology specific relevant sources. Clinical experts considered IFS to be a type of paediatric STS. CMN was grouped based on small numbers (n=1).
Breast	Treatment of physician's choice including vinorelbine, gemcitabine, paclitaxel, doxorubicin and docetaxel	TA423[57]	Treatment of physician choice arm from the EMBRACE Phase III RCT[58] in women with locally advanced or metastatic breast cancer who had at least two previous chemotherapy regimens.	Considered representative of larotrectinib patient population according to proposed marketing authorisation.
Cholangiocarcinoma	Gemcitabine + cisplatin	Valle et al 2010[46]	Gemcitabine and cisplatin arm from the ABC-02 Phase III RCT, in patients with unresectable, recurrent, locally advanced or metastatic biliary tract cancer.	Considered representative of larotrectinib patient population according to proposed marketing authorisation.
Pancreatic	5-FU + leucovorin	TA440[59]	5-FU and leucovorin arm of the NAPOLI-1 Phase III RCT in metastatic pancreatic ductal adenocarcinoma previously treated with gemcitabine-based therapy.	Considered representative of larotrectinib patient population according to proposed marketing authorisation.
Gliomas (CNS)	Lomustine	Batchelor et al 2010[60]	Lomustine arm from Batchelor et al. Phase III RCT in patients with recurrent glioblastoma previously treated with temozolomide-containing chemotherapy or radiation	Considered representative of larotrectinib patient population according to proposed marketing authorisation.
Thyroid	Best Supportive care	TA535[61]	Placebo arm of the DECISION RCT[62] in advanced hepatocellular carcinoma previously untreated with systemic therapy	Considered representative of larotrectinib patient population according to proposed marketing authorisation.

ASCO, American Society of Clinical Oncology; CMN, congenital mesoblastic nephroma; CRC, colorectal cancer; EORTC, European Organisation for Research and Treatment of Cancer; GIST, gastrointestinal stromal tumor; IFS, infantile fibrosarcoma, NCCN, National Comprehensive Cancer Network; NICE, National Institute for Health and Care Excellence; NSCLC, non-small cell lung cancer; RCT, randomized controlled trial; STS, soft tissue sarcoma.

The ERG has previously highlighted the challenges in validating the company's approach to model treatment effectiveness of the historical comparator (see Section 4.4.1). The ERG noted that the company's approach to selecting suitable comparator data was reasonable, however it is likely to be sensitive to the choice of data, or the analysis method used. Importantly, the use of data extracted from a range of previous NICE TAs and other published sources to inform the survival estimates of the historical comparator requires strong assumptions in terms of the comparability of the populations to the larotrectinib integrated efficacy analysis. The comparability of the larotrectinib and comparator population in terms of characteristics that may impact on prognosis (e.g. ECOG status, age, disease stage, etc.), which may bias the estimates of treatment effectiveness, cannot be assessed. The company does not report patient baseline characteristics for the comparator data sources, and interpretation of the potential impact of these characteristics would have been difficult, given the large number of tumour histologies and data sources. Therefore, the use of an unadjusted naïve comparison between tumour site specific comparators and larotrectinib may introduce confounding bias of unknown magnitude and direction in the estimates of cost-effectiveness.

One important potential source of confounding affecting the treatment effectiveness estimates are the treatments provided once patients have progressed. A number of patients treated with larotrectinib in the ePAS2 population who progressed were subsequently treated with LOXO-195 ([redacted] out of [redacted]), a second-generation selective TRK inhibitor, while [redacted] patients continued treatment with larotrectinib after disease progression. In the company's base case analysis, the gains in post-progression survival for larotrectinib far exceed the gains in PFS ([redacted] vs [redacted] life years gained (LYG)), suggesting that survival gains may be driven by post-progression treatments. Patients in the studies informing the effectiveness of the historical control would not have had access to LOXO-195 or larotrectinib, so the use of the historical data to inform comparator may bias estimates of treatment effectiveness in favour of larotrectinib.

Another important potential source of bias relates to NTRK status in the comparator population. As NTRK fusions were previously not considered actionable mutations, the NTRK fusion status was not routinely collected for the majority of tumour histologies, and the prevalence of these mutations in the historical data is unknown. In Section 2.1, the ERG notes that the evidence on the prognosis of patients with an NTRK fusion is generally mixed and sparse, and it is unclear whether the NTRK fusion status has independent prognostic value. This further hinders ascertaining the comparability of the intervention and comparator populations. The company conducted a scenario analysis to explore the impact on cost-effectiveness of assuming that NTRK fusion status has prognostic value and adjusting survival estimates in accordance. In this scenario, the company applied a hazard ratio (HR) comparing the risk of death between patients with and without NTRK fusions (HR=2.17) from a study in colorectal

cancer to adjust the OS and PFS curves of the comparator. This assumes that patients carrying the NTRK have an increased risk of dying compared to the ones who do not, and also that the comparator population did not include any NTRK+ patients. Since the company also applied the HR to PFS, it is further assumed that NTRK positivity has the same detrimental impact on disease progression as for death. The company applied this HR to the survival curves of the colorectal comparator in one analysis, and to the all comparators with NTRK prevalence lower than 25% (the CS incorrectly refers to this as greater than 25%). The first scenario reduces the ICER to [redacted] per additional QALY, while the second has greater impact reducing the ICER to [redacted] per additional QALY (see Table 38). The ERG notes that the company only explored the potential detrimental impact of NTRK status on survival outcomes. However, there is evidence suggesting that for at least one other tumour site NTRK status may improve prognosis [4]. Therefore, the impact of NTRK status on disease prognosis remains unknown.

The company considered two alternative approaches to generate a comparator. Both were considered unfeasible by the company, and were not implemented: i) unanchored indirect treatment comparison using a dataset of comparable patients, and ii) matched-adjusted indirect comparison, where a similar dataset would be adjusted by propensity score matching to minimise imbalances between comparator and intervention. The company's systematic reviews (see Section 5.1.) did not identify any published sources that adequately reflects the cohort of patients enrolled in the larotrectinib integrated efficacy analysis, and therefore the company could not implement the first approach. The company justifies not implementing the second approach because all the available comparator data was tumour site specific, and therefore, matching on the basis of tumour site would result in the loss of the vast majority of larotrectinib patients, before other covariates could be considered. The company's modelling methods review (Section B.3.2, CS) identified two alternative approaches to assess larotrectinib treatment effectiveness when an appropriate historical control is not available: i) comparison vs non-responders, and ii) comparison vs previous line of treatment. These two approaches were implemented in the economic model and two scenario analysis examine their impact on the estimates of cost-effectiveness. The alternative comparator approaches are described below.

The first uses the non-responders from larotrectinib integrated efficacy analysis (those with stable or progressed disease, n=35) as a proxy for patients not receiving an active treatment, with the observed time-to-event data for these patients extrapolated to derive OS and PFS. All other tumour site specific HRQoL and resource use and costs data sources and assumption remain the same as for the historical comparator. The approach taken to extrapolate comparator survival based on larotrectinib non-responder survival data is detailed in Section 5.2.6.2. The comparison against non-responders assumes that patients pre-treated with larotrectinib that did not respond represent an untreated population and that lack of treatment response is equivalent to non-exposure to treatment. The company states that

the advantage of this approach is that all patients in the non-responder subgroup met the same trial inclusion and exclusion criteria, and are receiving the same line of treatment. The limitations of the approach are outlined in the CS (p148, and Table 33). The small numbers of patients and events informing the survival models introduce considerable uncertainty in the extrapolation of OS and PFS. Furthermore, the approach requires strong assumptions, namely that there are no differences other than response status between responders and non-responders that explain the survival outcomes, and that non-responders derive no benefit and no harm from treatment with larotrectinib. The ERG notes that evidence suggests response appears to be systematically correlated with tumour type as shown by ERG the analysis of response rates by tumour site (see Section 4.6.1), so the assumption that survival is dependent on response status is unlikely to hold. The assumption of no treatment benefit or harm may also not hold, as some patients may receive some treatment benefit from larotrectinib, even if they do not have a partial or complete response.

The second approach uses data taken from the larotrectinib trial patients' previous line of treatment to derive OS and PFS curves. In this approach, the inverse of the ratio between average time-to-treatment progression (TTP) on their previous therapy and the mean extrapolated PFS with larotrectinib (also called the GMI multiplier) is applied to all health outcomes (total LYG and QALYs) for larotrectinib. This crude adjustment assumes that larotrectinib is more effective in terms of both PFS and OS than the comparator by the same proportion as the GMI multiplier. Therefore, the resulting GMI adjusted total mean LYG and QALYs are assumed to correspond to comparator outcomes and applied in the calculation of the ICER (based on LYG and QALYs). The company states that this approach is likely to be conservative due to considerable censoring of PFS for larotrectinib and the comparison being established against unrestricted TTP from the previous line. The bias against the later line treatment (i.e. larotrectinib), is also partly caused by patient's baseline status being likely to decline over the course of disease. Another limitation of this approach is that it provides no comparative OS estimates (as all patients with a previous line have survived to receive larotrectinib), and so additional assumptions on OS are required. Finally the company states that the patients in the previous-line of therapy received active treatments that would have varied substantially and may not have been reflective of treatments received in clinical practice in England.

The ERG notes that the company's approach to use a within-study previous line of treatment comparator could have been implemented in a more formal way by directly adjusting the larotrectinib total LYG and QALYs to produce the corresponding comparator outcomes. Hatswell and Sullivan[63] have outlined an approach whereby the mean gain in TTP comparing the intervention PFS to the previous line TTP is assumed to correspond to the treatment effect. The key difference between this approach and the

company's, is that this one assumes that the treatment effect is limited to delaying disease progression only.

ERG commentary

The ERG considers that the company appropriately explored alternative approaches to model the comparator survival outcomes. All approaches have limitations and may result in biased estimates of treatment effectiveness. The non-responders control approach may provide a more transparent and potentially flexible alternative to modelling comparator effectiveness than the pooled historical comparator. The non-responder approach is less affected by confounding from subsequent lines of treatment and imbalance of patients characteristics, and it is easier to assess how deviations from its key assumptions may impact on the cost-effectiveness results. The ERG performs further analysis to explore the uncertainty surrounding the treatment effect in Section 6, where both the historical comparator and the non-responder control approaches are utilised.

In principle, the previous line of treatment control as described in the literature[63] is also a valid approach to reduce the confounding of treatment effect, even if it relies on a different set of assumptions than the non-responder control. The ERG did not, however, attempt to implement it in the model, as it was not considerable feasible given data availability and time constraints.

5.2.5 Perspective, time horizon and discounting

A lifetime horizon is used in the economic model. For model engines considering adult patients only a 40 year time horizon is used. For paediatric populations (STS paediatric comparator) and the larotrectinib engine (which pools adult and paediatric patients) an 80 year time horizon is used. The use of differential time horizons is only required, because the company has pooled survival outcomes across adults and paediatric patients for larotrectinib. In Section 5.2.6, the ERG outlines concerns about the suitability of pooling effectiveness data across adult and paediatric populations (as well as study designs), and how heterogeneity at this level may be driving the survival benefits for larotrectinib. The ERG considers that it might have been more appropriate to model adult and paediatric patients separately, and use appropriate time horizons for each model. The ERG did not implement a common time horizon in the company's model, as this would not have addressed the potential bias introduced by pooling the survival outcomes of two groups of patients with substantial differences in terms of life expectancy.

Both costs and benefits were discounted at an annual rate of 3.5%, as per the NICE reference case. The CS also presents a scenario analysis using a discount rate of 1.5% per annum for costs and benefits (Table 57, CS).

A National Health Service and Personal and Social Services perspective is used.

5.2.6 Treatment effectiveness and extrapolation

5.2.6.1 Larotrectinib treatment effect

The main effectiveness inputs included in the company's model were PFS and OS. The survival estimates for larotrectinib were derived from uncontrolled pooled data (n=102) collected in three single arm trials in patients treated with larotrectinib (the integrated efficacy analysis). The larotrectinib survival data was extrapolated by fitting standard parametric distributions (Weibull, exponential, Gompertz, log-logistic, log-normal and generalised gamma) to the PFS and OS Kaplan KM curves based on the integrated efficacy analysis data. A wider range of survival models including Cox and spline/piecewise models were initially explored by the company in an earlier data cut (n=73, no cut-off date presented), and are described in Appendix L. The company states that, given that the more complex models did not suggest a better fit to data (based on the Akaike and Bayesian Information Criterion [AIC and BIC]) in the earlier data cut, only standard parametric models were examined for the data cut used to inform the cost-effectiveness model. The assessment of appropriate parametric models, used to inform the partitioned survival analysis considered the i) visual inspection of the Kaplan Meier curve and log cumulative hazard plot, ii) visual fit of extrapolated models to observed data, iii) statistical fit of the survival models based on AIC and BIC, and iv) clinical plausibility of the extrapolation. It is unclear to the ERG why the company considered it more informative to conduct a more comprehensive exploration of survival models for a smaller and more immature data cut of effectiveness data, rather than the later data cut used to inform the economic model. Nevertheless, given that the integrated efficacy analysis is still immature and has small sample size (n=102), it is unlikely that model fit of more complex survival model would have improved markedly for this data cut.

The integrated efficacy analysis data constitutes three separate studies which differed in terms of study design and patient populations (see section 4.2). Given these differences between the studies, pooling the survival data across the three studies may introduce bias in the overall OS and PFS estimates. The ERG notes that there is lack of consistency between the estimates of median PFS for the ePAS2 population (n=93) in the integrated efficacy analysis (i.e. excluding CNS patients), 27.4 months, and for an earlier data cut (n=47, 17th July 2017) of the NAVIGATE trial (p78 of the corresponding clinical study report), [redacted] months. In response to the request by the ERG for further explanation, the company stated that the different data-cuts from which these estimates were sourced is a factor in the difference in median PFS (see company's clarification additional response), but did not provide evidence to support this statement (e.g. median PFS by study for the 30th July 2018 data-cut). The median OS duration [redacted] for the integrated efficacy analysis (30th July

2018 data cut) or NAVIGATE at the 17thJuly 2017 data cut and so it cannot be compared, but it is likely that heterogeneity will also affect OS estimates. The CS did not explore or discuss the potential impact of heterogeneity in survival estimates or explore its impact on the estimates of cost effectiveness.

The pooled clinical trial data for larotrectinib is also affected by heterogeneity across tumour site. While the company claims that there is no suggestion of heterogeneity in treatment effect of larotrectinib for any of the subgroups for which data was requested (response to A3 in points for clarification), the ERG exploratory analyses on response data suggest the opposite (see Section 4.6.1). These analyses showed that response outcomes for larotrectinib could vary considerably across tumour sites, with higher ORRs for tumour sites where NTRK fusions are highly incident (e.g. IFS, salivary gland cancer) and for paediatric cancers. The analyses also highlighted that the company's mean ORR estimate (72%, 95% CI 62 to 81) for the ePAS2 analysis (across all sites except CNS tumours) (see Table 12, CS) may have overestimated response for patients receiving larotrectinib by not reflecting heterogeneity across tumour sites. The ERG estimated an average response rate across all sites (except CNS tumours) of 64% (95% CrI 29 to 83) using a Bayesian hierarchical model to account for tumour site response heterogeneity. When CNS tumour were also included in the analysis, the estimated average response rate decreased to 57% (95% CrI 23 to 80).

The company has not provided OS and PFS data for larotrectinib by tumour site, age, response status, response category, fusion type and mutation isoform, thus precluding further exploration (see company's response to Points for Clarification and clarification additional response). The ERG requested data across a variety of formats so as to allow some exploration of heterogeneity in the time-to-event outcomes. The ERG further requested the KM PFS and OS curves by study included in the larotrectinib integrated efficacy analysis. The company did not accede to this request either (see company's clarification additional response), and claimed that the two main reasons for not providing survival data by study were:

"*(1) Given the mode of action and strong biologic rationale for the histology-independent indication, an analysis of data "by study" is not considered to be meaningful as compared to an analysis "pooled by study, but specific for tumour type" and (2) the Phase 1 study only contributed 8 patients of various tumour types into the initial primary analysis set."*

The ERG presents further analyses in Section 6 exploring the potential impact of heterogeneity in the treatment response on the cost-effectiveness results.

5.2.6.2 Comparator treatment effect

Comparator PFS and OS were modelled independently for each tumour site in the base-case, and informed by historical data, namely previous NICE technology appraisals (TAs), or publications identified by the company as representative of the efficacy of the standard of care for each of the tumour locations (see Section 5.2.4.2). The approach followed is detailed in Appendix M.

For tumour sites NSCLC, melanoma, colorectal/appendix, GIST, non-GIST/bone sarcoma, breast, pancreas and thyroid, the company states that they have attempted to model survival data as closely as possible to the appraisal committee's preferred approach. For tumour sites (salivary, STS paediatrics/ IFS / congenital mesoblastic nephroma, cholangiocarcinoma and glioma/CNS), for which a previous NICE TAs was not available or was considered inappropriate to inform the survival estimates, the survival data was sourced from published data identified in systematic reviews and OS and PFS Kaplan Meier curves were digitised and IPD recreated, to allow for fitting of parametric distributions to extrapolated OS and PFS. For tumour sites in which the company did not identify evidence to inform survival estimates, equivalence between these tumour sites and other for which treatment at the relevant line of treatment was similar was assumed (company's response to Points for Clarification and Appendix M). However, it is not clear whether the company considered the comparability of patient's prognosis across the grouped tumour sites, even if patients receive similar treatments. For example, patients with IFS and with congenital mesoblastic nephroma were assumed to be comparable to other paediatric STS patients, and comparator survival data was sourced from a study in patients with relapsed or progressed rhabdomyosarcoma.

The ERG discusses in Section 5.2.4, the challenges and potential biases of using an uncontrolled comparison to inform treatment effectiveness. On balance, the ERG considers that alternative approaches to examine the comparator effectiveness should be jointly explored when estimating treatment effect, as all approaches have limitations and may result in biased estimates of treatment effectiveness. Both the pooled historical comparator and non-responder control approach are suitable for the purpose of exploring uncertainty in the treatment effect in the absence of a controlled comparison. The non-responders control, however, may provide a more transparent approach to establish magnitude and direction of bias affecting treatment effectiveness estimates. Importantly, it can be modified so that larotrectinib's OS and PFS are modelled conditional on response rates. This is advantageous as it provides a flexible framework to explore the potential impact of heterogeneity on response rates across tumour sites (or any other potential source of heterogeneity) on the cost effectiveness estimates results. In section 6, the ERG explores a series of scenario analyses using a dual partition survival model that allows modelling responders and non-responders separately.

The approach taken by the company to model treatment effectiveness (PFS and OS) for larotrectinib, the historical comparator and the non-responder comparator is described in greater detail in the next sections.

5.2.6.3 Progression free survival

Larotrectinib

Table 15 summarises the survival models investigated to extrapolate larotrectinib PFS, along with the main justification provided by the company for use in their base-case analysis.

Table 15 Summary of company's justification for PFS extrapolation curves selection – larotrectinib

Parametric distribution	Goodness of visual fit	Best statistical fit	Clinically plausible
Exponential	Similar fit to the observed PFS data for all distributions, with no distinguishable better visual fit.	[redacted]	[redacted]
Weibull		[redacted]	[redacted]
Log normal		[redacted]	[redacted]
Log logistic		[redacted]	[redacted]
Generalised Gamma		[redacted]	[redacted]
Gompertz		[redacted]	[redacted]

The base-case parametric distribution is highlighted in bold.

The company fitted 6 standard parametric models to extrapolate the PFS data from the integrated efficacy analysis. In this dataset, 37 (36%) of patients had experienced either progression or death, with the remaining patients censored (n=65).

The company selected the [redacted] distribution to extrapolate PFS in the base-case analysis, and tested other parametric models in sensitivity analysis. The parametric model choice was mostly driven by the clinical plausibility of predicted survival estimates. Goodness of visual and statistical fit was similar across the different survival models. The mean time required for 10% and 1% of patients to remain progression free for each of the parametric distributions are shown in Table 16. [redacted] were considered implausible, based on the length of time in which patients were predicted to remain in PFS. The long-term projections of the

[redacted] and [redacted] distributions were considered more clinically plausible. The company also highlighted that these two models underestimated PFS compared to the observed data (see Figure 12). The company preferred the [redacted] distribution for the base case analysis as the [redacted] assumes a constant hazard risk and does not account for the change in risk with aging.

Table 16 Estimated PFS by parametric form (years from start of treatment) (Appendix L, p399)

Model	10% PFS	1% PFS
Exponential	[redacted]	[redacted]
Weibull	[redacted]	[redacted]
Log normal	[redacted]	[redacted]
Log logistic	[redacted]	[redacted]
Generalised Gamma	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]

Figure 12 Larotrectinib PFS extrapolated curves (Appendix L, p391)

[redacted]

The ERG notes that, given the small patient numbers and the immaturity of the data, the extrapolation of PFS is considerably uncertain, and that this is likely to be resulting in clinically implausible predictions for survival models that have similar goodness of fit to the observed data. The two distributions considered to be more plausible by the company still predict that 1% of patients remain in PFS after [redacted] and [redacted] years for the [redacted] and [redacted] distributions, respectively. The cost effectiveness results appear to be sensitive to the choice of distributional form used to extrapolate PFS for larotrectinib with the incremental cost effectiveness ratio (ICER) varying between [redacted] ([redacted]) and [redacted] per additional QALY ([redacted]) for alternative parametric functions (company's response to Points for Clarification, p76). The differences in ICERs appear to be driven by differences in costs (importantly the cost of treatment) rather than by differences in QALY gains.

Historical comparator

Table 17 summarises the company's base case survival approach to extrapolate PFS by tumour site for the pooled historical comparator, and the rationale for model selection.

Table 17 Summary of company's justification for PFS extrapolation curves selection – historical comparator

Tumour Type	Survival model	Approach	Rationale
NSCLC	[redacted]	The ERG to TA374 did not report the parameters of the survival function used to model the PFS of the placebo arm of Shepherd et al, 2005 [64]. The company digitised the exponential tail fitted by the ERG to TA374, and back-calculated the coefficient for the exponential distribution applied.	Approach used in TA374[45]
Salivary	[redacted]	No KM available from published source (Airoldi et al., 2001[47]); assumed exponential with parameter calculated from median survival.	No previous TA. Airoldi et al., 2001[47]; reported outcomes of salivary gland cancer patients treated with cisplatin + vinorelbine, and was considered to be reflective of last line of treatment for these patients.
Melanoma	[redacted]	The company digitised the chemotherapy arm KM curve from the TA357 CS and fitted a distribution to the end of the KM.	Approach used in TA357[48]
Colorectal/Appendix	[redacted]	The company digitised the placebo KM curve from the ERG pooled analysis and fitted standard parametric curves to model PFS.	Approach used in TA405[50]
GIST	[redacted]	The company digitised the placebo KM curve from the TA488 documentation and fitted standard parametric curves to model PFS.	Approach used in TA488 [52]
Non-GIST/Bone sarcoma	[redacted]	All patients entered the model with progressed disease	Approach used in TA185[54]
STS paediatrics/IFS/ congenital mesoblastic nephroma	[redacted]	The company digitised the irinotecan + vincristine arm KM curve from Mascarenhas et al., 2010[56], and fitted standard parametric curves to model PFS.	No TA found in the STS included in the model or in IFS and congenital mesoblastic nephroma. The PFS was informed by a study in patients with relapsed or progressed rhabdomyosarcoma, a tumour histology that was not observed in the patients included in the larotrectinib clinical trial. The survival model selected is said to have the best statistical fit.
Breast	[redacted]	The company digitised the treatment of physician's choice arm KM from TA423 documentation.	Approach used in TA423[57]
Cholangiocarcinoma	[redacted]	The company digitised the gemcitabine and cisplatin arm KM curve from Valle et al., 2010[46], and fitted standard parametric curves to model PFS.	No TA found in cholangiocarcinoma. The PFS was informed by a study in patients with gallbladder cancer and cholangiocarcinoma. The distribution was selected based on goodness of statistical fit.

Glioma (CNS)	[redacted]	The company digitised the lomustine arm KM curve from Batchelor et al., 2010[60], and fitted standard parametric curves to model PFS.	No TA found that reflected current CNS tumour treatment. The PFS was informed by a study in patients with recurrent malignant glioma. The distribution was selected based on best-statistical fit and fit by visual inspection
Pancreas	[redacted]	The company digitised the 5-FU + leucovorin arm KM in the TA440 submission and fitted an exponential tail to the end of the KM.	Approach used in TA440[59]
Thyroid anaplastic, follicular and papillary	[redacted]	The company digitised the placebo arm KM in the TA535 CS and fitted an exponential tail to the end of the KM.	Approach used in TA535 [61]
KM, Kaplan Meier; TA, technology appraisal			

The ERG identified a few inaccuracies and inconsistencies in the reported approaches to extrapolate comparator PFS. For example, the company justifies the use of a [redacted] distribution to extrapolate PFS of the comparator for STS paediatrics/ IFS / congenital mesoblastic nephroma based on its goodness of fit. However, the [redacted] distribution has a better statistical fit based on AIC (414.0397 vs 418.1099, Table 159, Appendix M) and BIC is not reported. The company did not report a scenario analysis where comparator PFS for this tumour site was extrapolated with the [redacted] distribution, and this distribution was not fully implemented in the economic model. The survival distribution selected to extrapolate PFS for the GIST comparator was reported as being the [redacted] and the [redacted] in different places of the CS (p506 of Appendix M and Table 37, respectively). In the economic model the [redacted] was implemented for the base-case analysis. Another example of inconsistency of report was identified for the pancreatic cancer comparator for which PFS is stated [redacted]_5-FU + leucovorin arm KM curve in the TA440 submission [59] on Table 37 of the CS, and [redacted] on p548 of Appendix M.

The ERG notes that, while the inaccuracies and inconsistencies noted above were identified, it was not considered feasible to fully validate the company's approach to estimate the historical comparator's effectiveness. This would require not only reviewing all of the original data sources for accuracy, but also to validate numerous assumptions on population comparability.

Non-responder comparator

The company stratified the larotrectinib integrated efficacy analysis PFS data by response status, with a responder defined as having achieved a complete response or partial response, and a non-responder having stable or progressive disease. The company compared the observed PFS by response status with the log rank test and tested for proportional hazards using the Schoenfeld test in a Cox proportional hazards model, concluding that survival times for responders vs non-responders were i) statistically different (greater survival times for responders), and ii) the assumption of proportional hazards does not hold between responders and non-responders (see Appendix L).

The company fitted an accelerated failure time model to the observed larotrectinib PFS data with response status as a covariate and tested six standard parametric curves. Table 18 summarises the survival models investigated to extrapolate PFS for the response-based model, along with the main justification provided by the company for its use in the non-responder control scenario analysis (p213, CS, and Appendix L).

Table 18 Summary of company's justification for PFS extrapolation curves selection – Responder/Non-Responder Analysis

Parametric distribution	Goodness of visual fit	Best statistical fit	Clinically plausible
Exponential	Similar fit to the observed PFS data for all distributions, with no distinguishable better visual fit.	[redacted]	[redacted]
Weibull		[redacted]	[redacted]
Log normal		[redacted]	[redacted]
Log logistic		[redacted]	[redacted]
Generalised Gamma		[redacted]	[redacted]
Gompertz		[redacted]	[redacted]

The parametric distribution used in the company's responder control scenario analysis is highlighted in bold.

The company selected the [redacted] distribution to extrapolate PFS for the non-responder control to remain consistent with the base-case assumption for the PFS of larotrectinib and because it was considered the most clinically plausible distribution. The [redacted] distribution had the best statistical fit, but was considered clinically implausible given the lengthy survival estimates (see Table 19). The company notes that despite the assumption of proportional hazards appearing not to hold for PFS, it was still reasonable to prefer the [redacted] distribution given the limited number of data points and the relative uncertainty of the statistical analysis.

Table 19 Estimated PFS by parametric form (years from start of treatment) (Appendix L, p460)

Model	10% progression-free		1% progression-free	
	Responders	Non-responders	Responders	Non-responders
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log normal	[redacted]	[redacted]	[redacted]	[redacted]
Log logistic	[redacted]	[redacted]	[redacted]	[redacted]
Generalised Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]

Figure 13 shows the visual fit of the alternative extrapolation assumptions for response stratified PFS.

Figure 13 Parametric model fittings: PFS stratified by response status (Appendix L, p459)

The ERG notes that, as per the extrapolation of the non-stratified larotrectinib PFS, issues of sample size and data immaturity also affect the response-based PFS extrapolation, and are compounded by the stratification. The responder extrapolated PFS curves generate more uncertain survival projections compared to non-responders despite the higher number of responders (n=67 vs n=35) due to fewer events in responders (see Figure 107, Appendix L).

5.2.6.4 Overall survival

Larotrectinib

Table 20 summarises the survival models investigated to extrapolate OS for larotrectinib, along with the main justification provided by the company for use in their base-case analysis.

Table 20 Summary of company's justification for OS extrapolation curves selection – larotrectinib

Parametric distribution	Goodness of visual fit	Best statistical fit	Clinically plausible
Exponential	Similar fit to the observed OS data for all distributions, with [redacted]	[redacted]	[redacted]
Weibull		[redacted]	[redacted]
Log normal		[redacted]	[redacted]

Log logistic	providing better fit to later time points in the KM.	[redacted]	[redacted]
Generalised Gamma		[redacted]	[redacted]
Gompertz		[redacted]	[redacted]

The base-case parametric distribution is highlighted in bold.

The company followed an approach similar to that used to extrapolate PFS and described in Section 5.2.6.1.

The [redacted] distribution was also preferred to extrapolate OS for larotrectinib in the company's base-case. The visual fit of the extrapolated curves to the KM curve was considered to be similar across the six standard parametric models explored by the company (see Figure 14 below, and Figure 68 in Appendix L), as was the statistical fit based on AIC and BIC (Table 108 in Appendix L). Therefore, the choice of parametric distribution was again driven by the clinical plausibility of the OS predictions. The company states that only the [redacted] and the [redacted] distributions are clinically plausible, as all others predict that patients would live for a biologically implausible amount of time over the 80 year time horizon (Table 110, Appendix L).

Figure 14 Larotrectinib OS extrapolated curves (Appendix L, p413)

The ERG notes that the larotrectinib OS data is even more affected by uncertainty than the PFS data, due to immaturity of observed data and the small number of events. Only 14 deaths were observed in the integrated efficacy analysis data and 88 patients were censored. Similar to the PFS extrapolation, OS survival models with similar goodness of fit, result in considerably different survival projections, with expected total discounted life years ranging between [redacted] (exponential) to [redacted] years (Gompertz) (company's response to Points for Clarification, p75-76). The ERG also notes that the OS gains for larotrectinib are driven by the long post-progression survival, with considerable separation between extrapolated PFS and OS curves as shown in Figure 15. As discussed at the beginning of Section 5.2.6, the ERG considers that survival gains may be driven by post-progression treatments, and that is likely to bias the comparison in favour of larotrectinib. In the base case analysis, treatment effectiveness in the cost-effectiveness analysis was established using an uncontrolled naïve

comparison between larotrectinib and the pooled historical comparator, in which patients do not have access to LOXO-195 or larotrectinib. Therefore the bias introduced by larotrectinib patients receiving post progression treatment with LOXO-195 or larotrectinib was not adjusted for. The ERG requested a scenario adjusting the treatment effect to reflect current NHS clinical practice, i.e. excluding LOXO-195 and other therapies not currently available and/or recommended in the NHS (B5, Points for Clarification). The company did not provide this scenario, and argued that to adjust survival for further lines of treatment would be inappropriate and out of line with clinical practice as patients without other treatment options have access to experimental treatments (A3, company's response to Points for Clarification). This, however, ignores that the comparator in the company's base-case is a historical comparator and that patients would not have had access to either larotrectinib or LOXO-195.

Figure 15 Observed and extrapolated survival curves for larotrectinib with company's preferred assumptions (adapted from company's model)

Historical Comparator

Table 21 summarises the company's base case survival approach to model OS by tumour site for the pooled historical comparator, and the rationale for model selection.

Table 21 Summary of company's justification for OS extrapolation curves selection – pooled historical comparator

Tumour Type	Survival model	Approach	Rationale
NSCLC	[redacted]	The ERG to TA374 did not report the parameters of the survival function used to model the OSS of the placebo arm of Shepherd et al, 2005[64]. The company digitised the exponential tail fitted by the ERG to TA374, and back-calculated the coefficient for the exponential distribution applied. The exponential coefficient was then calibrated to predict a similar number of life years were estimated by the ERG model and this analysis.	Approach used in TA374[45]
Salivary	[redacted]	The company digitised the cisplatin + vinorelbine arm KM curve from Airoldi et al., 2001[47], and fitted standard parametric curves to model OS.	No previous TA. Airoldi et al., 2001[47] reported outcomes of salivary gland cancer patients treated with cisplatin + vinorelbine, and was considered to be reflective of last line of treatment for these patients. The survival model was selected based on goodness of statistical fit.
Melanoma	[redacted]	The company digitised the chemotherapy arm KM curve from the TA357 CS and fitted standard parametric curves to model OS.	Approach used in TA357 [48]
Colorectal/Appendix	[redacted]	The company digitised the placebo KM curve from the ERG pooled analysis and fitted standard parametric curves to model OS.	Approach used in TA405[50]
GIST	[redacted]	The company digitised the placebo KM curve from the TA488 documentation and fitted standard parametric curves to model PFS.	Approach used in TA488[52]
Non-GIST/Bone sarcoma	[redacted]	The company used the standard parametric distribution parameters reported in TA185 for the historical control to generate OS curves.	Approach used in TA185[54]
STS paediatrics/IFS/ congenital mesoblastic nephroma	[redacted]	The company digitised the irinotecan + vincristine arm KM curve from Mascarenhas et al., 2010[56], and fitted standard parametric curves to model OS.	No TA found in the STS included in the model or in IFS and congenital mesoblastic nephroma. The OS was informed by a study in patients with relapsed or progressed rhabdomyosarcoma, a tumour histology that was not observed in the patients included in the larotrectinib clinical trial. The survival model was selected based on goodness of statistical fit.

Breast	[redacted]	The company digitised the treatment of physician's choice arm KM from TA423 documentation and fitted an exponential tail to the end of the KM.	Approach used in TA423[57]
Cholangiocarcinoma	[redacted]	The company digitised the gemcitabine and cisplatin arm KM curve from Valle et al., 2010[46], and fitted standard parametric curves to model OS.	No TA found in cholangiocarcinoma. The PFS was informed by a study in patients with gallbladder cancer and cholangiocarcinoma. The distribution was selected due to its best statistical fit (AIC and BIC).
Glioma (CNS)	[redacted]	The company digitised the lomustine arm KM curve from Batchelor et al., 2010[60], and fitted standard parametric curves to model PFS.	No TA found that reflected current CNS tumour treatment. The OS was informed by a study in patients with recurrent malignant glioma. The distribution was selected based on goodness of statistical fit.
Pancreas	[redacted]	The company digitised the 5-FU + leucovorin arm in the TA440 submission and fitted an exponential tail to the end of the KM.	Approach used in TA440[59]
Thyroid anaplastic, follicular and papillary	[redacted]	The company digitised the placebo arm KM in the TA535 CS and fitted an exponential tail to the end of the KM.	Approach used in TA535[61]
KM, Kaplan Meier; TA, technology appraisal			

The company followed the same approach to select data sources to inform the pooled historical comparator OS extrapolation as used for PFS extrapolation (see Section 5.2.6.1). Thus, the OS extrapolation is affected by the same issues in terms of bias of unknown direction and magnitude due to potential lack of comparability between the populations in the larotrectinib studies and the historical comparator.

Non-responder comparator

The company followed the same approach to model larotrectinib OS data stratified by response status, as for PFS (see Section 5.2.6.1). Unlike the PFS extrapolation, the assumption of proportional hazards between responders and non-responders appeared to hold. Table 22 summarises the survival models investigated to extrapolate OSS for the response-based model, along with the main justification provided by the company for its use in the non-responder control scenario analysis (p213, CS, and Appendix L).

Table 22 Summary of company's justification for OS extrapolation curves selection – Responder/Non-Responder Analysis

Parametric distribution	Goodness of visual fit	Best statistical fit	Clinically plausible
Exponential	Similar fit to the observed OS data for all distributions, with no distinguishable better visual fit.	[redacted]	[redacted]
Weibull		[redacted]	[redacted]
Log normal		[redacted]	[redacted]
Log logistic		[redacted]	[redacted]
Generalised Gamma		[redacted]	[redacted]
Gompertz		[redacted]	[redacted]

The parametric distribution used in the company's responder control scenario analysis is highlighted in bold.

The company also selected the [redacted] distribution to extrapolate OS for the non-responder control to remain consistent with the base-case assumption for the OS of larotrectinib and because clinical plausibility. All parametric models were considered by the company to have a similar visual fit to the KM curve (see Figure 16), but generated considerably different long-term predictions. The [redacted] distribution had the best statistical fit based on AIC and BIC. The company judged the clinical plausibility of the extrapolated OS curves based on the projected survival estimates shown in Table 23, and considered the [redacted] distribution to generate the only clinically plausible OS estimates.

Figure 16 Parametric model fittings: OS stratified by response status (Appendix L, p466)

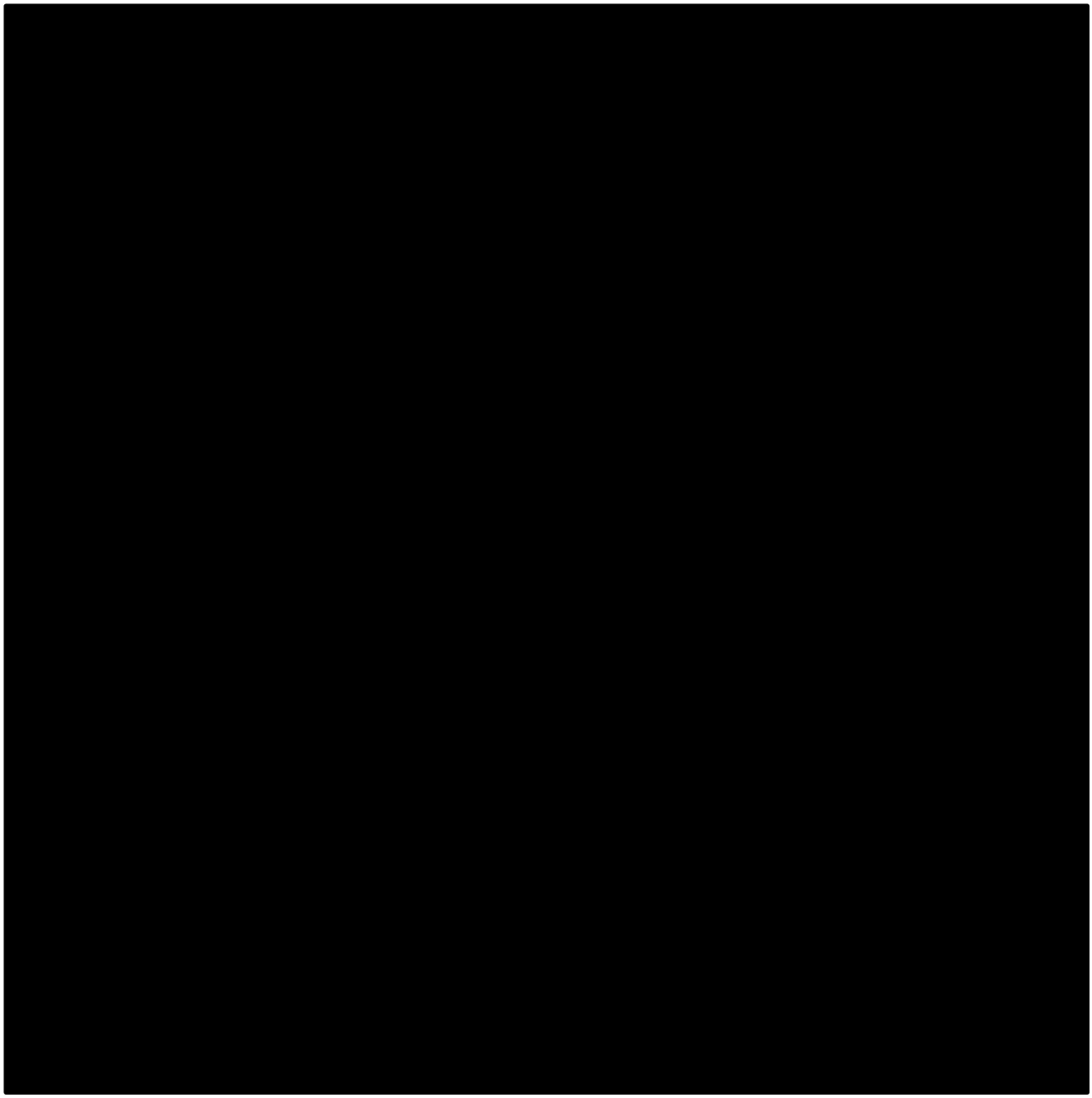
Table 23 Estimated OS by parametric form (years from start of treatment) (Appendix L, p466)

Model	10% alive		1% alive	
	Responders	Non-responders	Responders	Non-responders
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log normal	[redacted]	[redacted]	[redacted]	[redacted]
Log logistic	[redacted]	[redacted]	[redacted]	[redacted]
Generalised Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]

The ERG notes that while the OS data is immature for both responders and non-responders, the responder group has fewer events and heavier censoring (see Figure 17). Thus, the uncertainty in

terms of OS predictions is greater for the responder group compared to non-responders, with a greater number of parametric models predicting clinically implausible predictions. This is similar to what was observed when standard parametric models were fitted to the full larotrectinib integrated efficacy analysis OS data (see beginning of Section 5.2.6.1), and only two survival models were considered by the company to predict clinically plausible survival estimates. The use of a response-based survival model demonstrates that the responders OS is the main contributor to the uncertainty surrounding the larotrectinib extrapolation, which was not explicit from the non-stratified larotrectinib OS analysis. Overall, this suggests that there is important heterogeneity in survival outcomes linked to response status that translates into highly uncertain OS predictions.

Figure 17 OS Kaplan Meir curves by response status (Figure 111, Appendix L)



The ERG also notes that, despite the company's claims that the [redacted] distribution is the only model that generates clinically plausible OS estimates, the [redacted] distribution could be argued to be equally plausible. Figure 18 shows that both survival models have a similar fit to the observed data for non-responders, but the [redacted] distribution imposes more conservative long-term OS predictions for responders than the [redacted] distribution. It can also be argued that, although the OS predictions for responders appear more conservative for the [redacted] compared to the [redacted] distribution (1% of patients alive after [redacted] vs [redacted], see Table 23), the [redacted] may still

overestimate survival for responders, as the post-progression survival gains are still greater than those accrued in PFS. The ERG conducts further analysis to explore uncertainty in the survival extrapolation in Section 6.

Figure 18 Observed and extrapolated OS curves for larotrectinib by response status (adapted from company's model)

5.2.6.5 All-cause mortality

Overall survival in the model for all treatments was adjusted for UK all-cause mortality age and sex specific [65]. This was implemented in the model to correct long survival tails predicted by some parametric distributions. The company states that these adjustments to mortality were not triggered for any of the comparators. The average age and the male to female-ratio was sourced from the larotrectinib integrated efficacy analysis for both larotrectinib and the weighted comparator (p166, CS). A common male to female-ratio was assumed for both larotrectinib and comparator, and tumour site specific average ages were applied for the historical comparator. The model tracked age at each

cycle to determine the background mortality hazard rate within each cycle, which was used where the background mortality hazard was observed to be greater than the specific survival curve.

5.2.6.6 Adverse events

The company's decision model includes adverse events for larotrectinib and the comparator arm. This section describes the proportions of adverse events assumed in further detail. The HRQoL loss and costs associated with treatment of adverse events are detailed in subsequent sections (Section 5.2.7 and 5.2.8, respectively).

Larotrectinib

The company states that only grade 3 to 4 treatment emergent adverse events that occurred in at least 5% of patients in the pooled safety population from the larotrectinib integrated efficacy analysis, are included in the cost-effectiveness analysis. This population from which these adverse events are determined is composed of larotrectinib patients who have received at least one dose of larotrectinib, as of 30 July 2018, regardless of whether evaluable for efficacy (n=137; 82 patients from NAVIGATE, 10 from LOXO-TRK-14001 and 45 patients from SCOUT) (see p108, CS). The company applied in the model a [redacted] adverse event rate for both anaemia and neutropenia

The ERG notes that the adverse events included in the model are not the treatment emergent ones (as stated in the CS), but rather the all causality adverse events occurring in at least 5% of the safety population (see Table 28, CS).

Comparator

Adverse events included for the comparator vary across tumour sites, so as to reflect the different treatments that constitute standard of care for histological type. The adverse events that are applied in the model for each tumour site correspond to those that had a weighted rate of at least 5%. These weighted adverse event rates were calculated by weighting the rates of adverse events in the treatment arm of the source documents (Appendix M, same sources used to inform treatment effectiveness) by the tumour distribution in the larotrectinib clinical trial programme (Table 32, CS). Table 24 summarises the tumour specific adverse events rates for events occurring in at least of the 5% of the patients in the source documents by tumour site. It also reports the weighted adverse event rates, which were used to guide inclusion of adverse events in the model, alongside the tumour distribution weights. The following tumour sites are omitted from the table, as no adverse events were reported as occurring for at least 5% of the patients in the source documents: GIST, STS adults (non-GIST) and thyroid.

In the base case analysis, only anaemia and neutropenia are included in the model, as these are the only adverse events for which the weighted rate was equal or greater than the 5% threshold. The ERG note that after identifying which adverse events to include based on the weighted rates, the model applies the unweighted rates by tumour site to each tumour engine.

The company considers this approach conservative as the weighting approach reduces the chance of a given adverse event meeting the 5% threshold. An alternative assumption is explored by the company in a scenario analysis whereby all adverse events that have occurred in at least 5% patients in the source documents are included (i.e. selection of adverse events for inclusion in the model is based on the unweighted adverse event rates by tumour site). This scenario generates an ICER of [redacted] per QALY favouring the cost-effectiveness of larotrectinib compared to base-case assumptions, as both the costs and QALY loss from adverse events increase for the comparator due to the inclusion of a wider range of adverse events. Given the low impact of alternative assumptions on the adverse event rates on cost effectiveness results, the ERG does not explore the uncertainty surrounding these estimates any further.

Table 24 Comparator adverse event rates by tumour site and weighted

AEs (Grade 3-4)	Weighted	NSCLC	Salivary	Melanoma	CRC	STS children	Breast	Cholangiocarcinoma	CNS	Pancreas
Abnormal liver function	[redacted]							[redacted]		
Anaemia	[redacted]			[redacted]	[redacted]	[redacted]		[redacted]		[redacted]
ALT increased	[redacted]				[redacted]			[redacted]		
Anorexia	[redacted]	[redacted]								
Diarrhoea	[redacted]					[redacted]				
Fatigue	[redacted]	[redacted]			[redacted]			[redacted]	[redacted]	
Febrile neutropenia	[redacted]					[redacted]				
Increase alkaline phosphatase level	[redacted]				[redacted]					
Increase creatinine level	[redacted]				[redacted]					
Increase in total bilirubin	[redacted]				[redacted]					
Infection	[redacted]	[redacted]						[redacted]		
Leukopenia	[redacted]		[redacted]					[redacted]		
Lymphopenia	[redacted]								[redacted]	
Nausea	[redacted]		[redacted]							
Neutropenia	[redacted]					[redacted]	[redacted]	[redacted]		
Pulmonary embolism	[redacted]								[redacted]	
Thrombocytopenia	[redacted]				[redacted]			[redacted]	[redacted]	
Vomiting	[redacted]							[redacted]		
Tumour weights	100%	7%	17%	7%	7%	25%	1%	2%	9%	1%
AEs, adverse events; ALT, Alanine aminotransferase; CRC, colorectal cancer										

5.2.7 Health related quality of life

5.2.7.1 Systematic review of HRQoL studies

Appendix H of the company submission contained a description of the sources searched to identify studies of health-related quality of life (HRQoL) in tumours sites/locations know to harbour NTRK gene fusions (p338-40). The search strategies were not provided in the submission. The company provided the search strategies in Appendix 2 of their response to the points for clarification raised by the ERG.

The following databases were searched: MEDLINE (via PubMed), EMBASE, and the Cochrane Library. The interface/provider was not reported for EMBASE or the Cochrane Library. All database searches took place between May – August 2018 and were updated during January - March 2019. The searches of the Cochrane Library in August 2018 and during January - March 2019 would not have included searches of the Database of Abstracts of Reviews of Effects (DARE), the NHS Economic Evaluations Database (NHS EED), or the Health Technology Assessment (HTA) database, although these databases are listed as being searched on page 339, Appendix H of the submission. These three databases were removed from the Cochrane Library in August 2018. The Cochrane Controlled Register of Trials (CENTRAL) and the Cochrane Database of Systematic Reviews (CDSR) would have been searched via the Cochrane Library. Date limits (2015-2019) were applied to retrieval of conference abstracts in some of the search strategies for EMBASE.

In addition to the database searches, a comprehensive set of sources were searched for unpublished, grey literature. Five clinical trials registers were searched (Clinicaltrials.gov, ISRCTN register, International Clinical Trials Registry Platform (ICTRP), EU Clinical Trials Register and KlinischePrufungen (PhamNet.Bund, AMIS – Offentilicher Teil)). The following Conference proceedings were searched: American Society of Clinical Oncology (ASCO) and the European Society of Medical Oncology (ESMO) and the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) International and European Annual Meetings. Searches of several key international HTA websites were undertaken, including the NICE website. Further details on how the searches for unpublished, grey literature were carried out, the search terms used, the date of the search or any limits applied, were not reported.

14 sets of database searches were undertaken in total, one for each of the tumour sites/locations. The searches were structured appropriately and matched the inclusion criteria specified in Tables 92 and 93 (p.341-43) of the submission. Most of the searches included terms for the population combined with a set of terms to limit to HRQoL studies. In some of the search strategies, terms for the population were limited to advanced forms of the particular cancer/tumour (line #1 in utilities

strategies for non-small cell lung cancer, colorectal cancer, melanoma, pancreatic cancer, glioma, biliary cancer, gastrointestinal stromal tumours, bone sarcoma, and appendix cancer). A reasonable variety of terms and subject headings relating to HRQoL were included in the strategies, however validated search filters designed to retrieve health state utility studies do not appear to have been used. Use of these search filters within the strategies may have improved retrieval of relevant studies.

All searches were conducted appropriately, search lines were combined correctly and no major errors were found by the ERG. Some of the update searches carried out during January - March 2019 in EMBASE and MEDLINE were limited to articles published in 2018 or 2019. Therefore, any relevant studies added to the databases since the last search but with a publication year pre-2018, may not have been identified by the searches presented.

The Company presents a short summary of the identified HRQoL publications, by tumour site, at the end of Appendix H and states the general principles for study selection, when more than one study was identified, in the main body (p. 173-174, CS). If a previously published NICE TA was available, then HRQoL values were extracted from the official document, considering the Committee's preferred assumptions. When no previous appraisal was identified, additional targeted literature searches were conducted (salivary and STS paediatric cancers). If this approach was also unsuccessful, then other tumour sites were used as proxy (e.g. the weighted average of HRQoL values for the other tumour sites was assigned for cholangiocarcinoma). The company states that the results of the systematic review are used to inform the model structure and assumptions (p173, CS).

5.2.7.2 Health State Utilities

The model assigns treatment specific utilities to the progression free and progressed health states, while for death a utility of zero is applied to all patients.

Larotrectinib

Health state utilities for larotrectinib are derived from HRQoL data collected in two of the studies in the larotrectinib clinical trial programme (NAVIGATE and SCOUT). EQ-5D-5L utility measurements of adult patients were collected on the NAVIGATE study (every 8 weeks during the first year of follow-up, and every 12 weeks afterwards), and mapped to EQ-5D-3L utilities[66]. Paediatric HRQoL was collected in the SCOUT trial with two versions of the PedsQL instrument, according to the age of the patient (PedsQL Infant Scales (PedsQL IS) for infants up to 24 months, and PedsQL Generic Core Scales (PedsQL GCS) for children over 2 years of age). Paediatric HRQoL data was collected at pre-treatment screening and on the first day of every 28-day treatment cycle, up to treatment discontinuation. Only data collected with the PedsQL GCS was included in the analysis, as the company could not identify an algorithm that allowed mapping PedsQL IS measurements to EQ-5D-3L utilities.

PedsQL GCS measurements were mapped to EQ-5D-3L utilities [67]. The two samples of EQ-5D-3L were pooled with SCOUT contributing [redacted] assessments from [redacted] paediatric patients and NAVIGATE [redacted]assessments from [redacted] paediatric patients, and the data was analysed by estimating unadjusted mean values and using two alternative regression models (ordinary least squares and Mixed Model Repeated Measures (MMRM)). The company excluded the baseline measurements from the analysis sample, as it considered that the effects of larotrectinib might not yet be felt by the patient. The regression models estimated mean utility scores for patients considered to be either progression free (and on treatment) or having had disease progression. The company justifies regression model selection on the basis of two criteria: the method's ability to reflect the repeated nature of measurements. The MMRM model was selected to inform the health state utility parameters. The company provides details on the data sources and analysis performed inform the health state utilities for larotrectinib in Appendix N.

A summary of the larotrectinib utility values applied in the model for the base-case analysis and scenarios is presented in Table 25, alongside the pooled health state utilities for the comparator.

Table 25 Health state utility values in the model

	Utility value Mean (standard error)			
	Larotrectinib		Weighted comparator	
	Progression-free	Progressed disease	Progression-free	Progressed disease
Base-case	[redacted]	[redacted]	[redacted]	[redacted]
Scenarios – alternative assumptions on larotrectinib				
Utilities estimated from patients aged ≥11 years old in the clinical studies	[redacted]	[redacted]		
Same utility values as for weighted comparator	[redacted]	[redacted]		
Progressed disease utility estimated by applying the ratio between progression free and progressed disease utilities for the weighted comparator to larotrectinib's progression free utility value	[redacted]	[redacted]		

The larotrectinib health state utilities were estimated in a subset of patients (HRQoL sample) from the larotrectinib pooled integrated efficacy analysis ([redacted] out of 102 patients). The company presents a comparison of baseline patient characteristics between the integrated efficacy analysis and the HRQoL in Table 202 (Appendix N), and states that the two populations are similar with the exception of age differences (as patients 2 years old and younger were excluded)

The utility estimates for larotrectinib are affected by several sources of uncertainty. First, the analysis is informed by small numbers of observations, especially for the progressed disease health state

([redacted]assessments in [redacted]patients (p601 and Table 206, Appendix N)). The majority of post-progression utility observations were also from paediatric patients ([redacted]) (p602, Appendix N). The unadjusted mean utility values by health state and cycle in Table 203 (Appendix N) shows that mean utility values are often higher at post-progression than for progression-free at the same cycle for the NAVIGATE trial. Although these values are unadjusted and do not account for correlation between repeated measures, they suggest potential data inconsistencies. Second, the use of mapped utilities introduces uncertainty in the utility estimates. Uncertainty at this level is further compounded by the use of a mapping algorithm [67] to map PedsQL GCS scores to EQ-5D-3L utilities that was validated in healthy school aged children aged 11 to 15 years old. Not only does this require an assumption of comparability between the population in which data (SCOUT trial) was collected and the one used to validate the algorithm, but it also meant that the company had to make further assumptions to use the mapping algorithm since it requires patients to have completed the school domain of the PedsQL GCS. For children not of school age in the SCOUT trial, the company assumed that their school domain would correspond to the average of the remaining questionnaire domain scores. A scenario analysis was performed by the company whereby children younger than 11 years old were excluded from the HRQoL sample (The tumour site specific utilities for the comparator are sourced preferably from previous NICE TAs. The company did not identify suitable NICE TAs for the following tumour sites: cholangiocarcinoma, salivary gland, and paediatric STS. For cholangiocarcinoma, it was not possible to retrieve health-state utility values, and patients were assigned the weighted average HRQoL from the other tumour sites. The utility estimates for salivary gland cancer were sourced from a cost-utility model in head and neck cancer [68] identified in a targeted literature search. For STS paediatric utility estimates were sourced from two cost-effectiveness studies[69, 70] in a general STS population (i.e. not necessarily including children). Details of the comparator-specific utility values are presented in Appendix M of the CS. Tumour site specific health state utility values and data applied in the model for the comparator are shown in Table 26.

Table 26), which resulted on an increased ICER of [redacted] per additional QALY (B3, company's response to points for Clarification) QALY compared to base-case ([redacted] per additional QALY). Third, it is unknown how the pooling of utility estimates from trials with different designs, different study populations and different HRQoL instruments may bias utility estimates.

While the company acknowledged the majority of these limitations, the ERG considers that the potential impact of uncertainty surrounding the post-progression utility estimate was not sufficiently explored. The ERG also questioned why the company had not presented a clinical rationale that justifies the use of differential utility weights for post-progression for larotrectinib and comparator treatments (B4, Points for Clarification). The company replied that this was justified by the more favourable safety profile of larotrectinib compared to chemotherapy, and the potentially irreversible long term toxicities

of chemotherapy. It is worth noting that the comparator therapies for a number of sites were non-active treatments, so their toxicity is likely to be low. The ERG considers that the use of post-progression treatments such as larotrectinib and LOXO-195 in [redacted] of the 34 patients who had progressed (A5, company's response to points for Clarification) may, however, contribute to the high utility value in post-progression. The company varied the health state utility sources in two scenario analyses, one assuming the same utility estimates as for the weighted comparator and the other assuming the same proportional reduction in utility from progression free to progressed disease for larotrectinib as for the weighted comparator. Both analyses had considerable impact on the ICER, with the first increasing it to [redacted] per additional QALY and the second to [redacted] per additional QALY (B3, company's response to points for Clarification), as they both reduce the QALY gains for larotrectinib. The ERG notes that the impact of the alternative HRQoL assumptions in these scenarios may be magnified by the large and potentially too optimistic survival benefit for larotrectinib. The ERG further explores this area of uncertainty in the context of alternative post-progression survival assumptions in Section 6.

Comparator

The tumour site specific utilities for the comparator are sourced preferably from previous NICE TAs. The company did not identify suitable NICE TAs for the following tumour sites: cholangiocarcinoma, salivary gland, and paediatric STS. For cholangiocarcinoma, it was not possible to retrieve health-state utility values, and patients were assigned the weighted average HRQoL from the other tumour sites. The utility estimates for salivary gland cancer were sourced from a cost-utility model in head and neck cancer [68] identified in a targeted literature search. For STS paediatric utility estimates were sourced from two cost-effectiveness studies[69, 70] in a general STS population (i.e. not necessarily including children). Details of the comparator-specific utility values are presented in Appendix M of the CS. Tumour site specific health state utility values and data applied in the model for the comparator are shown in Table 26.

Table 26 Tumour site specific health state utilities in the model

Tumour Site	Progression-free Mean (95% CI)	Progressed Disease Mean (95% CI)	Source
NSCLC	[redacted]	[redacted]	[redacted]
Salivary Gland	[redacted]	[redacted]	[redacted]
Melanoma	[redacted]	[redacted]	[redacted]
Colorectal	[redacted]	[redacted]	[redacted]
Adult STS (GIST)	[redacted]	[redacted]	[redacted]
Adult STS (nGIST)	[redacted]	[redacted]	[redacted]
STS Paediatric	[redacted]	[redacted]	[redacted]

Breast	[redacted]	[redacted]	[redacted]
Cholangiocarcinoma	[redacted]	[redacted]	Weighted average of health state utilities for other tumour sites
Pancreas	[redacted]	[redacted]	[redacted]
Gliomas (CNS)	[redacted]	[redacted]	[redacted]
Thyroid	[redacted]	[redacted]	[redacted]

The company's approach to source the utility estimates for the tumour site specific comparator therapies appears to be consistent with how clinical inputs were informed. The ERG did not attempt to validate the full set of utility estimates against the sources references, but could not, for example, replicate the progression free utility estimate for salivary gland cancer from the values reported in the original publication[68].

5.2.7.3 Adverse Events Utilities

As discussed in section 5.2.6.4, the economic model only considers grade 3 or 4 adverse events which are reported for at least 5% or more of the patients in the larotrectinib clinical trials safety population (n=137, 30th July 2018 cut-off date) or patients in the comparator arm. Tumour-specific adverse utility decrements were then applied to these adverse events. These decrements where sourced from previous NICE Technology Appraisals, where possible, or from the results of the systematic literature review performed by the company (details in Table 41, CS, and Appendix M).

The company states that it was not possible to evaluate the impact of adverse events on HRQoL based on the larotrectinib clinical trial programme data, by including adverse events as a covariate in the regression model, given the small number of events. Thus, the company instead assumes that the disutility from adverse events for larotrectinib is the maximum disutility across all tumour sites for each of the adverse events included (anaemia and neutropenia). This approach is taken to avoid bias favouring larotrectinib, according to the company.

Utility decrements applied in the company's base-case analysis are shown in Table 27, alongside the event rates in the source data and weighted by tumour distribution. The QALY loss from AE events is [redacted] QALYs for larotrectinib and [redacted] QALYs for the historical comparator.

Table 27 Utility decrements applied for each treatment and tumour site in the company's base-case

AEs	Utility decrement (mean)	Source/Justification	AE rate	Tumour weighted AE rate
Larotrectinib				
Anaemia	[redacted]	[redacted]	[redacted]	
Neutropenia	[redacted]		[redacted]	

Melanoma				
Anaemia	[redacted]	[redacted]	[redacted]	[redacted]
Colorectal cancer				
Anaemia	[redacted]	[redacted]	[redacted]	[redacted]
STS paediatrics				
Anaemia	[redacted]	[redacted]	[redacted]	[redacted]
Neutropenia	[redacted]	[redacted]	[redacted]	[redacted]
Breast				
Neutropenia	[redacted]	[redacted]	[redacted]	[redacted]
Cholangiocarcinoma				
Anaemia	[redacted]	[redacted]	[redacted]	[redacted]
Neutropenia	[redacted]	[redacted]	[redacted]	[redacted]
Pancreas				
Anaemia	[redacted]	[redacted]	[redacted]	[redacted]

The company does not explicitly assume any adverse event duration and applies the disutility estimate (adjusted by the adverse event rate) as a one-off QALY loss to all patients at model entrance. It is not clear in the CS whether the disutility estimates sourced from previous NICE TAs and other published sources are disutility weights or have been adjusted for duration of adverse events, i.e. represent QALY losses. If no adjustment for duration of adverse events has been made, then the CS implicitly that assumes that all adverse events were incurred for one full year.

5.2.8 Resources and costs

Resource use and costs are described in detail in the CS (Section B.3.5). These constitute: drug acquisition costs, drug administration costs, costs related to the health states and adverse events. Unit costs were sourced form national drug tariff and fee schedules, and expressed as 2017/18 pound sterling. The approach to estimate comparator costs is the same regardless of how the comparator effectiveness was modelled.

5.2.8.1 Systematic review of resource use and costs

The company conducted systematic reviews of the literature to identify published studies on cost and healthcare resource use data in tumour sites/locations known to harbour NTRK gene fusions that were studied in the larotrectinib integrated efficacy analysis.

The sources searched were reported in Appendix I (p. 350-52). The search strategies were not included as part of the initial CS, but the company later provided them in Appendix 2 of their response to the points for clarification raised by the ERG.

The following databases were searched: MEDLINE (via PubMed), EMBASE, and the Cochrane Library. The interface/provider was not reported for EMBASE or the Cochrane Library. All database searches took place between May – August 2018 and were updated during January - March 2019. The searches of the Cochrane Library in August 2018 and during January - March 2019 would not have included searches of the Database of Abstracts of Reviews of Effects (DARE), the NHS Economic Evaluations Database (NHS EED), or the Health Technology Assessment (HTA) database, although these databases are listed as being searched on page 351, Appendix I of the submission. These three databases were removed from the Cochrane Library in August 2018. The Cochrane Controlled Register of Trials (CENTRAL) and the Cochrane Database of Systematic Reviews (CDSR) would have been searched via the Cochrane Library. Date limits were applied to some of the search strategies presented. A date limit was applied to restrict retrieval of articles published from 2008 onwards in some of the strategies and retrieval of conference abstracts in EMBASE were limited to 2015 onwards.

In addition to the database searches, a comprehensive set of sources were searched for unpublished, grey literature. Five clinical trials registers were searched (Clinicaltrials.gov, ISRCTN register, International Clinical Trials Registry Platform (ICTRP), EU Clinical Trials Register and KlinischePrufungen (PhamNet.Bund, AMIS – Offentilicher Teil)). The following Conference proceedings were searched: American Society of Clinical Oncology (ASCO) and the European Society of Medical Oncology (ESMO) and the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) International and European Annual Meetings. Searches of several key international HTA websites were undertaken, including the NICE website. Further details on how the searches for unpublished, grey literature were carried out, the search terms used, the date of the search or any limits applied, were not reported.

14 sets of database searches were undertaken in total, one for each of the tumour sites/locations. The searches were structured appropriately and matched the inclusion criteria specified in Tables 95 and 96 (p.353-55) of the company submission. Most of the strategies included terms for the population combined with a set of terms to limit to cost or resource use studies. A variety of free text terms and subject headings were included to cover the various terms relating to costs and resource use that are used in the literature. In some of the search strategies, terms for the population were limited to advanced forms of the particular cancer/tumour.

All searches were conducted appropriately, search lines were combined correctly and no major errors were found by the ERG. Some of the update searches carried out during January - March 2019 in EMBASE and MEDLINE were limited to articles published in 2018 or 2019. Therefore, any relevant

articles added to the databases since the last search but with a publication year pre-2018, may not have been identified by the searches.

The company presents the methods for the systematic review in Appendix 2 (company's response to points for clarification) in a format similar to that of the cost-effectiveness systematic review by tumour site (see section 5.13), but without a quality assessment. No conclusions are presented for the systematic review, other than that no published estimates of healthcare resource use were identified for patients with TRK Fusion cancer that could inform the larotrectinib group. The company states that the results of the systematic review are used to inform the model structure and assumptions (p136, CS).

5.2.8.2 Drug acquisition and administration costs

Larotrectinib

The costs of treatment with larotrectinib only constitutes drug acquisition costs. Administration costs were not included, as larotrectinib is administered orally.

The cost of acquisition for larotrectinib in the model is based on an expected list price that varies across the different presentations of the technology. Larotrectinib will be available as 100 mg and 25 mg capsules, and an oral solution (20mg/mL) (Table 28).The acquisition costs of larotrectinib were initially calculated in the model based on a price list per 30 days of treatment rather than the actual expected list price. The company clarified that this was due to the pack size for each presentation not being finalised at the time of submission (see p95, company's response to points for clarification).

On the 23rd July 2019, the company submitted a change to the proposed NHS list price for the paediatric formulations of larotrectinib (25 mg capsules and 20 mg/mL oral solution) with an updated model that reflected these changes (see the company's Larotrectinib price change and model updates). The company did not submit any updated cost-effectiveness results based on this price change until the 5th August 2019, and the updated model only contained the updated deterministic base-case results. Compared to the original base-case results the ICER based on the updated paediatric acquisition costs was reduced from [redacted] to [redacted] per additional QALY. The company also reported in the "Larotrectinib price change and model updates" document a proposed Patient Access Scheme (PAS) price for each formulation that consisted of [redacted]

The proposed list prices (A) contained within the CS and the subsequently updated list prices without (B) and with PAS (C) for each presentation of larotrectinib expected to be available in the market are summarised in Table 28.

Table 28 Larotrectinib's expected cost

Formulation	Pack size	Total mg per pack	Expected cost					
			A. CS proposed NHS list price (June 2019)		B. Proposed NHS list price (July 2019)		C. Proposed PAS price (July 2019)	
			Per pack	Per mg	Per pack	Per mg	Per pack	Per mg
100mg capsules	56	5,600	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
25mg capsules	56	1,400	[redacted]	[redacted]	[redacted]		[redacted]	
Bottle of solution (20mg/ml)	100 mL	2,000	[redacted]	[redacted]	[redacted]		[redacted]	

The cost per cycle of larotrectinib in the model was calculated based on the expected price list (June 2019) and the average daily dose of larotrectinib in the integrated efficacy analysis (p178-179, CS). The company calculated a weighted cost per day for paediatric and adult patients according to the proportion of patients in each cohort in the integrated efficacy analysis (33% paediatric and 67% adults at the start of the model). The cost per day for adults assumed the adult average dose of larotrectinib of [redacted] mg/day, which combined with the initially expected list price for larotrectinib (100 mg capsules) resulted in a price per day of [redacted], which at the start of the model was applied to [redacted] of patients.

For paediatric patients, the model tracked the age and the formulation of larotrectinib that each individual patient in the larotrectinib integrated efficacy analysis received, in order to calculate a weighted cost per formulation. The formulation received by each of 34 paediatric patients was assumed to remain the same as at the start of the cycle unless the patient reached 18 years at which point they would switch to 100 mg capsules. At model entry [redacted] and [redacted] of paediatric patients received the oral solution, 25 mg capsules and 100 mg capsules, respectively (see electronic version of the model). A cost per day for each of the formulations was calculated based on a common paediatric average dose of larotrectinib of [redacted] mg/day (from the larotrectinib integrated efficacy analysis), and the cost per mg for each formulation.

Table 29 Larotrectinib acquisition costs based on initially expected list price (adapted from Table 42, CS)

Larotrectinib formula	Formulation	Pack size	Total mg per pack	Expected cost per pack	Average dose per day (mg)	Expected cost per day	Proportion of patients receiving each formulation at cycle 0
Adults	100mg capsules	56	5,600	[redacted]	[redacted]	[redacted]	[redacted]
Paediatric	100mg capsules	56	5,600	[redacted]	[redacted]	[redacted]	[redacted]
	25mg capsules	56	1,400	[redacted]		[redacted]	[redacted]

	Bottle of solution (20mg/ml)	100 mL	2,000	[redacted]		[redacted]	[redacted]

The company states that the acquisition costs of larotrectinib were applied at the start of the treatment cycle, with no half-cycle correction to account for treatment wastage due to patients discontinuing treatment during the 7 days cycle for any reason and that this is in line with clinical practice (p146, CS). The ERG questioned whether this is in line with clinical practice given that clinical practice is not yet established, and requested additional scenarios to reflect the uncertainty across alternative patterns of treatment discontinuation (see B.12, points for clarification). The company provided two scenarios assuming that patients are prescribed a 2 and 4 weeks treatment supply in which the ICER increased to £[redacted] and £[redacted] per additional QALY, respectively (see B.12, company's response to points for clarification). The company also noted that specialist oncology medicines would not be dispensed in large quantities and that unused products would be returned.

The company calculates acquisition costs for both adults and paediatric patients based on the respective average dose of larotrectinib in the integrated efficacy analysis s. For adults, the average dose of larotrectinib ([redacted] mg/day) in the trials corresponded to [redacted] of the recommended daily dose (100 mg, twice a day). For paediatric patients, the average dose of larotrectinib ([redacted] mg/day) in the trials corresponded to [redacted] of the recommended daily dose (100 mg/m^2; [redacted] mg/daily assuming the paediatric average BSA in the trials of [redacted] m^2). The company explored a scenario analysis to test the impact of adherence to the adult recommended dose (200 mg/day), which caused the ICER to increase from [redacted] to [redacted] per additional QALY (Table 57, CS). The ERG notes that some patients received in the larotrectinib integrated efficacy analysis received lower than therapeutic doses of larotrectinib, as two of the trials (SCOUT and LOXO-14001) had dose-escalation components as part of their safety assessment (see Table 5, CS). The SCOUT trials in paediatric patients, which was a Phase I/II trial, may have been particularly affected by sub-therapeutic doses, with the CS (Table 5) stating that "At least 25/34 paediatric patients received 100mg/m^2 twice daily dosing of larotrectinib". This is in contrast with adult patients for whom at least 66/68 received 100 mg of larotrectinib and the other two patients 150 mg twice daily. Therefore, the ERG considers that the paediatric average dose in the larotrectinib integrated efficacy analysis may underestimate the larotrectinib dose in the paediatric population, and explores alternative assumptions in Section 6.

In the base-case, patients are assumed to continue treatment with larotrectinib until disease progression. However, [redacted] patients in the ePAS2 dataset (n=93) received larotrectinib beyond disease progression ([redacted]months on average). The assumption of treatment until progression is relaxed in one of the company scenario analysis where larotrectinib's treatment duration is based on an extrapolated

time to treatment discontinuation curve ([redacted], considered the most clinically plausible, Appendix L) informed by the larotrectinib integrated efficacy analysis data. Under this alternative assumption the ICER rises to [redacted] per additional QALY (Table 57, CS). The ERG considers that it is uncertain whether patients in the NHS would receive until disease progression or similarly to treatment in the larotrectinib clinical programme trial (i.e. treatment while the clinicians perceives that the patient is still deriving benefit from the drug). The NICE scope and the expected marketing authorisation do not explicitly state any treatment stopping rules. Furthermore, the clinical advisers to the ERG stated that where patients do not have other treatment options after disease progression, they will sometimes continue to receive the pre-progression treatment to avoid a sharp decline in their condition. However, the impact of post-progression treatment on cost-effectiveness estimates needs to be examined, both in terms of costs and health outcomes. The ERG explores this area of uncertainty in Section 6.

Comparator

The company estimates tumour site specific acquisition and administration costs according to the comparator treatments that are included in the historical comparator. Unit costs are sourced from electronic market information tool (eMIT)[73] and British National formulary (BNF) [74], while dose and posology were sourced from the same previous NICE TAs and other published sources used to inform the comparator effectiveness (Table 43, CS). The company states where there were unit costs for multiple presentations of the drugs, the least expensive cost per mg of drug was used to represent unit cost. The base-case analysis assumes no wastage for the comparators. The company also assumes that for tumour locations with no active treatment, both treatment arms receive current standard management and larotrectinib is an add-on therapy. Therefore, no costs of treatment are attributed to the comparator for these tumour sites.

The dose for some of the comparator treatments is dependent on patient BSA. For these treatments, average BSA was sourced where available from the previous NICE TAs and other published sources used to inform the comparator effectiveness. For the two tumour sites (salivary and IFS) for which these sources did not provide an average BSA estimate, the average BSA according to age group from the larotrectinib integrated efficacy analysis is assumed (p179, CS).

Comparators administered orally are assumed to have no administration costs. Administration costs for IV drugs are assumed to be incurred every 7-days cycle and correspond to a simple parenteral chemotherapy administration from NHS reference costs (£228.99 per cycle) (ref) in line with previous NICE TAs (refs) (Table 44, CS).

The company considers a weighted treatment cost (including drug acquisition and administration costs) for comparators consisting of mixed treatment. Total treatment costs per cycle are presented in Table 30.

Table 30 Comparator treatment costs per cycle

Drug	Cost per cycle (week)	Source**
Comparators with no active treatment		
NSCLC	[redacted]	Assumption
Colorectal/Appendix	[redacted]	
GIST	[redacted]	
Thyroid anaplastic, follicular and papillary	[redacted]	
Non-GIST/Bone sarcoma	[redacted]	
Active treatments		
Melanoma	[redacted]	TA357 [48]
Breast	[redacted]	TA423 [57]
Gliomas	[redacted]	Batchelor et al, 2013[60]
Pancreas	[redacted]	TA440[59]
Salivary	[redacted]	Airoldi et al, 2001[47]
STS paediatric/IFS/ Congenital mesoblastic nephroma	[redacted]	Mascarenhas et al, 2010 [56]
Cholangiocarcinoma	[redacted]	Valle et al, 2010[46]

* STS paediatric treatment dosing is irregular from week-to-week (See Table 43, CS for details); ** Source refers to treatment dosing.

Treatment duration was sourced from the same previous NICE TAs and other published sources used to inform the comparator effectiveness. For the comparators informed by NICE TAs, the company assumed the same approach to determine treatment duration. Where other published sources were used, the company states that treatment duration data was extracted, and assumed either a fixed treatment schedule, or point estimate (as published). The company further states that where only maximum treatment durations were reported, the treatment duration was capped by the fixed schedule or the maximum duration for patients that had not yet progressed.

The ERG notes that when sourcing unit costs from the BNF [74], the company did not always select the least costly unit cost for the selected drug formulation. The cost-effectiveness results are not, however, sensitive to this, as the ICER increased marginally compared to the company's base-case when these unit costs were corrected ([redacted] vs [redacted] per QALY gained).

5.2.8.3 Health state unit costs and resource use

Health state resource use and costs also differ across tumour sites for the comparator. Consistent with the approach taken to inform treatment costs, resource use data was extracted preferably from previous NICE TAs and the Committees preferred assumptions were utilised. In the absence of

previous NICE TAs, other published sources were used to inform these parameters (systematic literature reviews described in Section 5.2.8.1). Further targeted searches (not described in the submission) were conducted to inform the resource use estimates for tumour sites for which no previous NICE TAs or other published studies were identified via systematic review. Unit costs were sourced from NHS Reference costs 2017-2018 [75], Monthly Index of Medical Specialties (MIMS)[76], Personal and Social Services Research Unit (PSSRU)[77], BNF [74] and previous NICE guidance[45, 78] (Table 47, CS).

The company considers different 'progression free' and 'progressed' health state costs for patients who enter the state (start-up costs) and those who remain in these states. A start-up cost is defined as the one-time cost of health resources required for assessment and/or treatment initiation when patients enter a health state. The company assumed that the start-up cost is £0 if the source does not mention any resource use details or if health state costs are presented in an aggregate form. The remaining costs for each health state are recurrent costs per cycle.

Costs in the 'death' state are attributed to patients who enter the state at each cycle as a one-off lump sum, and correspond to the costs of delivering end of life care.

Table 31 summarises the health state costs applied in the model for larotrectinib and the historical comparator (by tumour site), alongside sources. The health state cost break-down for each tumour site is detailed in Appendix M of the CS.

Table 31 Health state costs per cycle (adapted from p98, company's response to Points for Clarification, and Appendix M)

	Cost per cycle				**Source**	**Cost**	**Source**
Tumour locations	**Progression-free, start-up**	**Progression-free, per cycle**	**Progressed, start-up**	**Progressed, per cycle**		**Death/End-of-life**	
Larotrectinib	[redacted]	[redacted]	[redacted]	[redacted]	Weighted average of comparators	[redacted]	Weighted average of comparators
NSCLC	[redacted]	[redacted]	[redacted]	[redacted]	TA 374[45]	[redacted]	TA 374[45]
Colorectal/Appendix	[redacted]	[redacted]	[redacted]	[redacted]	TA405[50]	[redacted]	TA405[50]
GIST	[redacted]	[redacted]	[redacted]	[redacted]	TA 488[52]	[redacted]	TA 488[52]
Thyroid anaplastic, follicular and papillary	[redacted]	[redacted]	[redacted]	[redacted]	TA 535[61]	[redacted]	TA 535[61]
STS (Non-GIST)/Bone sarcoma	[redacted]	[redacted]	[redacted]	[redacted]	TA185[54]	[redacted]	Round et al., 2015[79].
Melanoma	[redacted]	[redacted]	[redacted]	[redacted]	TA268[80]	[redacted]	TA268[80]
Breast	[redacted]	[redacted]	[redacted]	[redacted]	TA 423[57]	[redacted]	TA 423[57]
Glioma[1]	[redacted]				TA 23[71]	[redacted]	Round et al., 2015[79].
Pancreas	[redacted]	[redacted]	[redacted]	[redacted]	TA 440[59]	[redacted]	TA 440[59]
Salivary	[redacted]	[redacted]	[redacted]	[redacted]	TA 490 [81]	[redacted]	TA 490 [81]
STS paediatric/IFS/Congenital mesoblastic nephroma	[redacted]	[redacted]	[redacted]	[redacted]	Amdahl et al., 2014[70],	[redacted]	Round et al., 2015[79].
Cholangiocarcinoma	[redacted]	[redacted]	[redacted]	[redacted]	Weighted average of other comparators	[redacted]	Round et al., 2015[79].
1, Not cost per cycle, but total costs applied as a one-off cost to glioma health states. The source TA reported monitoring cost over the treatment period by a fixed schedule and did not fit a per-cycle calculation.							

The majority of 'progression-free' or 'progressed' costs were informed by the same NICE TAs or published sources from which treatment effectiveness data was sourced. The company assumes that although the comparator treatment in TA 23 (procarbazine in monotherapy) was not reflective of the standard of care for CNS tumours and, therefore, not appropriate to inform effectiveness estimates, it was suitable to inform health state costs. Health state costs for melanoma were informed by an earlier NICE TA [80] rather than the one informing treatment effectiveness [48], as the former was cited by TA 357[48] as source of these costs. For salivary cancers, the company assumes that patients with head and neck cancer have similar health state costs as in salivary cancers and sources costs from TA 490[81]. As the treatment effectiveness data source for STS paediatric (including IFS and congenital mesoblastic nephroma) did not report health state resource use or costs, the company sourced health state costs from Amdahl et al., 2014 [70], a cost-effectiveness in advanced soft tissue sarcoma patients from the UK perspective. This study does not, however, appear to include paediatric patients. The company states that assuming costs in children would be similar to adults in STS was considered a conservative assumption at the clinical validation exercise (p525, Appendix M). The ERG notes that the 'progression-free' and 'progressed' costs in the model are similar for STS in paediatric and adult patients.

The company did not identify any sources of health care resource use data to inform the progression-free' or 'progressed' costs for cholangiocarcinoma, and assumes these costs to be a weighted average of the corresponding costs for all other tumour site comparators.

End-of life costs applied to patients who newly transition to the 'death' state were sourced where available from the same source as for the other health state costs. Where the source used to inform other health state costs did not provide estimates of End-of-Life care costs (STS (Non-GIST)/Bone sarcoma, STS paediatric/IFS/congenital mesoblastic carcinoma, and cholangiocarcinoma), these were sourced from a modelling study identified through a targeted literature review (non-described)[79]. The company states that End-of-Life care costs are included in the model even if not included in the NICE TA informing the other health state costs for each tumour site, so as to *"align with the resources accounted for in other TAs for consistency"* (p517, Appendix M).

The company did not identify any source to inform health state costs for larotrectinib and, therefore, these costs for larotrectinib are assumed equal to the weighted average of the comparators costs, using the tumour site distribution in the larotrectinib integrated efficacy analysis as weights. The company validated this approach with UK clinicians as part of the clinical validation (section B.3.10, CS), and states that the clinicians considered this an appropriate assumption given the data available, and expected this would likely be conservative, and overestimate health care resource use for larotrectinib.

5.2.8.4 Adverse events

Section 5.2.6.6 describes the adverse events included in the model. The costs of treating adverse events (adjusted by the corresponding adverse event rates) were applied as a one-off cost to all patients at the start of the model. The adverse events included in the company's base-case analysis for both larotrectinib and the historical comparator are anaemia and neutropenia. The company assumes the same unit cost for each adverse event irrespective of tumour site, and states that this is in line with previous NICE TAs. Table 32 summarises the adverse events costs and rates applied in the company's base-case analysis for larotrectinib and historical comparator.

Table 32 Adverse event costs and rates applied in the company's base-case analysis

Adverse event	AE rates		Cost per event	Source
	Larotrectinib	Historical comparator		
Anaemia	[redacted]	[redacted]	[redacted]	TA405
Neutropenia	[redacted]	[redacted]	[redacted]	NHS reference costs 2017/18

The unit cost for anaemia corresponds to an activity weighted average of NHS reference costs 2017/18 [75]healthcare resource group (HRG) codes [redacted] and was based on the cost used on NICE TA 405 [50] For neutropenia the company assumes the cost of [redacted].

5.2.8.5 NTRK fusion testing

The CS did not include the test of testing for NTRK fusions in determining the cost-effectiveness of larotrectinib. This was omitted on the basis that including the cost of testing for NTRK fusions in the CE analysis is not in line with the NICE process guide and that WGS would be available for all tumour types, as part of routine cancer care as specified in the NHS Long Term Plan. The ERG requested that the company provide a scenario that models the cost of testing for NTRK (B15, Points for Clarification), but the company did not provide it.

A variety of alternative testing algorithms have been proposed for the identification of NTRK fusions, which recommend that the testing approach should be tailored according to the prevalence of NTRK fusions, and depending on the provision of genomic testing currently available[30, 31]. In Section 2.2.2.1, the ERG provides details on the current provision of NTRK fusion testing and the potential testing approaches that may be implemented in the NHS. In section 6, the ERG explore the addition of testing costs to larotrectinib.

The ERG notes that there are also issues of capacity to roll out genomic and histological testing for NTRK fusions. The introduction of vast numbers of tests to NHS pathology services may have practical and infrastructural implications to service provision that may limit the roll-out of testing unless significant investment is made to increase testing capacity. Furthermore, the CS does not consider the diagnostic accuracy of tests for NTRK fusion, which may also impact on the clinical and cost-effectiveness of larotrectinib. The ERG exploratory analyses do not address these matters, but consider both are considered an area of uncertainty.

5.2.9 Cost effectiveness results

This section presents the company's cost-effectiveness results as reported in the original CS. On the 5th August 2019, the company submitted an updated CS and electronic version of the model that reflected a reduction to the proposed NHS list price for the paediatric formulations of larotrectinib (25 mg capsules and 20 mg/mL oral solution) (see Section 5.2.8.2). The ERG was not able to validate the full set of results in the updated CS or model, as these were submitted less than three weeks before the ERG report was due for submission. Thus, this section focus on the results in the original CS, and only presents updated results for key analyses (deterministic and probabilistic base-case), so as to illustrate the impact of the price reduction for paediatric formulations of larotrectinib.

The company also submitted on the 5th August 2019 cost-effectiveness results including a PAS price, which consisted [redacted] for the three larotrectinib formulations. The ERG was also unable to validate the full set of results with the PAS price. The results for key analyses are reported in Appendix 10.2.

5.2.9.1 Base-case results

The company's base-case results in the original CS and the updated CS (with price reduction for paediatric formulations of larotrectinib) are summarised in Table 33. These results are exclusive of the confidential PAS.

Table 33 Company base-case deterministic cost-effectiveness results (adapted from Table 53, CS) exclusive of the confidential PAS

	Original CS			Updated CS		
	Larotrectinib	**Comparators**	**Incremental**	**Larotrectinib**	**Comparators**	**Incremental**
Treatment cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progression-free costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressed disease costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Adverse event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
End of life care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progression-free life years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressed disease life years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total life years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progression-free QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressed disease QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse events	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ICER (per QALY)			[redacted]			[redacted]

The cost difference between larotrectinib and comparator is largely driven by the costs of treatment, larotrectinib treatment costs on average [redacted] higher than the comparator. When the updated list price for larotrectinib is considered, mean incremental treatment costs reduce by [redacted], and the ICER decreases to [redacted] per QALY gained. The health state costs for 'progression-free' and, particularly, for 'progressed' are also higher with larotrectinib than the comparator (mean incremental cost difference of [redacted] and [redacted] for 'progression-free' and 'progressed', respectively), as the company predicts that larotrectinib extends both progression free and overall survival. The end-of-life costs are, however, lower for larotrectinib compared to the pooled weighted comparator ([redacted] vs [redacted]). Larotrectinib patients are predicted to die later in time than patients receiving the comparator, and, therefore, the end-of-life costs for larotrectinib are more affected by discounting (lower present value for larotrectinib costs). The cost of adverse events is low for both larotrectinib and comparator, but slightly higher for the latter ([redacted] vs [redacted]), reflecting the higher incidence rate of adverse events for the comparator.

The majority of QALY gains for larotrectinib were generated within the 'Progressed' state ([redacted] of mean incremental QALYs), after treatment with larotrectinib is assumed to have stopped. In Section 5.2.6, the ERG highlighted that the considerable survival gains for larotrectinib compared to the historical comparator may be driven by post-progression treatments such as LOXO-195 and larotrectinib, which were delivered to a proportion of patients in the larotrectinib integrated efficacy analysis, but are unavailable for the historical control. The ERG also noted the uncertainty surrounding the OS extrapolation for larotrectinib, which is based in immature and sparse data. The post-progression health outcomes for larotrectinib are both a driver of cost-effectiveness and a key area of uncertainty.

The QALY loss from adverse events is slightly higher for larotrectinib than for the historical comparator (mean incremental difference [redacted] QALYs), despite the comparator having higher incidence rates for both adverse events included in the base-case analysis. This is due to the assumption that the utility decrements applied to larotrectinib correspond to the maximum disutility across all tumour sites for each of the adverse events included.

5.2.9.2 Sensitivity analyses

Deterministic sensitivity analysis

The CS presents a series of one-way deterministic sensitivity analyses to assess the impact of varying model input parameters on the ICER. The full list of parameters and the range over which they are varied is reported in Table 197 (Appendix M). The company varied the model inputs +/- 20% over the point estimate, but does not state why this range was selected. The results of the one-way deterministic sensitivity analyses are presented in a tabular form (Table 54, CS) and in a tornado diagram (Figure 19) for the 20 parameters that the ICER was most sensitive to. These results are exclusive of the confidential PAS.

Figure 19 Deterministic sensitivity analysis tornado diagram (Figure 32, CS)

The ERG notes that the list of parameters in Table 197 (Appendix M) includes parameters that are not utilised in the model, as for example 'NTRK total patients', and the QALY loss due to adverse events of larotrectinib (adjusted by the corresponding adverse event rates) was mislabelled as 'adverse event disutility (trial based)'. The tornado diagram also reports results for one parameter, 'model adult start age' that is not listed in Table 197 (Appendix M).

The company considers the cost-effectiveness results to be most sensitive to the scale (ICER range: [redacted] per QALY gained) and shape (ICER range: [redacted] per QALY gained) parameters of the parametric distribution used to extrapolate OS for larotrectinib, and that this is expected given the immaturity and small number of events in the observed data. The results are also sensitive to variation in the shape parameter of the STS paediatric comparator OS curve (ICER range: [redacted] per QALY gained), but the company does not discuss this finding. The variation of other parameters resulted on a change in the ICER of less than £2,500 per QALY.

Scenario analysis

Scenario analysis in the CS

The company conducted a number of scenario analyses to test the base-case main assumptions. The key scenario analyses are described in Table 34.

Table 34 Summary of scenario analysis (Table 52, CS)

Scenario analysis	Scenario description	Justification
Discount rate	Replace 3.5% discount rates for cost and outcomes with 1.5% rate	Investigate the long term uncertainty and impact of discounting
Utility	Replace larotrectinib utilities with alternative utility model: Revised patient pool excluding paediatrics under age 11	Investigate the uncertainty surrounding the utility values derived from the small patient numbers
Drug costs	For adults, base case will use actual trial dose, and scenario will test the full daily dose ([redacted] will be cost out as 200 mg)	Investigate the impact of 100% adherence to treatment dose
	Use of larotrectinib TTD curves	To test the impact of alternative treatment assumptions
Time horizon	10 years, 20 years	Investigate impact of using shorter time horizon
Health state costs	Replace tumour location specific health state costs with consistent costs for every tumour location	Investigate the impact of the inconsistency and uncertainty of health state costs across tumour locations
	Remove health state costs if not reported in the source documents	Investigate the outcomes if model follows the original sources exactly instead of making assumptions to fill data gaps
Survival	Different comparator and larotrectinib survival curves where possible (PFS, OS)	Investigate the uncertainty and sensitivity of alternative parametric fits to survival curves
	Alternative comparator survival data for STS non-GIST; pazopanib (following clinical validation)	
AEs	Alternative AE inclusion criteria; all AE with individual 5% rates reported in source publication	Investigate the uncertainty of adverse event rates for the historical comparator.
NTRK prognosis	Results from the SLR conducted to consider evidence on NTRK prognosis	Used to explore how a prognostic effect of being NTRK positive may affect CE results.
Alternative modelling methods	Stratified responder/non-responder analysis, with non-responder representing the comparator arm	Investigate the uncertainty of the overall results using alternative survival modelling methods to represent efficacy.
	Use of GMI as relative risk applied to larotrectinib health outcomes to represent a previous line of therapy comparator. See section B.2.6.	

The results of the scenario analysis are reported in Tables 57 to 61 (CS). The originally submitted version of the company's model did not allow the full set of scenarios reported in the CS (B1, Points for Clarification) to be performed by the ERG, and the ERG could not replicate some of the company's scenario analysis results due to reporting errors in Table 57 (B2, Points for Clarification). Furthermore, the results for the scenarios listed in Table 34 were not all reported in the CS. The company subsequently submitted a fully functioning version of the model and an updated version of Table 57 in the CS (Table 1, p74-79, response to Points for Clarification).

The cost-effectiveness results are most sensitive to the choice of alternative survival functions to extrapolate the OS and PFS of larotrectinib, as shown in Table 35 (exclusive of the confidential PAS). The mean total QALYs (and LYG) vary widely for the alternative OS distributions for larotrectinib, while mean costs are considerably more stable. This is because all the alternative survival functions, except the exponential, predict greater survival benefits in the 'progressed' health state than the base-case assumption, and the cost per cycle in post-progression is lower than in pre-progression as it does not include the cost of treatment with larotrectinib. Therefore, all alternative OS scenarios result in ICERs lower than the base-case (per QALY) except when the exponential is applied (per QALY). This illustrates the impact of uncertainty in the larotrectinib OS extrapolation, on the cost-effectiveness estimates.

The selection of functional form to extrapolate PFS data for larotrectinib also has considerable impact on the cost-effectiveness results with the ICER varying between and per QALY gained across scenarios. The impact on cost-effectiveness is driven by the expected total costs for larotrectinib, which vary between and per QALY gained. In these scenarios, the average time spent in 'pre-progressed' varies considerable (to LYG) even if mean total life-years are stable, and therefore, so does the time on treatment with larotrectinib (the main component of costs in pre-progression). Although the larotrectinib PFS data is more mature than the OS data, it is still very uncertain, and the cost-effectiveness results are sensitive to alternative extrapolation assumptions for both OS and PFS of larotrectinib.

Table 35 Scenario analysis results for alternative survival models for larotrectinib OS and PFS

Scenario	Description	Larotrectinib			Historical comparator			ICER (per QALY)
		Costs	QALYs	LYG	Costs	QALYs	LYG	
	Base case results							
14	Larotrectinib OS - Exponential							
15	Larotrectinib OS - Gompertz							
16	Larotrectinib OS - Log-logistic							
17	Larotrectinib OS - Log-normal							
18	Larotrectinib OS - Gen Gamma							
19	Larotrectinib PFS - Exponential							
20	Larotrectinib PFS - Gompertz							
21	Larotrectinib PFS - Log-logistic							

Scenario	Description	Larotrectinib			Historical comparator			ICER (per QALY)
		Costs	QALYs	LYG	Costs	QALYs	LYG	
22	Larotrectinib PFS - Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
23	Larotrectinib PFS - Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

The company concludes that the base case ICER is robust to alternative OS and PFS extrapolation models for the historical comparator, as it varies between [redacted] and [redacted] per QALY gained across scenarios 24 to 71 (Table 1, p74-79, response to Points for Clarification). The company also explored the use of an alternative source of survival data to inform PFS and OS for the STS adult (non GIST) tumour histology [82] proposed by the company's clinical advisers (Section B.3.10, CS) in Scenario 13 (Table 1, p74-79, response to Points for Clarification), but this had a modest impact on the ICER ([redacted] per QALY gained).

The company explored two alternative approaches to model the effectiveness of the comparator, the first of which uses data from non-responder group in the larotrectinib integrated efficacy analysis as a proxy for patients not receiving an active treatment (see Section 5.2.4). In Section 5.2.6, the ERG describes how survival models were selected to extrapolate OS and PFS for non-responders. In this scenario, [redacted] distributions were chosen for both PFS and OS, and used to inform health state transitions for each tumour site specific comparator engine. This scenario retained all other base-case assumptions and data sources. Results are presented in Table 36 (exclusive of the confidential PAS).

Table 36 Alternative modelling methods: using non-responding patients as a control (Table 58, CS)

Technologies	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER (£/QALY)
Larotrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Comparator	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

In this scenario comparator survival outcomes are lower than under base-case assumptions (mean [redacted] vs [redacted] LYG), resulting into greater mean incremental QALYs for larotrectinib compared to comparator ([redacted] vs [redacted] QALYs for the scenario and base-case respectively. The average total costs for the comparator are lower than in the base case ([redacted]) due to the shorter expected survival. The ICER for this scenario is [redacted] per additional QALY. The company did not explore the use of alternative parametric distributions to extrapolate PFS and OS in this scenario, despite the considerable uncertainty associated with the survival data collected within the larotrectinib integrated efficacy analysis (see Section 5.2.6).

The second scenario that explored an alternative modelling approach for comparator PFS and OS was the use of previous line of therapy to adjust control The company estimated the GMI multiplier using two different samples of the larotrectinib integrated efficacy analysis: i) 52 patients for whom previous treatment was for metastatic disease setting); and ii), all 73 patients who had received at least one prior therapy. The company considers that using all patients who had had a previous line of treatment was highly conservative, as patient's baseline status is likely to decline over the course of disease. The restriction of the sample to only patients who were already metastatic at the previous line of treatment attempts to make patients at a previous line of treatment more comparable to the same patients after initiating treatment with larotrectinib (p149 and p215, CS). Table 37 shows the results of the scenario analysis for the two GMI multipliers calculated by the company (exclusive of the confidential PAS).

Table 37 Alternative modelling methods: using previous line of therapy to adjust control (Table 59, CS)

GMI source	GMI value	Larotrectinib			Historical comparator			
		Costs	QALYs	Life years	Costs	QALYs	Life years	ICER
Base case		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
All patients who received a prior systemic therapy (mean GMI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
All patients receiving prior systemic therapy in the metastatic disease setting (mean GMI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Both scenarios result in a higher ICER compared to the base-case analysis, as the incremental mean QALYs between larotrectinib and the comparator are reduced under the scenarios' assumptions. The company considers this finding to be realistic, as patients on a previous line of treatment would potentially have less advanced disease than a direct comparison with their current standard of care. Furthermore, the company consider that the results of the two alternative comparator modelling approaches (i) Non-responders as control and ii) using previous line of treatment to adjust control) provide further evidence that the base-case analysis results are robust to alternative assumptions on the survival outcomes of the comparator.

The ERG considers that the company ran extensive scenario analyses to explore the impact of alternative comparator effectiveness assumptions on the cost-effectiveness results. The comparator survival outcomes appear to be fairly consistent across the range of alternative survival models applied for the historical control, and alternative approaches to generate survival outcomes for the comparator. In contrast, the cost-effectiveness results appear to be very sensitive to alternative survival extrapolation for larotrectinib, as this is informed by immature data and low number of events (particularly for OS).

There is also uncertainty as to the extent to which the OS benefits for larotrectinib may be driven by post-progression treatments, and that may lead to confounding of larotrectinib's treatment effect when comparison is established against the historical control. This important area of uncertainty has not been sufficiently addressed by the company. The ERG conducts further scenario analyses in Section 6 to explore the potential consequences to cost-effectiveness of uncertainty in the extrapolation of survival and the magnitude of post-progression survival benefits.

Results for the remaining scenarios in the CS are shown in Table 38 (exclusive of the confidential PAS).

Table 38 Other scenario results in the CS (adapted Table 1, response to clarification, and Tables 60 & 61, CS)

Scenario number	Scenario category	Description	Larotrectinib			Historical comparator			ICER (£/QALY)
			Costs	QALYs	LYG	Costs	QALYs	LYG	
0		Base case results	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
1	Discount rate	Replace 3.5% discount rates for cost and outcomes with 1.5% rate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2	Utility	Replace larotrectinib utilities with weighted comparator utilities for progression-free health state	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
3		Replace larotrectinib utilities with alternative utility model: Revised patient pool excluding paediatrics under age 11	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
4		Replace larotrectinib utility for progressed disease state only with literature based relative reduction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
5	Drug costs	Full daily dose for larotrectinib adults (200mg)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
6		Larotrectinib TTD curve for time on treatment (Weibull)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
7		Larotrectinib TTD curve for time on treatment (Exponential)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
8	Time horizon	10 year time horizon	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
9		20 year time horizon	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
10	Health state costs	Replace tumour location specific health state costs with consistent costs for every tumour location; weighted average of all tumour location sources	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
11		Remove health state costs if not reported in the source documents for each tumour location	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12	Adverse events	Alternative AE inclusion criteria; all AE with individual 5% rates reported in source publication	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
-	NTRK prognosis	Survival adjustment for NTRK+ only applied to colorectal cancer	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
-		Survival adjustment for NTRK+ applied to all tumour sites where NTRK incidence (<25%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario analysis requested by the ERG

The ERG also requested that the company reported the following scenario analyses:

1. Treatment effect adjusted to reflect current NHS clinical practice, i.e. excluding LOXO-195 and other therapies not currently available and/or recommended in the NHS (B5, Points for Clarification)
2. 'Replace larotrectinib utility for progressed disease state only with literature based relative reduction' (B7, Points for Clarification) – company scenario included in the model, but not described in the CS
3. 'Use of weighted comparator health state utilities for the larotrectinib arm' (B8, Points for Clarification) – company scenario included in the model, but not described in the CS
4. Alternative wastage assumptions (B12, Points for Clarification) based on a:
   4.1. Two weekly prescribing pattern
   4.2. Four-weekly prescribing pattern
5. Inclusion of costs of testing for NTRK fusions

The company did not submit any scenario where the treatment effect of larotrectinib was adjusted to reflect current practice, as it was not considered appropriate *"as it could potentially compromise the validity of the ITT approach adopted in the larotrectinib trials"*. Furthermore, the company noted that patients in the NHS might have compassionate access to medicines not yet licensed or drugs approved via a system of individual funding requests (B5, company's response to Points for Clarification). The ERG highlighted in Section 5.2.6 that patients in the historical comparator would not have had access to either larotrectinib or LOXO-195, and noted concerns that the large post-progression survival gains predicted for larotrectinib in the company's base-case are likely to be at least partly driven by these treatments.

The scenario including the costs of testing for NTRK fusions was also not provided by the company, who argued that WGS will be available in the NHS as part of routine cancer care and that the costs of NTRK gene fusion testing would not be unique to larotrectinib (B15, company's response to Points for Clarification).

Results for the two utility scenarios (point 2 and 3) are presented in Table 38 (exclusive of the confidential PAS), alongside the other utility scenario considered in the CS. The company stated that the results of these scenarios were not initially included in the CS, because the utility data from the larotrectinib clinical trial programme was considered more representative of HRQoL of the modelled population than literature derived estimates. The scenario in point 2 is however mislabelled, as it effectively assumes that the larotrectinib health state utilities in both pre-progression and post-

progression are the same as for the historical comparator (using the literature source estimates by tumour site and weighting them by tumour distribution). Under this scenario, the mean total QALYs for larotrectinib are reduced by 22% and the ICER rises to [redacted] per QALY gained. The other health state utility scenario (point 3) adjusts larotrectinib progressed disease utility by applying the ratio between progression free and progressed disease utilities for the weighted comparator to larotrectinib's progression free utility value. This scenario also results in a reduction in mean total QALYs accrued by larotrectinib (10%) with the ICER increasing to [redacted] per additional QALY.

The company conducted the two scenarios requested by the ERG to further examine the potential impact of larotrectinib drug wastage on cost-effectiveness (point 4). These scenarios assume that patients receive a supply of larotrectinib every i) 2 weeks, or ii) 4 weeks, and that any unused drug due to treatment discontinuation is wasted. Full results for these scenarios are presented in the company's response to points for clarification (p96). The ICER increased to [redacted] and [redacted] per QALY gained when a 2 weeks and 4 weeks prescribing pattern was assumed, respectively.

Probabilistic sensitivity analysis

The company performed a probabilistic sensitivity analysis (PSA) where parameters were sampled probabilistically from distributions based on 1,000 simulations. Probabilistic distributions were fitted to the following groups of parameters:

- Model characteristics (discount rate, time horizon, age)
- Parametric survival models
- Adverse event costs and disutilities
- Health state utilities
- Health state costs

The full list of parameters varied in the PSA are listed in Table 158 (Appendix M). The company states that disutilities and health state costs were assumed to follow a normal distribution, while beta distributions were fitted to utility estimates. Multivariate normal functions were fitted to the parametric survival models used to extrapolate OS and PFS for larotrectinib (with the exception of the exponential distributions, which used a normal distribution) using an Excel function developed by the Centre for Bayesian Statistics in Health Economics [83]. Univariate normal distributions were fitted to the tumour site specific PFS and OS parametric survival models implemented to extrapolate the specific PFS and OS for the comparator. Therefore, the simulated values for OS and PFS of the comparator do not account for the correlation between parameters in each survival function, and may be biased.

The company does not justify the choice of probability distributions selected for each type of parameters. The ERG notes that the use of the normal distribution to obtain random draws from cost data may lead to the use of implausible values in the simulations as the normal distribution is not bound

at zero and may return negative values (unlike the more commonly used gamma distribution) [84]. The use of the normal distribution for disutilities may result on positive values (i.e. QALY gains from adverse events).

The company also does not justify the selection of parameters that are varied in the PSA. The ERG notes that the company includes discount rates on costs and QALYs, the time horizon and age on the PSA. These parameters are not affected by parameter uncertainty and reflect either normative judgments (e.g. discount rates) or variability (e.g. start age), and, therefore, should not have been implemented stochastically in the PSA.

Furthermore, the company does not detail in the CS the sources used to inform standard errors of parameters varied in the PSA. On examination of the electronic version of the model, the standard errors for a large number of parameters were calculated by assuming the parameter is normally distributed and assuming that the 95% confidence intervals correspond to -/+ 10% of the mean value of the parameter. The majority of parameters for which this appears to have been assumed are the ones for the tumour site specific comparator, and include the survival, utility and cost parameters. It is not clear whether such value ranges actually represent the true uncertainty around these parameters.

Table 39 summarises the probabilistic cost-effectiveness results as reported in the original and updated CS (with price reduction for paediatric formulations of larotrectinib, exclusive of the confidential PAS).

Table 39 Company base-case probabilistic cost-effectiveness results (adapted from Table 55, CS)

	Original CS			Updated CS		
	Larotrectinib	**Comparators**	**Incremental**	**Larotrectinib**	**Comparators**	**Incremental**
Mean costs (£)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean LYG	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ICER (per QALY)			[redacted]			[redacted]

The probabilistic ICER is higher than that from the deterministic analysis for the original and the updated CS ([redacted] and [redacted] per QALY).

The company also presents the cost effectiveness acceptability curve and the probabilistic results in the cost-effectiveness plane (Figures 34 and 35, CS). For the expected list price of larotrectinib in the original CS, larotrectinib has a [redacted] probability of cost-effectiveness at £50,000 per QALY gained

(Table 56, CS). The probability of larotrectinib being cost-effectiveness at £50,000 per QALY gained is reported in the updated CS as [redacted] (Table 56, updated CS).

5.2.10 Model validation and face validity check

The company's validation and face validity check included the following elements:

1. Comparison of larotrectinib's observed and model predicted PFS and OS outcomes
2. Scoping of the economic model
3. Validation of the economic model
   3.1. Clinical validation
   3.2. Economic validation

The first element consisted of a comparison of the observed percentage of patients who were event free at 3, 6, 12 and 24 months in the larotrectinib clinical trial programme for PFS and OS against the model predicted outcomes (Table 62, CS). The company noted that the model and observed outcomes were similar, with a slight underestimation of time-to-event outcomes for larotrectinib at later time points in the model.

The scoping of the economic model included a review of previous NICE TA involving histology independent treatments conducted by an independent health economic and outcomes research consultancy. A consultancy with expertise in economic analysis [redacted] is also stated to have provided *"economic analysis and insight into best modelling practices and advised on the modelling structure and methodology"*. The company also held a formal advisory board meeting of 8 unnamed academic health economists and statisticians whose advice was sought regarding the modelling methodologies that were used to inform the analysis.

The clinical validation consisted of interviews with a number of UK clinical experts to validate approaches, data sources and assumptions. The interviews were led by a Bayer health economist and a Bayer clinician and facilitated by a medical communications agency. The company states that two data sources used to inform the comparator treatments were questioned by the experts, namely the source used to inform the adult STS (non GIST) and the salivary gland cancer sites. One scenario analysis was conducted using an alternative source was proposed for the efficacy data for the comparator arm (see Section 5.2.10.2) adult STS (non GIST). For salivary gland cancer, an alternative standard of care was suggested by the experts, but the company could not identify any data to inform a sensitivity analysis. The company states that the assumption that a weighted average of comparator health state resource use would be reflective of resource use for larotrectinib was reasonable but potentially conservative as larotrectinib is a targeted therapy.

An initial validation of the economic model is stated to have been undertaken by health economists who had not developed the model and included: checks on the selection and results of different modelling options, calculation spot checks, cross checks against source data and extreme value scenarios to check if the model behaved logically. This validation exercise identified small inaccuracies which were corrected. The company states that a parallel final model validation was performed by an independent health economic and outcomes research consultancy, consisting of a review of the analytical approach to determine whether it was fit for purpose, a number of quality control checks and face validity checks of model outcomes.

The ERG notes that despite the two validation exercises conducted by the company, the version of the economic model submitted did not allow performing the PSA simulations, and the Visual Basic for Applications macro for the one way deterministic sensitivity analysis did not generate results for some of the parameters (B1, company's response to points for clarification). These and other minor errors were corrected by the company and a new version of the model was submitted alongside the response to points for clarification. All modifications made by the ERG to the company's model were conducted on this version of the model (submitted on the 8th July 2019).

At a later stage, the company changed the pricing structure for larotrectinib (see Section 5.2.8.2) and submitted an updated model. Following, two further versions (with and without a PAS over the expected list price for larotrectinib) of the model were submitted on the 5th August 2019. The ERG was not able to check and validate any of these versions of the model, given their late submission. The ERG updated the model version submitted on the 8th July 2019 with the most recently proposed list price for larotrectinib and with the proposed PAS, and was able to replicate the company's results reported in the company's 'Larotrectinib price change and model updates' document.

5.3 Conclusions of the cost effectiveness section

The data available to model the cost-effectiveness of larotrectinib for *NTRK* fusion–positive solid tumours, presents a significant challenge. The analysis of PFS and OS for larotrectinib, given small numbers and data immaturity, means that estimates are subject to considerable uncertainty and are likely to be sensitive to assumptions made to extrapolate the integrated efficacy analysis results over the model time horizon. In addition, the absence of a control group for larotrectinib makes estimation of the treatment effect difficult. The company sought to explore alternative scenarios to generate a comparator arm, however, the ERG's general view is that all approaches have limitations and may result in biased estimates of treatment effectiveness. The choice of the historical comparator for the base case analysis is likely to be subject to confounding bias, due to the unknown NTRK status of patients in the comparator studies. In generating estimates of OS, the ERG also noted that the long post progression survival benefits may be driven to some extent by treatments available in the

larotrectinib arm that are not available in routine practice, i.e. are not available to the historical comparator. The response-based model is less affected by confounding from subsequent lines of treatment and imbalance of patients characteristics. The previous line of treatment control is also a valid approach to reduce the confounding of treatment effect, however it was not appropriately applied in the company submission.

The company present a single ICER across all tumour types. The ERG's view is that this potentially conceals significant variation in the tumour specific ICERs, driven by a combination of factors, particularly variability in relative effectiveness between tumour types. In particular, the ERG suggests the company could have explored variability in the treatment effect across tumour types, and how testing costs are likely to impact on the cost-effectiveness of specific tumour types. The company submission excluded any testing costs associated with NTRK fusion–positive solid tumours.

An overview of the key uncertainties identified by the ERG are presented below. These issues are explored in further detail in section 6.

1. Choice of historical control in the company base case

The available effectiveness evidence for larotrectinib was from three single arm studies (the integrated efficacy analysis) and therefore it was necessary to 'artificially' generate an appropriate comparator to determine the cost-effectiveness of larotrectinib compared to standard practice. The historical comparator utilises previous NICE TAs and other published sources to inform survival estimates for the twelve tumour site specific comparator model engines. The cost-effectiveness estimates for each tumour site were then weighted by the distribution of tumour types in the integrated efficacy analysis. While the ERG considers the company's approach to selecting suitable comparator data was reasonable, the comparability of the larotrectinib and comparator population in terms of characteristics that may impact on prognosis (e.g. ECOG status, age, disease stage, etc.), may bias the estimates of treatment effectiveness. The implications of this bias cannot be assessed and the company does not report patient baseline characteristics for the comparator data sources. Crucially, NTRK status in these twelve populations is unknown. Larotrectinib is only licensed for patients that have proven *NTRK* fusion–positive solid tumours, and therefore, a significant proportion of these comparator patients may not be eligible for treatment with larotrectinib.

The second approach to assess larotrectinib's relative treatment effectiveness, the responder analysis, utilises non-responders from larotrectinib integrated efficacy analysis. As discussed above this ensures that all patients in the non-responder subgroup met the same trial inclusion and exclusion criteria, and are receiving the same line of treatment as larotrectinib patients. However, the small

numbers of patients and events informing the survival models introduces considerable uncertainty in the extrapolation of OS and PFS. It is also unlikely that only the difference in response status explains survival outcomes (see Section 4.6.1) or that non-responder patients do not receive some benefit or harm from having been previously treated with larotrectinib. The fact that they have not responded to larotrectinib means that they are not a larotrectinib eligible population, as per the license requirements.

The ERG consider that, given the issues noted with both the historical comparator and the non-responder comparator, it is important to explore the implications of both approaches on the cost-effectiveness results**.** The ERG could not appropriately implement the previous line of treatment comparator in the model, as it was not considerable feasible given data availability and time constraints.

2. Extrapolation of larotrectinib survival outcomes

The choice of parametric model to extrapolate from observed data for larotrectinib, is a key driver of the survival gains. The ERG highlights that the observed data for larotrectinib was immature, with median OS not yet met. As such, there is significant uncertainty regarding the longer-term survival benefits of larotrectinib. As discussed in Sections 5.2.6.3-5.2.6.4, the company did not comprehensively explore the use of alternative survival models, for larotrectinib, for the later data cut used to inform the economic model. Particularly for OS, the gains for larotrectinib are driven by the long post-progression survival, with considerable separation between extrapolated PFS and OS curves. The ERG considers that, by only exploring alternative parametric distribution for the historical comparator analysis (See Section 6.3.3), the company do not fully explore the implications of the large variation in survival times for larotrectinib and the ICER according to different parametric distributions.

3. Potential confounding of subsequent lines of therapy

The ERG have concerns that post-progression gains in survival for larotrectinib, may be driven, at least in part, by treatments only available to patients in the larotrectinib comparator: continued larotrectinib post progression and LOXO-195. In the base case analysis, no additional cost of larotrectinib is assumed for patients in the post-progression state. Similarly no costs are assumed for LOXO-195. The bias is likely to be in favour of larotrectinib.

Given that the gains in post-progression survival for larotrectinib far exceed the gains in PFS ([redacted] vs [redacted] LYG) and also exceed the OS for the historical comparator ([redacted] vs [redacted] LYG, Table 33), the ERG considers it is important to explore the implications of the confounding in post-progression gains on the cost-effectiveness results.

4. Heterogeneity in treatment effect

The ERG conclude in Section 4 that there is potential for heterogeneity in the treatment effect across tumour types, as well as across other clinical characteristics such as age (paediatric vs adults) and fusion type. The implications of this heterogeneity for cost-effectiveness results are unknown.

As demonstrated in the ERG exploratory analyses on response data (see Section 4.6.1), there is evidence to suggest that the treatment effect is heterogeneous across tumour types. Furthermore, the predicted ORR using a BHM generates different ORRs for the ePAS2 only and the full population (including SAS3 patients): 64% for ePAS2 population and 57% for the full population (section 4.6.2).

The company do not explore the issue of heterogeneity in response and/or survival times, nor do they consider heterogeneity in ORR between the ePAS2 and the full population or by individual study in the integrated efficacy analysis.

The ERG consider the issue of heterogeneity in treatment effect to be a fundamental issue in determining the cost-effectiveness of larotrectinib and has concerns regarding the validity of the ICERs generated by the company, given the lack of analysis to reflect these sources of heterogeneity..

5. Lack of testing costs

The ERG also has substantive concerns regarding the lack of *NTRK* fusion testing costs assumed in the company submission. NTRK gene fusion testing is currently not performed routinely in the UK for all tumour sites, and it is unclear how adding tests costs for these populations will impact on the ICER for larotrectinib. The cost of testing is also an important sources of heterogeneity, as different tumour types require different numbers of patients to test, to identify NTRK fusions, given differences in prevalence rates. The ICER may therefore vary widely by including testing costs weighted according to the tumour types observed in the integrated efficacy analysis.

6 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

6.1 Overview

This section details the ERG's further exploration of the assumptions and uncertainties raised in the review and critique of the company's cost-effectiveness analysis, presented in Section 5. This section is organised in four parts. Section 6.2 details the impact of adjustments to the company base-case analysis in terms of dose and acquisition costs of larotrectinib.

Section 6.3 details a series of scenario analyses exploring the robustness of the cost-effectiveness results to specific assumptions and additional uncertainties identified by the ERG. These analyses were conducted within the company corrected base-case analysis as presented in Table 40 in Section 6.2. The scenario analyses presented in Section 6.3 focus on exploring the following issues and uncertainties associated with the economic analysis:

- Extrapolation of larotrectinib survival data
- Assumptions regarding the gains in post-progression survival and HRQoL with larotrectinib
- Heterogeneity in response rate

In Section 6.4, the ERG alternative base-case is presented based on a combination of the exploratory analyses presented in Section 6.3. Further exploratory analyses in the context of the ERG alternative base-case analysis are also presented in Section 6.5. First, exploring the impact of tumour site specific response rates, utilising the response-based analysis model and assuming a common distribution of PFS and OS across tumour sites, conditional on response status. This analysis is intended to illustrate the potential impact of heterogeneity in response rates, on survival times and the ICER for larotrectinib. A second analysis is conducted using the ERG base-case, again, to enable some exploration of heterogeneity by tumour type, and also in response to the lack of testing costs reflected in the company model. In this scenario the ERG include testing costs by tumour type, reflecting numbers needed to test and different testing strategies according to tumour type (Appendix 10.4).

Due to the model inflexibility, data limitations and time constraints, ICERs based on the deterministic analysis are presented throughout this section. All results are exclusive of the confidential PAS. Results using the confidential PAS are presented ion Appendix 10.5.

6.2 ERG adjustments to the company's base case model

As discussed in Section 5.2.8, subsequent to the CS, the company submitted a revised model. This reflected a new pricing strategy that resulted on a flat price per mg across the three larotrectinib formulations of [redacted]. Section 5.2.9.1 shows the updated base case results reflecting the price change

(exclusive of the confidential PAS). A separate results document was not made available by the company at the time of submitting the revised model. Given the late submission of the revised model (23rd July 2019) the ERG was not able to validate the electronic version. Therefore, the updated pricing was implemented in the version of the model submitted by the company in response to the points for clarification (8th July 2019). All modifications to the company's model subsequently described in Section 6 were implemented on that same version.

In Section 5.2.8.1, the ERG noted concerns around the use of the average paediatric dose of larotrectinib from the larotrectinib integrated efficacy analysis ([redacted] mg/day). The ERG considers that this dose is likely to underestimate the dose of larotrectinib for paediatric patients in clinical practice, as it was derived from data that included Phase I safety assessment doses. To calculate the revised paediatric dose, the ERG applied the ration of [redacted]% ([redacted]/200 mg/day) to the recommended paediatric dose (100 mg/m^2 twice daily). The recommended paediatric dose was calculated assuming the paediatric average BSA in the larotrectinib integrated efficacy analysis ([redacted]*200mg/m^2 day=[redacted] mg/day).

Table 40 ERG adjusted company base-case cost-effectiveness results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
Comparator	[redacted]	[redacted]	[redacted]				
Larotrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

The adjustment to larotrectinib's paediatric dosage increase the ICER to [redacted] per QALY gained compared to the company's base-case with the updated pricing strategy ([redacted] per QALY gained).

6.3 Additional ERG analyses

6.3.1 ERG adaptation of the company cost-effectiveness model

In order to explore a number of uncertainties highlighted in Section 5.3, the ERG further adapted the company's response-based scenario model, which only included alternative survival distributions for the non-responder (proxy for comparator survival in one of the company's scenario). A dual-partitioned response-based model is implemented, which distinguishes between responders and non-responders to larotrectinib. As previously discussed, this approach assumes non-responder patients are a proxy for patients receiving comparator treatments. Importantly, for further exploration of the impact of heterogeneity in response rates on survival outcomes and cost-effectiveness, it assumes a surrogate relationship between response and PFS and OS.

The ERG acknowledges the limitations inherent in this approach, particularly in terms of the maturity and small sample size of the effectiveness data. However, compared to the company's responder analysis, it provides a more flexible framework to explore alternative assumptions on the predicted survival outcomes for larotrectinib, which have been shown to be highly uncertain throughout Section 5.

It also provides a means to more transparently model, compared to the historical comparator, the comparator survival outcomes, in order to assess the magnitude and direction of bias affecting treatment effectiveness estimates in the absence of randomised evidence.

As mentioned above, this approach requires an assumption regarding the relationship between response and PFS and OS. While the FDA evaluation of larotrectinib considered that these surrogate relationships were reasonably likely to predict meaningful benefit[5], this assumption introduces uncertainty in the analysis of unknown magnitude, i.e. we cannot reflect how well response predicts survival. A review of the relationship between the more long-term outcomes of PFS and OS suggested that it varies considerably by cancer type and is not always consistent even within one specific cancer type [85]. Ideally, the use of this approach would need to be accompanied by a review of studies in NTRK fusion patients to consider the extent to which response-based outcomes can be considered a robust surrogate endpoints for PFS and OS, and appropriately quantify these relationships. The ERG considers, however, that in the absence of further data to explore the heterogeneity in survival outcomes, this approach allows exploratory analysis to be conducted that may provide insight into the likely impact of heterogeneity on estimates of cost-effectiveness.

The company had partially implemented the response-based survival models described in Appendix L1.4 of the CS by parameterising six standard parametric distributions (Weibull, exponential, log-normal, log-logistic, Gompertz and generalised gamma) to extrapolate the survival outcomes of non-responders, so as to inform the survival outcomes of the comparator in one of the scenario analysis (see Section 5.2.4).

The company did not report in the CS the exact functional forms used for each parametric distributions or the R software package used to fit the survival models to the observed data. Instead they reported generic functional forms in the model, the R regression output and intermediate calculations necessary to implement the survival curves for both responders and non-responders (see Appendix 10.3). The ERG cannot fully validate these models without knowing the software package used and, therefore, how the regression output is expressed. Thus, the ERG followed the company's approach to parameterise the non-responder extrapolated survival curves to generate survival curves for responders and applying the coefficient for the response status covariate. The company did not

report the variance-covariance matrices for these models, and therefore, the ERG could not fit probabilistic distributions to the data to allow joint parameter uncertainty in survival outcomes to be propagated and therefore reflected in the cost-effectiveness estimates. Hence, all the results presented in this section are deterministic estimates and are likely to underestimate the uncertainty surrounding the cost-effectiveness results.

For the analyses using the response-based model, the comparator survival was assumed to be equivalent to that of the non-responder patients, and the larotrectinib survival was estimated as a weighted average of survival in the responder and non-responder patients, weighted by either the estimated response rate of 72% reported by the company for the ePAS2 population (Section 6.3.3 or the alternative response rates from the BHM described in Section 4.6.1 (64% or 57% for the ePAS2 or full integrated efficacy analysis populations, respectively) (Section 6.3.6). Scenarios were also explored for specific tumour types, representing a high response rate tumour and a low response rate tumour (section 6.5.3).

6.3.2 ERG additional analyses

In Section 5.3, the ERG identified a range of uncertainties relating to the company's base-case cost-effectiveness results. The ERG performed a number of adjustments to the company's revised model (post-clarification questions) to explore the implications of several of these uncertainties.

The following sections describe the model adjustments and report the results of the additional exploratory analyses performed by the ERG to explore the areas of uncertainty identified in Section 5. All scenarios results are reported deterministically. The assumptions varied for each scenario are summarised in Table 41.

Table 41 Overview of ERG's additional analyses

Scenario	Area of uncertainty	Variation from company's base-case assumptions
1-4. Response-based survival approach	Extrapolation of survival outcomes	A response-based dual partition survival approach is used to generate PFS and OS for larotrectinib, and the following parametric distributions are explored: 1. Weibull for OS & PFS 2. Exponential for OS & PFS 3. Gompertz for OS & PFS 4. Gompertz for OS & Weibull for PFS
5-8.Company's preferred survival approach		The company's survival approach is used to generate PFS and OS for larotrectinib, for same set of parametric distributions as in scenario 1-4: 5. Weibull for OS & PFS 6. Exponential for OS & PFS 7. Gompertz for OS & PFS 8. Gompertz for OS & Weibull for PFS
9. Post-progression survival equal for larotrectinib and comparator	Larotrectinib post-progression survival gains	The extent of post-progression survival for larotrectinib is truncated, and costs and QALYs are adjusted to reflect the shorter survival gains.
10. Post-progression survival for larotrectinib equal to OS for comparator		
11. Post-progression utility independent of treatment	Post-progression utility for larotrectinib	Progressed disease health state utility is set equal to [redacted] for both intervention and comparator, and builds on the assumptions for: 11.1 Scenario 4 11.2 Scenario 10
12. Alternative response rates	Heterogeneity in response rates	Builds on the assumptions of scenario 4, and assumes the response rates estimated by the Bayesian hierarchical model on: 12.1 ePAS2: ORR=64% 12.2 Integrated efficacy analysis: ORR=57%
13. Tumour site	Heterogeneity by tumour site	Builds on the assumptions of scenario 4, estimates tumour site specific ICERs for the following, using ORRs from the BHM: 13.1 IFS 13.2 Colorectal cancer

14. NTRK testing costs	Heterogeneity NTRK testing costs	Includes the cost of NTRK fusion testing to the ERG base-case results: 14.1 Assuming a ORR=64% 14.2 Assuming a ORR=57%

6.3.3 Extrapolation of survival outcomes

The ERG have concerns regarding the face validity of the survival projections used in the company cost-effectiveness model. Particularly the ratio between progression -free and post-progression survival. In the company's base case analysis, the gains in post-progression survival for larotrectinib far exceed the gains in PFS ([redacted] vs [redacted] life years). The ERG noted in Section 5.2.6 that the uncertainty in larotrectinib survival outcomes is largely driven by the survival outcomes of responders, in which there are a low number of events (progressed for PFS and deaths for OS). Given the immaturity of the data for PFS and OS, the choice of parametric model to extrapolate from observed data for larotrectinib, is a key driver of the survival gains. The company examined selected the [redacted] distribution to extrapolate PFS and OS in the base-case analysis. The [redacted] distribution was also considered plausible, however the [redacted] distribution was used in the base-case on the basis that it provides closer estimates to the later points of the KM than the exponential. The clinical plausibility of predicted survival estimates was also cited. Alternative parametric distribution were considered in scenarios analyses (see Table 57, CS), for the historical comparator analysis. These show large variation in survival times for larotrectinib (Figure 27, CS) and the ICER according to different distributions. The plausibility of post-progression survival gains for larotrectinib, given pre-progression survival gains, was not discussed in the CS.

The dual-partitioned response-based model implemented by the ERG (see Section 6.2) allows further exploration of alternative assumptions regarding the survival extrapolation of larotrectinib. The following scenarios explore alternative assumptions regarding survival outcomes and discuss the plausibility of post-progression survival gains for larotrectinib given pre-progression survival gains, using the ERG adjusted company's model (see Section 6.2). The survival models explored include the two distributions considered clinically plausible by the company ([redacted]), and the [redacted] distribution which was identified by the ERG as producing potentially plausible survival estimates when considering the company's response-based survival models described in Section 5.2.6. In scenarios 1 to 4 the larotrectinib survival curves assume a 72% response rate to weight the responder and non-responder. The ERG sourced the response rate from the ePAS2 population as reported in the CS. The response rate for the full integrated efficacy analysis that informs the company's approach is lower (66-67% vs 72%). The ERG chose to use the ePAS2 response rate due to the issues with assessing progression in the SAS3 population (Section 4.2) and because the number of responders in the SAS3 populations is reported inconsistently by the company (n=[redacted] in Table 12, CS, n=[redacted] in Table 114, Appendix L).

Cost-effectiveness results for scenarios 1 to 4 are presented in Table 42.

Table 42 Cost effectiveness results for alternative extrapolation assumptions using response-based survival approach

Scenario	1 - Weibull for OS and PFS			2 - Exponential for OS and PFS			3 - Gompertz for OS and PFS			4 - Gompertz for OS and Weibull for PFS		
	Laro	**Comp**	**Inc**	**Laro**	**Comp**	**Inc**	**Laro**	**Comp**	**Inc**	**Laro**	**Comp**	**Inc**
LYG												
Progression-free	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total LYG	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
QALYs												
Progression-free	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AEs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Costs												
Progression-free	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
End of life care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AEs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ICER (per QALY)			[redacted]			[redacted]			[redacted]			[redacted]
Comp, comparator (non-responder control); Inc, incremental; Laro, Larotrectinib (72% response weighted survival)												

The discounted mean LYG for the comparator appear to be fairly stable across the scenarios, varying between [redacted] and [redacted] and between [redacted] and [redacted] for the progression-free and progressed free health states, respectively. This is in contrast with the survival outcomes for larotrectinib, with discounted mean LYG for larotrectinib ranging between [redacted] and [redacted] depending on the OS parametric assumption compared with a range of [redacted] to [redacted] for the non-responder control.

Alternative choices of parametric models for larotrectinib also result into considerably different ratios between survival gains in PFS and OS. In scenarios 1 and 2, the company's preferred survival functions (PFS and OS both extrapolated with the i) [redacted] and ii) [redacted] distributions), suggest a post-progression survival benefit of [redacted] to [redacted] times greater than progression free benefit. When using the [redacted] to extrapolate OS this ratio decreases to [redacted] and [redacted] depending on whether PFS is extrapolated with the [redacted] or the [redacted], respectively.

The ERG considers that the survival extrapolation with the exponential function (Scenario 2) results in clinically implausible survival predictions for larotrectinib ([redacted] LYG). Scenario 1, which uses the company's preferred parametric survival assumptions ([redacted]) predicts survival outcomes similar to the company's base-case analysis (see Table 33), which the ERG had previously considered clinically implausible. Thus, the ERG has concerns that both Scenario 1 and 2 overestimate the survival benefits accrued by patients who receive larotrectinib. Importantly, both scenarios result in considerably large and unrealistic post progression survival periods, and, thus, the resulting ICERs ([redacted] and [redacted] per additional QALY for scenarios 1 and 2, respectively) represent extremely optimistic estimates of cost-effectiveness for larotrectinib.

Scenarios 3 and 4 predict more conservative survival estimates compared to scenarios 1 and 2. The key difference between scenario 3 and 4 appears to be how the survival benefit for larotrectinib is distributed between the progression-free and progressed disease health states. When PFS is extrapolated with the Gompertz distribution (scenario 3) the majority of the survival benefit is accrued in pre-progression (mean [redacted] LYG in pre-progression compared to [redacted] LYG in progressed disease). When PFS is extrapolated with the Weibull distribution instead (scenario 4), the survival benefit is more evenly distributed between the two health states (mean [redacted] and [redacted] LYG in pre-progression and progressed disease, respectively). However, the ERG notes that the shape coefficient for the PFS and OS Gompertz have different signs, suggesting decreasing hazard for the PFS and increasing hazard for the OS. These results in inconsistent PFS and OS curves, and thus, scenario 3 is not considered plausible. Scenario 4 yield an ICER of [redacted] per QALY gained.

The ERG considers that the shorter post-progression survival gains for larotrectinib in scenario 4 makes it more clinically plausible than scenarios 1 and 2. Importantly, these shorter post-progression

survival gains may address some of the concerns expressed by the ERG in terms of the potential contribution of post-progression treatments to these survival gains (see Section 5). While the use of the Gompertz distribution to extrapolate OS in scenario 4 may be argued to generate more conservative survival predictions, it is important to consider that, given the immaturity of the OS data and the small numbers of patients, this may result in optimistic survival projections across all scenarios. In light of this, scenario 4 may reflect a more conservative OS extrapolation assumption across the scenarios, but all four scenarios may overestimate overall survival.

On balance scenario 4 may provide more clinically plausible projections of post-progression survival for larotrectinib. Therefore, the ERG builds upon this scenario in subsequent analyses (Sections 6.3.4 to 6.5).

It is worth noting that while these analyses allow further exploration of the magnitude of post-progression survival and the ratio between progression-free and progressed life years, even the ERG preferred assumptions are unlikely to produce robust results. Therefore, results should be interpreted cautiously, given the significant uncertainty in both the PFS and OS curves for responders.

The ERG also examined the impact of the alternative parametric survival assumptions tested for larotrectinib in the response-based model (scenarios 1 to 4) using the company's preferred approach to model survival. In scenario 5 to 8, survival outcomes are informed by the full integrated efficacy analysis for larotrectinib using the historical comparator.

Table 43 Cost effectiveness results for alternative extrapolation assumptions using the company's survival approach

Scenario	5 - Weibull for OS and PFS			6 - Exponential for OS and PFS			7 - Gompertz for OS and PFS			8 - Gompertz for OS and Weibull for PFS		
	Laro	**Comp**	**Inc**	**Laro**	**Comp**	**Inc**	**Laro**	**Comp**	**Inc**	**Laro**	**Comp**	**Inc**
LYG												
Progression-free	[redacted]											
Progressed disease	[redacted]											
Total LYG	[redacted]											
QALYs												
Progression-free	[redacted]											
Progressed disease	[redacted]											
AEs	[redacted]											
Total QALYs	[redacted]											
Costs												
Progression-free	[redacted]											
Progressed disease	[redacted]											
End of life care	[redacted]											
AEs	[redacted]											
Treatment	[redacted]											
Total costs	[redacted]											
ICER (per QALY)			[redacted]			[redacted]			[redacted]			[redacted]

Comp, comparator (historical comparator); Inc, incremental; Laro, Larotrectinib (integrated efficacy analysis)

All of the scenarios using the company's preferred survival approach (5 to 8) generate long and clinically implausible projections of post-progression survival for larotrectinib. The OS extrapolation with the Gompertz predicts particularly large survival predictions (mean total [redacted] LYG), as the shape parameter is negative (see company's model) suggesting a monotonically decreasing death hazard over time. This is in contrast with the OS extrapolation with the Gompertz using the response-based model for which the shape parameter suggested a monotonically increasing death hazard, and therefore, does not predict a heavy survival tail.

In Section 5.2.6.2, the ERG outlined the limitations of establishing treatment effectiveness comparisons based on i) historical data, and ii) within study non-responder data. It was concluded that the use of historical data may introduce bias of unknown magnitude and direction, while non-responder data may underestimate the treatment effect of the intervention if non-responders derive some benefit from exposure to larotrectinib (even if response is not achieved). The survival outcomes with the historical comparator are higher than with the non-responder control across all alternative parametric assumptions in scenario 1 to 4. This implies a more conservative assumption of the effectiveness of the comparator compared to the historical comparator which should result in a bias in favour of larotrectinib.

6.3.4 Larotrectinib post-progression survival gains

These scenarios build on changes implemented in Section 6.3.3, specifically the ERG consider that scenario 4 provides a better framework to explore further uncertainties, as it generates more clinically plausible projections of post-progression survival for larotrectinib.

As discussed in 5.2.4, the gains in post-progression survival for larotrectinib for the company's base-case far exceed the gains in PFS ([redacted] vs [redacted] LYG) and also exceed the OS for the historical comparator ([redacted] vs [redacted] LYG, Table 33). These large post-progression survival gains may be artefacts of the highly uncertain extrapolation for larotrectinib. They may also be driven by the high proportion of patients in the integrated efficacy analysis that go onto receive treatments post progression that are not currently available in NHS routine practice ([redacted]) receive LOXO-195, and [redacted] continued treatment with larotrectinib after disease progression). For the company base case model, utilising a pooled historical control, the OS estimates are confounded, as patients in the studies informing the effectiveness of the historical control, would not have had access to LOXO-195 or larotrectinib. The bias is likely to be in favour of larotrectinib.

At clarification stage, the ERG requested that the company consider the relatively high proportion of larotrectinib patients that go onto receive LOXO-195, or receive treatment post-progression, and the implications for the historical comparator model (the company base-case). The company did not

provide any further analysis to address the issue and did not justify the use of this data to model progression free survival nor discuss its limitations (A5 and B5, company's response to Points for). The company did provide a scenario which included the cost of larotrectinib for patients receiving this post-progression (using the mean treatment duration estimate from ePAS2 population). This increased the company's original base-case ICER from £[redacted] to £[redacted] per additional QALY.

The ERG explores in this section two scenarios to further explore the uncertainty surrounding post-progression survival gains for larotrectinib. While the extrapolation approach in scenario 4 appears to reduce the extent to which larotrectinib post-progression survival exceeds pre-progression survival ([redacted] vs [redacted] LYG in PFS compared to the company's base-case), the extrapolation is itself uncertain due to the immaturity of the observed data and may overestimate the survival gains for larotrectinib. Thus, scenarios 9 and 10 examine two more conservative structural assumptions for larotrectinib post-progression survival. Both scenarios build on scenario 4, using the response-based model to inform the survival estimates of larotrectinib and the comparator.

Scenario 9 assumes that the mean discounted post-progression survival for larotrectinib is the same as for the comparator ([redacted] LYG), and also equalises the mean discounted costs ([redacted]) and QALYs ([redacted]) accrued in progressed disease.

Scenario 10 presents a less conservative assumption than scenario 9, and assumes that the mean discounted post-progression survival for larotrectinib is the same as the mean discounted overall survival for the comparator ([redacted] LYG). The larotrectinib mean discounted costs and QALYs are adjusted to reflect the reduction in progressed disease survival. The adjustment consists of multiplying the larotrectinib mean discounted progressed disease costs and QALYs as predicted by scenario 4 ([redacted] and [redacted] QALYs) by the ratio between larotrectinib post-progression survival in this scenario and larotrectinib post-progression survival in scenario 4 ([redacted] LYG).

The ERG notes that both scenarios apply crude adjustments to post-progression survival as the model is structured as a partitioned survival model, and, therefore it is not possible to separately track where the gains in post progression gains in survival for larotrectinib and the comparator occur. Furthermore, as the current model setup does not simultaneously produce undiscounted estimates of the outcomes, the ERG applied the model adjustments assuming equivalence between discounted and undiscounted outcomes. Since patients transition to progressed disease on average at an earlier point in time than for larotrectinib the present value of costs, LYG and QALYs for the comparator will be higher (less discounted). So the assumed post-progression costs, LYG and QALYs for larotrectinib will be slightly overestimated. The ERG did not, however, implement further corrections, as the scenarios were only generated for illustrative purposes.

The cost-effectiveness results for scenarios 9 and 10 are shown in Table 44 and Table 45.

Table 44 Cost effectiveness results for post-progression survival equal for larotrectinib and comparator

Scenario 9	Larotrectinib	Comparator	Incremental
LYG			
Progression-free	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]
Total LYG	[redacted]	[redacted]	[redacted]
QALYs			
Progression-free	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]
AEs	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]
Costs			
Progression-free	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]
End of life care	[redacted]	[redacted]	[redacted]
AEs	[redacted]	[redacted]	[redacted]
Treatment	[redacted]	[redacted]	[redacted]
Total costs	[redacted]	[redacted]	[redacted]
ICER (per QALY)			[redacted]

Table 45 Cost effectiveness results for post-progression survival for larotrectinib equal to OS for comparator

Scenario 10	Larotrectinib	Incremental	Comparator
LYG			
Progression-free	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]
Total LYG	[redacted]	[redacted]	[redacted]
QALYs			
Progression-free	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]
AEs	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]
Costs			
Progression-free	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]
End of life care	[redacted]	[redacted]	[redacted]
AEs	[redacted]	[redacted]	[redacted]
Treatment	[redacted]	[redacted]	[redacted]
Total costs	[redacted]	[redacted]	[redacted]
ICER (per QALY)			[redacted]

Both scenarios reduce the mean incremental costs and QALYs compared to scenario 4, and increase the ICER for larotrectinib. The impact on incremental QALYs ([redacted] and [redacted] reduction compared scenario 4 for scenarios 9 and 10, respectively) is, however, greater than for incremental costs ([redacted] and [redacted] reduction compared scenario 4 for scenarios 9 and 10, respectively). This is driven by the post-progression utility assumptions for larotrectinib, which are further explored in Section 6.3.5. Scenario 9 yields an ICER of [redacted] per QALY gained while the ICER for scenario 10 is [redacted] per QALY gained.

6.3.5 Post-progression utility for larotrectinib

In Section 5.2.7.2, the ERG outline the uncertainties surrounding the company's assumption differential utility weights for post-progression for larotrectinib ([redacted]) and comparator treatments ([redacted]). This assumption is tested in two scenario analyses. Scenario 11.1 adds the assumption that utility in post-progression is independent of treatment to Scenario 4. Scenario 11.2 adds the same assumption to the scenario which assumes the post-progression survival for larotrectinib is equal to OS for comparator (Scenario 10).

The post-progression utility is assumed to be equal to the comparator pooled progressed disease utility. This results in a decrease in post-progression utility for larotrectinib from [redacted] to [redacted]. Results are show in Table 46.

Table 46 Cost effectiveness results assuming post-progression utility is independent of treatment

	Total costs	Total LYG	Total QALYs	Inc costs	Inc LYG	Inc QALYs	ICER
Scenario 11.1 - same post progression utility							
Comparator	[redacted]	[redacted]	[redacted]				
Larotrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario 11.2 - same post progression utility and same post-progression survival as comparator OS							
Comparator	[redacted]	[redacted]	[redacted]				
Larotrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

The assumption that post-progression utility is treatment independent increases the ICERs for scenarios 11.1 and 11.2 compared to the scenarios 4 and 10respectively, as it decreases the incremental mean QALYs for larotrectinib. The impact is greater for the scenario that assumes greater post-progression survival for larotrectinib, i.e. scenario 11.1, with an ICER increase of [redacted] compared to scenario 4. The ICER increase for scenario 11.2 compared to scenario 10 is [redacted].

The ERG illustrates in this analysis the potential impact of overestimating post-progression utility for larotrectinib, but acknowledges that HRQoL beyond progression remains particularly uncertain given

the limitations of the evidence base (see section 5.2.7.2). Therefore, the assumptions of this analysis are not carried forward in subsequent analyses.

Similarly, the ERG considers the post-progression survival gains for larotrectinib uncertain. Given the strong assumptions imposed in scenarios 9 and 10 (see Section 6.3.3) to illustrate the potential impact of overestimating these survival gains, these are also not carried forward in subsequent analyses.

6.3.6 Heterogeneity in response rates

As discussed in Section 4.6.1. an alternative ORR can be estimated using a hierarchical approach. This approach recognises that there is potential heterogeneity between tumour histologies and generates a pooled posterior probability of response across all tumour types assuming that the probability of response is a random variable. As discussed in Section 4.2 there are differences between the ePAS2 only and the full integrated efficacy analysis population (including SAS3 patients). This generates different ORRs for these two populations using the BHM: 64% for ePAS2 population and 57% for the full integrated efficacy population. Each of these response rates are considered separately in the following scenarios, which build on changes implemented in scenario 4 in order to explore uncertainty in the ORR used in the response-based model. Results are presented in Table 47.

Table 47 Cost effectiveness results assuming alternative response rates from Bayesian hierarchical model

	Scenario 12.1 – ePAS2 ORR (64%)			**Scenario 12.2 –ORR (57%)**		
	Larotrectinib	**Comparator**	**Incremental**	**Larotrectinib**	**Comparator**	**Incremental**
LYG						
Progression-free	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total LYG	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
QALYs						
Progression-free	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AEs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Costs						
Progression-free	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
End of life care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AEs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ICER (per QALY)			[redacted]			[redacted]
Comp, comparator; Inc, incremental; Laro, Larotrectinib						

The lower response rates result in lower survival gains for larotrectinib in both scenarios (and fewer LYG compared to scenario 4 for scenarios 12.1 and 12.2, respectively), and consequently reduce the mean QALY gains. The reduction in response rates also shifts the balance between post-progression and progression-free survival gains, increasing the proportion of time in post-progression compared to pre-progression for scenario 12.1 and 12.2. Therefore, the mean QALY gains in these scenarios are reduced not only because of the OS reduction, but also due to patients spending proportionally more time in post-progression accruing fewer QALYs. Since the cost of treatment is incurred until progression, some of these QALY losses are offset by the reductions in treatment costs, but not sufficiently to improve the cost-effectiveness of larotrectinib. Overall, the reduction in response rates increases the ICERs to (64% ORR) and per QALY gained (57% ORR).

6.4 ERG base case model

The assumptions and survival modelling approach preferred by the ERG in Section 6.3.3, i.e. dual partition response model with Weibull for PFS and Gompertz for OS extrapolations (Scenario 4), were combined with the response rates estimated by the BHM for the ePAS2 population and the full integrated efficacy analysis (Section 6.3.6). Both estimates of ORR were both considered relevant, given that both account for potential heterogeneity between tumour histologies and the issues with assessing progression in the SAS3 population (Section 4.2). The ERG base-case corresponds to Scenarios 12.1 and 12.2 (Section 6.3.6). Table 48 shows summary results for the ERG base-case, assuming these two response rates of 64% and 57%.

Table 48 ERG base-case cost-effectiveness results

	Total costs	Total LYG	Total QALYs	Inc costs	Inc LYG	Inc QALYs	ICER
ERG base-case with 64% ORR							
Comparator							
Larotrectinib							
ERG base-case with 57% ORR							
Comparator							
Larotrectinib							

Depending on the population informing the response rate estimates, the ICER will range between (64% ORR) to per QALY gained (57% ORR) for the ePAS2 and full integrated efficacy, respectively.

6.5 Further exploratory analysis of heterogeneity using the ERG base-case analysis

6.5.1 Why heterogeneity matters

The cost-effectiveness of larotrectinib may depend on the characteristics of patients or groups of patients. This is termed heterogeneity [86, 87]. Heterogeneity matters for two main reasons: first, if benefits differ by patient characteristics then estimates of treatment benefit must match the patient population that is expected to receive the treatment (the target population) in routine clinical practice. For example, if treatment effects differ by tumour type, then in estimating the effect of a treatment the model must take account of the tumour distribution of the target population. The second reason is that there can be health benefits from making tailored decisions for particular groups of patients. Ignoring these differences could mean that a treatment which is not cost effective for the total population (combining all subgroups) may be cost effective in specific subgroups. Making a "one size fits all" recommendation would then result in a potentially cost effective treatment being withheld from a subset of patients for whom the treatment would represent an appropriate use of NHS resources. Conversely, a treatment which appears cost effective for the total population may not be cost effective in particular subgroups. In this case a "one size fits all" approach could result in the treatment being recommended in identifiable subgroups in which the value of providing the new treatment is lower than the opportunity cost. That is, the health gain for these specific subgroups is not sufficient to offset the potential health lost from a reduction in the provision of services elsewhere in the NHS that is necessary to fund the new treatment.

In the case of histology independent treatments such as larotrectinib, heterogeneity is particularly important to consider. This is because, an important (though not the only) source of heterogeneity is the difference in tumour histology. Though larotrectinib may be clinically effective across a range of tumour sites, there are theoretical and empirical reasons to expect that cost and health consequences could vary significantly by tumour type. This is in addition to the usual sources of heterogeneity (e.g. age, gender etc.) which are present in conventional treatments.

There are multiple potential sources of heterogeneity relevant to larotrectinib. First is the cost of patient identification across histologies. The frequency of mutations in different histologies differs markedly (see Table 55). In some "high frequency" histologies the relevant mutation is almost always present and so testing is very likely to result in a positive result and subsequent treatment with the appropriate therapy [34, 88]. In other "low frequency" histologies the opposite is the case. This can lead to wide disparities in patient identification (i.e. screening) costs across histologies. Second is heterogeneity in relative treatment effect. The relative effect of larotrectinib compared to standard care may be greater in some histologies compared to others. For a survival outcome this might imply a hazard ratio which differs according to histology. Third is heterogeneity in baseline risk. This is where different histologies will have different prognosis, regardless of the treatment they are given. For this reason, different histologies may have the same relative effect estimate (i.e. relative effect

homogeneity) but different baseline risks and so will differ in terms of their absolute outcomes with treatment. Fourth is health state and quality of life heterogeneity. This is where the disease costs and quality of life associated with different histologies differ. For example, more severe histologies may be associated with lower quality of life and larger disease costs.

For larotrectinib, it is possible that each of the four sources of heterogeneity could be present when considering different histologies. Even after taking account of heterogeneity in histologies (see Figure 3), additional heterogeneity could also exist in adults compared to children due to differences in baseline risks, disease costs and quality of life.

The CS provides a base case analysis which acknowledges histology heterogeneity in the comparator arm (historical control), however the CS does not fully explore heterogeneity in survival outcomes for larotrectinib and importantly the consequences of any differences for cost-effectiveness. Data from the integrated efficacy analysis are utilised in the base case, assuming that the population across the 3 studies can be pooled (Section 4.2). Differences in response rates observed across tumour types were not reflected in the base case analysis, which uses a historical comparator and therefore does not account for differences in PFS and OS for responders and non-responders. In the company scenario utilising the responder analysis, the non-response stratified full integrated efficacy analysis data was utilised to inform larotrectinib survival implicitly assuming the unadjusted response rate for this population, and thus, not recognising the heterogeneity in the sample. The response-based survival model was only used to inform the comparator by assuming equivalence between non-responders and non-active treatment.

An alternative approach would have been to use a BHM to estimate ORR, allowing for potential heterogeneity between histologies, a random effects model (Section 4.6.1), and apply these estimates in a dual partition response-based survival model that allows establishing a link between a response rate that incorporates heterogeneity and survival as illustrated in Section 6.3.6.

Another approach explored by the ERG to illustrate the potential impact of heterogeneity on survival outcomes is detailed in Section 4.6.2. Assuming that the distribution of survival times is common across tumour sites conditional on response status, it was possible to derive OS and PFS estimates by tumour site.

The ERG analyses were exploratory and require strong assumptions about the link between response and survival outcomes, but do serve to highlight how heterogeneity by tumour histology may impact on treatment effectiveness. Uncertainty at this level is compounded by the immaturity of the survival data in the pooled clinical trial data, especially for the OS data (only 14 deaths and 88 patients censored at the 30th July 2018 cut-off). The assumption of a common distribution of survival times

across tumour sites, conditional on response status, is highly uncertain and therefore the total impact of heterogeneity across tumour types on cost-effectiveness, is largely unknown.

The ERG do not consider that the CS made sufficient efforts to explore heterogeneity.

6.5.2 Data required to illustrate heterogeneity

At clarifications stage the ERG requested data to explore heterogeneity and quantify uncertainty around each subgroup, including tumour type, by age ORR status, response category, fusion type and Isoform (A3). The ERG requested this data in multiple formats: individual patient data (IPD) and aggregate data as KMs by subgroup or median outcomes (time on treatment, PFS and OS at 6 and 12 months). This would have allowed the ERG to explore the impact of heterogeneity in PFS and OS using a hierarchical approach similar to that described in Section 6.5.1. The company did not provide any of this data in their response to clarifications, stating that they did not believe that providing subgroup data is justified or helpful in terms of decision-making (A3, company's response to points for Clarification, and Clarification additional response). They also stated that there was no evidence of heterogeneity in treatment effect according to the subgroups requested and that patient numbers are too small for further analysis. Whilst the ERG, recognises that the patient numbers for many subgroups are small and the data for PFS and OS is immature, given the heterogeneity in response rates observed (see Section 4.6.1), it is important to consider the potential differences in PFS and OS for subgroups, including by tumour type, and the impact that these differences may have on cost-effectiveness. Not exploring this heterogeneity and its potential impact would completely ignore that fact that it exists.

Observed differences in survival outcomes were discussed in Section 5.6.2, however without the outcomes for both PFS and OS, using a consistent data cut, the ERG are unable to formally explore this heterogeneity. Data by study, as opposed to the integrated efficacy analysis was requested following the clarification response (via email to NICE and subsequent communication between NICE and the company). This would enable the ERG to explore potential differences between the adult and paediatric populations. This are potentially important subgroups, given the differences in PFS noted in Section 5.2.6. The company did not provide this evidence.

The current data available to the ERG does not provide sufficient evidence to allow stratified decisions at this stage, however a more explicit assessment of heterogeneity, particularly in the larotrectinib arm is still an important consideration at this stage. There are 2 main reasons for this: (i) it is possible that the distribution of patients treated in real life will be different to that observed in the trial - so we need some way to assess whether this could materially affect the ICER estimates and (ii) we need to better understand the potential importance of heterogeneity to help inform and prioritise

data collection (e.g. CDF data arrangements) that would facilitate a more robust assessment in the future.

6.5.3 Exploring heterogeneity in response rates according to tumour type

In order to explore heterogeneity and its impact on cost-effectiveness, the ERG specify two scenarios using tumour specific response rates as estimated in Section 4.6.1 from the BHM, utilised in the response-based model. This enables scenarios to be specified for extreme response examples. Two tumour types are chosen for scenarios: 13.1 IFS which has a very high response rate (87% estimated from the BHM) compared to the overall, 13.2 and colorectal cancer which has a very low response rate (43% estimated from the BHM) compared to the overall rate. For these two tumour types, tumour specific ICERs are estimated. Tumour specific health state utilities and costs are applied to the comparator.

These scenarios are likely to underestimate the impact of tumour type on the ICERs, as a common distribution of PFS and OS is applied, conditional on response. Without access to the data as requested, the ERG are unable to explore the validity of this assumption. As discussed in 6.3.1, it was also not possible to run these scenarios probabilistically, and therefore a value of heterogeneity framework [86] cannot be used to further explore the potential consequences to decision making of ignoring heterogeneity. Scenario results are presented in Table 49 and Table 50.

Table 49 Cost effectiveness results for IFS scenario (87% ORR estimated from the BHM)

Scenario 13.1	**Larotrectinib**	**Comparator**	**Incremental**
LYG			
Progression-free	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]
Total LYG	[redacted]	[redacted]	[redacted]
QALYs			
Progression-free	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]
AEs	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]
Costs			
Progression-free	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]
End of life care	[redacted]	[redacted]	[redacted]
AEs	[redacted]	[redacted]	[redacted]
Treatment	[redacted]	[redacted]	[redacted]
Total costs	[redacted]	[redacted]	[redacted]
ICER (per QALY)			[redacted]

Table 50 Cost effectiveness results for colorectal cancer scenario (43% ORR estimated from the BHM)

Scenario 13.2	Larotrectinib	Comparator	Incremental
LYG			
Progression-free	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]
Total LYG	[redacted]	[redacted]	[redacted]
QALYs			
Progression-free	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]
AEs	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]
Costs			
Progression-free	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]
End of life care	[redacted]	[redacted]	[redacted]
AEs	[redacted]	[redacted]	[redacted]
Treatment	[redacted]	[redacted]	[redacted]
Total costs	[redacted]	[redacted]	[redacted]
ICER (per QALY)			[redacted]

The IFS ICER is lower than ERG base-case estimates (£[redacted] vs £[redacted] and £[redacted] per QALY gained for ORR=64% and 57%, respectively) due to the increase in mean incremental QALY gains in both progression-free and post-progression and despite the increase in treatment costs.

For colorectal cancer, the ICER is higher than ERG base-case estimates (£[redacted] vs £[redacted] and £[redacted] per QALY gained for ORR=64% and 57%, respectively). In this scenario the survival gains for larotrectinib are reduce with fewer QALYs increased on average. Despite the reduction in costs of treatment, this scenario is less favourable for larotrectinib.

The ERG notes that these results are markedly uncertain given the small number of patients per tumour site. Furthermore, it is not the case that the ICERs will be consistently lower for tumour sites with high response rates or consistently higher for tumour sites with low response rates. This is due in part to the change in balance between post-progression and progression-free survival gains at different response rates, and the trade-off between increasing pre-progression QALY gains and treatment costs for larotrectinib as time in progression-free increases. The magnitude of ICER change is also difficult to predict due to the tumour site specific comparator health state costs and utilities, which vary widely across tumour sites.

6.5.4 Exploring heterogeneity in testing costs according to tumour type

In Section 5.2.8.5, the ERG highlights the importance of examining the potential impact of costs of NTRK testing in the cost-effectiveness of larotrectinib. Since the company did not address this area of uncertainty, the ERG explores a scenario including testing costs and examines its likely impact on cost-effectiveness.

NTRK gene fusion testing is currently not performed routinely in the UK for all tumour sites, and it is unclear how the current diagnostic pathway, for each tumour site that might harbour these mutations, may change to accommodate NTRK testing (see Section 2.2.2.1). Furthermore, genomic test provision is heterogeneous across NHS trusts, and even in tumours sites for which genomic testing is recommended, further local level criteria may restrict the set of patients to whom these tests are offered (Dr Helene Schlecht, personal communication July 2019). The NTRK testing costs implemented in this scenario represent hypothetical tumour site specific diagnostic pathways based on existing literature and clinical advice, and aim only to illustrate the potential impact of these costs the cost-effectiveness of larotrectinib. The ERG did not attempt to optimise diagnostic strategies in terms of cost-effectiveness. Appendix 10.4 details NTRK fusion testing costs calculations.

The weighted overall cost of testing for larotrectinib applied in the model is £18,670. In this scenario, the cost of testing was added as a one-off cost to the total costs of larotrectinib. Results are presented in Table 51.

Table 51 Cost-effectiveness results including NTRK fusion testing costs

	Total costs	Total LYG	Total QALYs	Inc costs	Inc LYG	Inc QALYs	ICER
Scenario 14.1: ORR=64%							
Comparator	[redacted]	[redacted]	[redacted]				
Larotrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario 14.2: ORR=57%							
Comparator	[redacted]	[redacted]	[redacted]				
Larotrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Considering the cost of NTRK gene fusion testing increases the ERG base case ICER to from £[redacted] per QALY gained when assuming a 64% ORR, and from £[redacted] to [redacted] per QALY gained, when assuming a 57% ORR.

6.6 Conclusions from ERG analyses

The ERG has presented a number of additional analyses to explore key uncertainties identified in the company base case cost-effectiveness results. These analyses were undertaken using the model submitted in the original CS, but updated, by the ERG, for the price changes for larotrectinib which were later submitted by the company. In addition the ERG revised the paediatric dose of larotrectinib in accordance with average BSA in the larotrectinib clinical trial programme.

The scenario analyses specified using this revised model addressed the following issues:

1. Parametric distribution fitted to larotrectinib PFS and OS data
2. Assumptions regarding the gains in post-progression survival with larotrectinib
3. Model used to estimate overall response rate

The scenarios associated with the greatest impact on cost-effectiveness outcomes related to alternative extrapolation assumptions applied to OS and PFS, using the ERG amended dual-partitioned response-based model. The ERG regards scenarios utilising the exponential and Weibull (the company's base case) parametric distributions for OS and PFS, generate clinically implausible survival predictions for larotrectinib ([redacted] and [redacted]LYG for the exponential and Weibull respectively). The ERG consider that scenario 4, utilising the Gompertz for OS and Weibull for PFS, generates the most clinically plausible projections survival for larotrectinib, including more conservative estimates of the ratio between pre- and post-progression survival benefits. This scenario generates an ICER for larotrectinib of [redacted] per additional QALY compared to the ERG adjusted company base-case of [redacted] per additional QALY. Alternative scenarios utilising the historical comparator were also explored by the ERG. All of these generate long and clinically implausible projections of post-progression survival for larotrectinib.

Scenarios relating to the post progression survival gains for larotrectinib increase the ICER to [redacted] per QALY gained and [redacted] per QALY gained, however these scenarios remain highly uncertain and therefore, the assumptions of this analysis are not carried forward in subsequent analyses. Utilising alternative ORRs from a BHM, only has a small impact on the ICER, in addition to the use of alternative survival models implemented in the response-based model (scenario 4).

The ERG base-case represents a combination of these scenarios and concludes that a dual partition response model with Weibull for PFS and Gompertz for OS extrapolations (Scenario 4) combined with response rates estimated by the BHM for the ePAS2 population and the full integrated efficacy analysis, represent the most appropriate scenarios. The ERG base case cost-effectiveness results generate an ICER for larotrectinib of [redacted] and [redacted] per QALY gained for the ePAS2 population and the full integrated efficacy analysis respectively.

The final part of this section carried out a further series of exploratory analyses that explored the impact of heterogeneity in response rates on survival estimates and therefore cost-effectiveness. The ERG specify two scenarios using tumour specific response rates as estimates in Section 6.5 and utilised in the responder analysis model. The ICER for the exemplar tumour site with a high response rate (IFS, ORR=85%) was lower than for the two ERG base-case ICERs (£[redacted] compared to £[redacted] and £[redacted] per QALY gained for the ePAS2 population (ORR=64%) and the full integrated efficacy analysis (ORR=57%), respectively). The ICER for the exemplar tumour site with a low response rate (colorectal cancer, ORR=43%) was lower than for the two ERG base-case ICERs (£[redacted] compared to £[redacted] and £[redacted] per QALY gained for the ePAS2 population (ORR=64%) and the full integrated efficacy analysis (ORR=57%), respectively). The ERG considers these results particularly uncertain, and notes that it is not possible to generalise the relationship between response rate and the ICERs, i.e. it is not the case that the ICER will be consistently lower for tumour sites with high response rates or consistently higher for tumour sites with low response rates.

Heterogeneity in testing costs is also explored by specifying a scenario including testing costs, weighted according to the prevalence of tumour types observed in the integrated efficacy analysis. Considering the cost of NTRK gene fusion testing increases the ERG base case ICER to from £[redacted] per QALY gained (ePAS2 population), and from £[redacted] to [redacted] per QALY gained (integrated efficacy analysis population)

There remain a number of uncertainties in determining the cost-effectiveness of larotrectinib, which were not possible for the ERG to explore, given the data provided by the company and the time constraints for the appraisal. In particular, the immaturity of the data and the small number of events for responders, means that even the ERG results exploring alternative survival extrapolations are unlikely to be robust. They do, however result in survival estimates they are more likely to be regarded as clinically plausible compared to the company base case. The ERG were unable to properly account for post progression treatments received by larotrectinib patients and therefore the scenarios specified by the ERG are likely to be subject to bias.

The ERG were extremely limited in analysis of the impact of heterogeneity on cost-effectiveness. The company did not provide the ERG with data or scenarios to specify survival outcomes by sub-group, including tumour histology, and therefore the ERG analysis of heterogeneity was limited to differences in response rate according to tumour type, assuming that the distribution of survival outcomes is independent of response. The ERG believe that there is evidence of heterogeneity in survival outcomes, for example observed differences in survival outcomes discussed in Section 5.2.6. that suggests potential differences in outcomes between the adult and paediatric populations. Without

the outcomes for both PFS and OS, using a consistent data cut, the ERG were unable to formally explore this heterogeneity.

The scenario incorporating NTRK testing costs is likely to be conservative as it does not reflect any implementation costs when expanding the service across the NHS. The roll out of WGS across the NHS is recommended in the NHS Long term Plan, however the timescale for this remains uncertain. The costs associated with wide scale roll out also remain uncertain. The implications of diagnostic accuracy to cost-effectiveness are also unknown. The ERG specify a scenario for NTRK testing cost based on clinical advice. The ERG did not attempt to optimise diagnostic strategies in terms of cost-effectiveness and therefore the current pathway may not represent the most cost-effective way of testing potential NTRK fusion patients. Pathways are also likely to differ by prevalence, where more conservative testing strategies are employed in tumour types where NTRK fusion prevalence is low.

7 End of life

The ERG notes that the lack of direct comparator data on survival without larotrectinib, and that tumour sites were not considered separately in the CS, makes end of life difficult to assess reliably.

Based on the pooled data across all tumour sites the median overall survival in patients receiving best supportive care (or a proxy to it) is around 400 days. Median PFS is around 100 days. Median OS on patients who did not respond to larotrectinib was around 12 months. Median PFS on larotrectinib was 27.4 months in the ePAS2 data; median overall survival has not yet been reached. Therefore, when data are pooled across tumour sites, the end-of-life criteria appear to be met. This does not apply to patients with primary CNS tumours (the SAS3 data set), where median PFS was [redacted] months, so it is unclear if there is a meaningful treatment benefit.

The ERG considered whether end-of-life criteria were met for each included tumour site. To do this the median PFS and OS based on assuming exponential survival distributions for the comparator data (data provided in the supplied economic model) were compared to the ERG's analysis of PFS and OS by tumour site (see Figure 7 to Figure 10). The results of this analysis are shown in Table 52.

This analysis suggests that expected survival is below 24 months for [redacted] survival is close to 24 months [redacted]. The only tumour sites where there is good evidence that PFS is also improved by at least three months (i.e. the 5th centile of the bootstrapped distribution of improvement in PFS exceeds 3 months) are: [redacted]. Results for improvement in OS are not used, given the lack of data on overall survival.

The ERG notes that this analysis is speculative, but considers that it gives an indication of where end-of-life criteria might be met. The ERG considers that the criteria are plausibly met [redacted] subject to current limitations in the data.

Table 52 Assessing end-of-life criteria by tumour site

Tumour site	Median overall survival (from comparator data)	Median increase in PFS (from ERG analysis)	Minimum (5th centile) increase in PFS (from ERG analysis)
Bone sarcoma	[redacted]	[redacted]	[redacted]
Breast	[redacted]	[redacted]	[redacted]
Cholangiocarcinoma	[redacted]	[redacted]	[redacted]
Colon	[redacted]	[redacted]	[redacted]
CMN	[redacted]	[redacted]	[redacted]
GIST	[redacted]	[redacted]	[redacted]
Infantile fibrosarcoma	[redacted]	[redacted]	[redacted]
Lung	[redacted]	[redacted]	[redacted]
Melanoma	[redacted]	[redacted]	[redacted]
Pancreas	[redacted]	[redacted]	[redacted]
Salivary gland	[redacted]	[redacted]	[redacted]
Soft tissue sarcoma	[redacted]	[redacted]	[redacted]
Thyroid	[redacted]	[redacted]	[redacted]

8 Overall conclusions

8.1 Clinical effectiveness

8.1.1 Trials of larotrectinib

The CS presented data from three single-arm trials of larotrectinib which recruited patients across a number of tumour sites and included both adult and paediatric patients. The primary analysis included 93 patients, with a further 9 patients with primary CNS tumours also considered. The ERG notes that this is a small sample on which to base any assessment, and samples are even smaller once the number within any tumour site included is considered; with a maximum of [redacted] patients in any tumour site group, or 17 when considering adults and children separately.

The analysis in the CS made the assumption that larotrectinib was equally effective in all tumour types, and all patients were analysed together without differentiating between trials, patient ages or tumour sites. The ERG considers this to be inappropriate, as heterogeneity is at least plausible, and basket trials are designed to investigate heterogeneity across tumour sites. The ERG requested data by tumour site to investigate possible heterogeneity, but the company declined to provide it, on the grounds that the data were too limited for such an analysis. The ERG disagrees with this, noting that investigating heterogeneity is particularly important where data are limited.

The ERG performed a Bayesian analysis of overall response rate, accounting for possible heterogeneity across tumour sites. This concluded that the best estimate of overall response rate across all tumour sites was 64% (95% CrI 29 to 83), somewhat lower than the estimate presented in the CS. The analysis identified evidence of heterogeneity in response across tumour sites, with a possibility that larotrectinib is more effective in tumours where NTRK fusion is common (e.g. MASC, IFS) and less effective where NTRK fusion is rare (e.g. pancreas, appendix, breast). The ERG notes three possible reasons for the observed heterogeneity:

1. Genuine clinical variation in effect in different tumour sites, perhaps as a consequence of NTRK fusion prevalence.

2. High rates of "false positives" (people who do not have NTRK fusion tumours, who cannot benefit from larotrectinib) in tumour sites where NTRK fusion is rare.

3. Chance finding, due to low numbers in tumour sites where NTRK fusion is rare.

It is unclear how this heterogeneity might impact on survival outcomes. The company declined to provide survival data categorised by tumour site. Speculative analyses performed by the ERG suggest that this heterogeneity in response rate may lead to heterogeneity in progression-free and overall

survival rates across tumour sites. However the limitations of the data mean these findings are uncertain.

8.1.2 Indirect comparisons

As all three trials had single arm designs, no direct comparison of larotrectinib with other interventions was possible. The CS presented three methods for comparing larotrectinib to other interventions (or best supportive care). These were used only to inform the economic analysis, and no comparative estimates of effect (e.g. hazard ratios) were calculated.

Survival curves were estimated for each tumour site using data from past NICE TAs or, in the absence of suitable data, other literature sources or by assuming equivalent survival across similar tumour sites. These TAs varied in what intervention was used; some used best supportive care, other an active intervention; there were some tumour sites with no TA, so assumptions were made by equating survival across similar tumour sites.. Because of these assumptions, the somewhat arbitrary selection of previous TAs, the variation in interventions considered, and the likelihood that people in previous TAs did not have NTRK fusions, the ERG considers this approach to have substantial limitations, and may not truly represent the survival expectations of people with NTRK fusion who do not receive larotrectinib.

Patients who did not respond to larotrectinib were considered as a proxy for patients not receiving larotrectinib. The ERG notes that this may be a more suitable analysis, as it uses actual patient data from the trials, and patients were eligible for larotrectinib. However, non-responders may have a worse survival profile than patients who never receive larotrectinib; the ERG's analysis supported this possibility for progression-free survival. For overall survival results may be biased because many non-responder continued to receive larotrectinib or were recruited into the trial of LOXO-195 (an experimental treatment for patients with larotrectinib resistance), and so may have better than expected survival. This possibility was confirmed by the ERG's analysis.

The third approach compared time to progression on previous lines of therapy to progression-free survival on larotrectinib. The ERG considers this to be of limited vale as it cannot inform overall survival, data were not available for all patients, and there was considerable heterogeneity across patients and tumour sites.

8.1.3 Identifying NTRK fusions

The ERG notes that there is considerable uncertainty in how many people will be eligible for larotrectinib. Estimates of the prevalence of NTRK fusion are highly uncertain, with varying estimates from different sources, heterogeneity across tumour sites, and uncertainty over exactly which types of

tumour might harbour NTRK fusions. There is also uncertainty as to where in the clinical pathway larotrectinib will be used. The approval is for whenever no satisfactory treatment options exists; hence use of larotrectinib may vary between tumour sites, where the range of available treatments will vary. The estimates from the BIA and ERG analysis suggest only a small number of patients per year (less than 10) will be eligible for larotrectinib, mostly with rare tumours where NTRK fusion is common.

The CS did not consider screening to identify NTRK fusions. The ERG considers this to be a critical omission, because successful screening for NTRK fusion is essential to identify patients eligible for larotrectinib. There currently appears to be little consensus on how genetic screening should be used. In some tumour sites it may already be widely used (but may require extra panels to identify NTRK fusion specifically); in others it may not be used at all. The ERG found that the numbers needed to screen to identify one NTRK fusion may be very high in tumour sites with low NTRK fusion prevalence, raising doubt as to the practicality of screening in such tumour sites.

There also appears to be uncertainty as to the accuracy of screening to identify NTRK fusions. The ERG notes that, for tumour sites with low NTRK fusion prevalence, even with a near perfect test (e.g. of 99% accuracy) the number of "false positives" (people who test positive for NTRK fusion but do not have it) may outnumber the true NTRK fusion cases. This will substantially reduce the observed effectiveness of larotrectinib, and casts doubt on whether larotrectinib can be used ethically in tumour sites with low NTRK fusion prevalence.

8.2 Implications for research

The ERG considers that the primary need for further research is to complete the NAVIGATE trial, focussing on recruiting patients from currently under-represented tumour sites. This would help resolve uncertainty over possible heterogeneity in response and survival rates across tumour sites.

Ideally further, independent basket trials of larotrectinib should be performed. These should include a suitable control arm. While a control arm of patients with NTRK fusion might not be feasible, a "basket" of patients without an NTRK fusion could be included, receiving the best appropriate alternative to larotrectinib, preferably with matching by tumour site, ECOG status etc. If NTRK fusion (after matching) does not affect survival prognosis this approach fits the requirements of "Mendelian randomisation"; that is, NTRK fusion status is essentially random, so can be used to "randomise" patients to trial arms. Hence such trials could be analysed, approximately, as if they were conventional RCTs.

This STA has highlighted major limitations in the broader evidence around NTRK fusion. In particular, around the prognosis of patients with NTRK fusion, its prevalence across tumour sites, and the implementation and diagnostic accuracy of screening to identify NTRK fusion. Observational studies, diagnostic accuracy studies, and audits of registry data are all needed to inform how larotrectinib might be used in practice.

8.3 Cost-effectivness

The ERGs adjusted company base-case ICER, for larotrectinib compared with established management, generated a single ICER of [redacted] per QALY (exclusive of the confidential PAS) to cover the anticipated marketing authorisation. The ERG has substantive concerns about the validity of the survival gains the company proposes for larotrectinib and also about the presentation of a single ICER across all tumour types. The ERG's view is that this potentially conceals significant variation in the tumour specific ICERs, driven by a combination of factors, particularly variability in relative effectiveness between tumour types. The company submission also excluded any testing costs associated with *NTRK* fusion–positive solid tumours.

The uncontrolled comparison used in the larotrectinib clinical studies and the immature and small evidence base, presents a significant challenge in establishing robust estimates of cost-effectiveness. The ERG do not believe that the company submission explores the uncertainty in survival outcomes sufficiently or accounts for potential heterogeneity in treatment effect and its implications for cost-effectiveness.

The ERG proposed an alternative base-case to address several of the key uncertainties identified. This explored the extrapolation of survival outcomes for larotrectinib, post progression survival gains and estimation of overall response rate using a Bayesian Hierarchical model. The ERG concluded that utilising the Gompertz distribution for OS and Weibull for PFS, within a dual portioned response-based model, generates the most clinically plausible projections survival for larotrectinib, including more conservative estimates of the ratio between pre and post progression survival benefits. The ERG base case cost-effectiveness results generate an ICER for larotrectinib of [redacted] and [redacted] per QALY (exclusive of the confidential PAS) gained for the ePAS2 population and the full integrated efficacy analysis respectively.

The company base case analysis utilises instead a historical comparator. The ERG consider that, while the response-based model is also subject to limitations and assumptions, compared to the historical comparator, however it includes a population more likely to match the characteristics of patients in the larotrectinib comparator. It is also flexible to allow for further exploration of the impact of heterogeneity in response rates on survival outcomes and cost-effectiveness,

The ERG was extremely limited in terms of the analysis of heterogeneity possible with the data provided by the company. Analysis of specific tumour types, representing low and high prevalence tumours, demonstrates that heterogeneity in the ICER is likely. This may be further compounded by any differences in the distribution of survival outcomes by tumour type, which the ERG believe are likely. In addition, the ERG consider that NTRK testing currently represents a cost to the NHS, which is also driven by the tumour types larotrectinib will be made available for, with lower prevalence tumour types requiring larger numbers needed to test to identify cases. The addition of testing costs further increases the ICER for larotrectinib to [redacted] and [redacted] per QALY (exclusive of the confidential PAS) gained for the ePAS2 population and the full integrated efficacy analysis respectively.

9 References

1. Amatu A, Sartore-Bianchi A, Siena S. NTRK gene fusions as novel targets of cancer therapy across multiple tumour types. *ESMO Open* 2016;**1**:e000023.
2. Bongarzone I, Vigneri P, Mariani L, Collini P, Pilotti S, Pierotti MA. RET/NTRK1 rearrangements in thyroid gland tumors of the papillary carcinoma family: correlation with clinicopathological features. *Clin Cancer Res* 1998;**4**:223-8.
3. Prasad ML, Vyas M, Horne MJ, Virk RK, Morotti R, Liu Z, et al. NTRK fusion oncogenes in pediatric papillary thyroid carcinoma in northeast United States. *Cancer* 2016;**122**:1097-107.
4. Vokuhl C, Nourkami-Tutdibi N, Furtwangler R, Gessler M, Graf N, Leuschner I. ETV6-NTRK3 in congenital mesoblastic nephroma: a report of the SIOP/GPOH nephroblastoma study. *Pediatr Blood Cancer* 2018;**65**.
5. Center for Drug Evaluation and Research. *NDA multidisciplinary review and evaluation NDA 210861 and NDA 211710 VITRAKVI (larotrectinib)*. Silver Spring, MD: U.S. Food and Drug Administration; 2016.
6. Sethi R, Kozin E, Remenschneider A, Meier J, VanderLaan P, Faquin W, et al. Mammary analogue secretory carcinoma: update on a new diagnosis of salivary gland malignancy. *Laryngoscope* 2014;**124**:188-95.
7. Farrimond J. *Analysis of the cancer registry combined database for use with the Brain and CNS Registry*: Eastern Cancer Registration and Information Centre; 2010.
8. Luk PP, Selinger C, Eviston T, Lum T, O'Toole SA, Clark JR, et al. Mammary analogue secretory carcinoma: an evaluation of its genetic and clinical characteristics. *Pathology* 2015;**47**:S69.
9. Orbach D, Rey A, Cecchetto G, Oberlin O, Casanova M, Thebaud E, et al. Infantile fibrosarcoma: management based on the European experience. *J Clin Oncol* 2010;**28**:318-23.
10. NHS. *The NHS Long Term Plan*; 2019.
11. NHS England. *National Genomic Test Directory for cancer*. NHS England; 2019. Available from: https://www.england.nhs.uk/publication/national-genomic-test-directories/ [accessed 24th July 2019].
12. Lewisham and Greenwich NHS Trust. *PD-L1 Freedom of Information Request*. London: Lewisham and Greenwich NHS Trust; 2017.
13. National Institute for Health and Care Excellence (NICE). *Gene expression profiling and expanded immunohistochemistry tests to guide selection of chemotherapy regimes in breast cancer management*: NICE; 2011.
14. Cross PC, Chandra A, Giles T, Johnson S, Kocjan G, Poller D, et al. *Guidance on the reporting of thyroid cytology specimens*. London: The Royal College of Pathologists; 2016.
15. Greater Manchester Lancashire and South Cumbria Strategic Clinical Network. *Referral and clinical guidelines for lung cancer*: NHS; 2014.
16. London Cancer Alliance (LCA). *LCA colorectal cancer clinical guidelines*. London: LCA; 2014.
17. Reid R, Bulusu R, Carroll N, Eatock M, Geh I, Judson I, et al. *Guidelines for the management of gastrointestinal stromal tumours (GIST)*. London: Association of Upper Gastrointestinal Surgeons of Great Britain and Ireland (AUGIS); 2009.
18. London Cancer. *Management of patients with suspected cholangiocarcinoma clinical guidelines*. London: London Cancer; 2014.
19. Ganesan R, Singh N, McCluggage WG. *Dataset for histological reporting of endometrial cancer*. London: The Royal College of Pathologists; 2017.
20. Fotopoulou C, Hall M, Cruickshank D, Gabra H, Ganesan R, Hughes C, et al. British Gynaecological Cancer Society (BGCS) epithelial ovarian/fallopian tube/primary peritoneal cancer guidelines: recommendations for practice. *Eur J Obstet Gynecol Reprod Biol* 2017;**213**:123-39.
21. Hirschowitz L, Albus A, Brown L, Das N, Ganesan R, Herrington CS, et al. *Histopathology reporting in cervical screening – an integrated approach*. Sheffield: NHS Cancer Screening Programmes; 2012.
22. Health and Social Care Information Centre. *National Head and Neck Cancer Audit 2014: report 7*. Leeds: Health and Social Care Information Centre; 2015.
23. London Cancer Alliance (LCA). *LCA oesophageal and gastric cancer clinical guidelines*. London: LCA; 2014.

24. Oxley J, Varma M, Berney D. *Dataset for histopathology reports for prostatic carcinoma* London: The Royal College of Pathologists; 2016.
25. London Cancer Alliance (LCA). *LCA brain/CNS cancer clinical guidelines*. London: LCA; 2014.
26. Okamura R, Boichard A, Kato S, Sicklick JK, Bazhenova L, Kurzrock RJpoN-. Analysis of NTRK alterations in pan-cancer adult and pediatric malignancies: implications for NTRK-targeted therapeutics (published online ahead of print November 15 2018). *JCO Precis Oncol* 2018: 10.1200/PO.18.00183.
27. Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med* 2017;**23**:703-13.
28. Sigal DS, Bhangoo MS, Hermel JA, Pavlick DC, Frampton G, Miller VA, et al. Comprehensive genomic profiling identifies novel NTRK fusions in neuroendocrine tumors. *Oncotarget* 2018;**9**:35809-12.
29. Schwarze K, Buchanan J, Fermont JM, Dreau H, Tilley MW, Taylor JM, et al. The complete costs of genome sequencing: a microcosting study in cancer and rare diseases from a single center in the United Kingdom (published online ahead of print July 30 2019). *Genet Med* 2019: 10.1038/s41436-019-0618-7.
30. Penault-Llorca F, Rudzinski ER, Sepulveda AR. Testing algorithm for identification of patients with TRK fusion cancer. *J Clin Pathol* 2019:doi:10.1136/jclinpath-2018-205679.
31. Marchio C, Scaltriti M, Ladanyi M, Iafrate AJ, Bibeau F, Dietel M, et al. ESMO recommendations on the standard methods to detect NTRK fusions in daily practice and clinical research. *Ann Oncol* 2019: https://doi.org/10.1093/annonc/mdz204.
32. Solomon JP, Hechtman JF. Detection of NTRK fusions: Merits and limitations of current diagnostic platforms. *Cancer Res* 2019;**79**:3163-8.
33. Kummar S, Lassen UN. TRK inhibition: a new tumor-agnostic treatment strategy. *Target Oncol* 2018;**13**:545-56.
34. Hsiao SJ, Zehir A, Sireci AN, Aisner DL. Detection of tumor NTRK gene fusions to identify patients who may benefit from TRK inhibitor therapy. *J Mol Diagn* 2019;**21**:553-71.
35. Drilon A, Laetsch TW, Kummar S, DuBois SG, Lassen UN, Demetri GD, et al. Efficacy of Larotrectinib in TRK Fusion-Positive Cancers in Adults and Children. *N Engl J Med* 2018;**378**:731-9.
36. Cunanan KM, Iasonos A, Shen R, Begg CB, Gonen M. An efficient basket trial design. *Stat Med* 2017;**36**:1568-79.
37. Hyman DM, Puzanov I, Subbiah V, Faris JE, Chau I, Blay J-Y, et al. Vemurafenib in multiple nonmelanoma cancers with BRAF V600 mutations. *N Engl J Med* 2015;**373**:726-36.
38. Groenendijk FH, Bernards R. Drug resistance to targeted therapies: deja vu all over again. *Mol Oncol* 2014;**8**:1067-83.
39. Hyman D, Kummar S, Farago A, Geoerger B, Mau-Sorensen M, Taylor M, et al. Abstract CT127: Phase I and expanded access experience of LOXO-195 (BAY 2731954), a selective next-generation TRK inhibitor (TRKi). *Cancer Res* 2019;**79**:CT127-CT.
40. Thall PF, Wathen JK, Bekele BN, Champlin RE, Baker LH, Benjamin RS. Hierarchical Bayesian approaches to phase II trials in diseases with multiple subtypes. *Stat Med* 2003;**22**:763-80.
41. Cunanan KM, Iasonos A, Shen R, Gonen M. Variance prior specification for a basket trial design using Bayesian hierarchical modeling. *Clin Trials* 2019;**16**:142-53.
42. Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol* 2012;**12**:9.
43. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996;**313**:275-83.
44. European Medicines Agency, Committee for Medicinal Products for Human Use. *Vitrakvi - Summary of opinion (initial authorisation)*. 25th July 2019. Available from: https://www.ema.europa.eu/en/documents/smop-initial/chmp-summary-positive-opinion-vitrakvi_en.pdf [accessed 1st August 2019].

45. National Institute for Health and Care Excellence (NICE). *Erlotinib and gefitinib for treating non-small-cell lung cancer that has progressed after prior chemotherapy [TA374]*. NICE; 2015. Available from: https://www.nice.org.uk/guidance/ta374 [accessed 23rd August 2019]
46. Valle J, Wasan H, Palmer DH, Cunningham D, Anthoney A, Maraveyas A, et al. Cisplatin plus gemcitabine versus gemcitabine for biliary tract cancer. *N Engl J Med* 2010;**362**:1273-81.
47. Airoldi M, Pedani F, Succo G, Gabriele AM, Ragona R, Marchionatti S, et al. Phase II randomized trial comparing vinorelbine versus vinorelbine plus cisplatin in patients with recurrent salivary gland malignancies. *Cancer* 2001;**91**:541-7.
48. National Institute for Health and Care Excellence (NICE). *Pembrolizumab for treating advanced melanoma after disease progression with ipilimumab*. NICE; 2015. Available from: https://www.nice.org.uk/guidance/ta357 [accessed 23rd August 2019]
49. Ribas A, Puzanov I, Dummer R, Schadendorf D, Hamid O, Robert C, et al. Pembrolizumab versus investigator-choice chemotherapy for ipilimumab-refractory melanoma (KEYNOTE-002): a randomised, controlled, phase 2 trial. *Lancet Oncol* 2015;**16**:908-18.
50. National Institute for Health and Care Excellence (NICE). *Trifluridine–tipiracil for previously treated metastatic colorectal cancer [TA405]*. NICE; 2016. Available from: https://www.nice.org.uk/guidance/ta405 [accessed 23rd August 2019]
51. Mayer RJ, Van Cutsem E, Falcone A, Yoshino T, Garcia-Carbonero R, Mizunuma N, et al. Randomized trial of TAS-102 for refractory metastatic colorectal cancer. *N Engl J Med* 2015;**372**:1909-19.
52. National Institute for Health and Care Excellence (NICE). *Regorafenib for previously treated unresectable or metastatic gastrointestinal stromal tumours [TA488]*. NICE; 2017. Available from: https://www.nice.org.uk/guidance/ta488 [accessed 23rd August 2019]
53. Demetri GD, Reichardt P, Kang Y-K, Blay J-Y, Rutkowski P, Gelderblom H, et al. Efficacy and safety of regorafenib for advanced gastrointestinal stromal tumours after failure of imatinib and sunitinib (GRID): an international, multicentre, randomised, placebo-controlled, phase 3 trial. *Lancet* 2013;**381**:295-302.
54. National Institute for Health and Care Excellence (NICE). *Trabectedin for the treatment of advanced soft tissue sarcoma [TA185]*. NICE; 2010. Available from: https://www.nice.org.uk/guidance/ta185 [accessed 23rd August 2019]
55. Van Glabbeke M, Verweij J, Judson I, Nielsen OS. Progression-free rate as the principal end-point for phase II trials in soft-tissue sarcomas. *Eur J Cancer* 2002;**38**:543-9.
56. Mascarenhas L, Lyden ER, Breitfeld PP, Walterhouse DO, Donaldson SS, Paidas CN, et al. Randomized phase II window trial of two schedules of irinotecan with vincristine in patients with first relapse or progression of rhabdomyosarcoma: a report from the Children's Oncology Group. *J Clin Oncol* 2010;**28**:4658-63.
57. National Institute for Health and Care Excellence (NICE). *Eribulin for treating locally advanced or metastatic breast cancer after 2 or more chemotherapy regimens [TA423]*. 2016. Available from: https://www.nice.org.uk/guidance/ta423 [accessed 23rd August 2019]
58. Cortes J, O'Shaughnessy J, Loesch D, Blum JL, Vahdat LT, Petrakova K, et al. Eribulin monotherapy versus treatment of physician's choice in patients with metastatic breast cancer (EMBRACE): a phase 3 open-label randomised study. *Lancet* 2011;**377**:914-23.
59. National Institute for Health and Care Excellence (NICE). *Pegylated liposomal irinotecan for treating pancreatic cancer after gemcitabine [TA440]*. NICE; 2017. Available from: https://www.nice.org.uk/guidance/TA440 [accessed 23rd August 2019]
60. Batchelor TT, Mulholland P, Neyns B, Nabors LB, Campone M, Wick A, et al. Phase III randomized trial comparing the efficacy of cediranib as monotherapy, and in combination with lomustine, versus lomustine alone in patients with recurrent glioblastoma. *J Clin Oncol* 2013;**31**:3212-8.
61. National Institute for Health and Care Excellence (NICE). *Lenvatinib and sorafenib for treating differentiated thyroid cancer after radioactive iodine [TA535]*. NICE; 2018. Available from: https://www.nice.org.uk/guidance/ta535 [accessed 23rd August 2019]

62. Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc JF, et al. Sorafenib in advanced hepatocellular carcinoma. *N Engl J Med* 2008;**359**:378-90.
63. Hatswell AJ, Sullivan WG. Creating historical controls using data from a previous line of treatment - Two non-standard approaches. *Stat Methods Med Res* 2019: https://doi.org/10.1177/0962280219826609.
64. Shepherd FA, Rodrigues Pereira JR, Ciuleanu T, Tan EH, Hirsh V, Thongprasert S, et al. Erlotinib in previously treated non–small-cell lung cancer. *N Engl J Med* 2005;**353**:123-32.
65. Office for National Statistics (ONS). *National life tables, UK: 2014 to 2016*. ONS; 2017. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/nationallifetablesunitedkingdom/2014to2016 [accessed 23rd August 2019]
66. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health* 2012;**15**:708-15.
67. Khan KA, Petrou S, Rivero-Arias O, Walters SJ, Boyle SE. Mapping EQ-5D utility scores from the PedsQL generic core scales. *Pharmacoeconomics* 2014;**32**:693-706.
68. Liberato NL, Rognoni C, Rubrichi S, Quaglini S, Marchetti M, Gorlia T, et al. Adding docetaxel to cisplatin and fluorouracil in patients with unresectable head and neck cancer: a cost–utility analysis. *Ann Oncol* 2011;**23**:1825-32.
69. Delea TE, Amdahl J, Nakhaipour HR, Manson SC, Wang A, Fedor N, et al. Cost-effectiveness of pazopanib in advanced soft-tissue sarcoma in Canada. *Curr Oncol* 2014;**21**:e748-59.
70. Amdahl J, Manson SC, Isbell R, Chit A, Diaz J, Lewis L, et al. Cost-effectiveness of pazopanib in advanced soft tissue sarcoma in the United Kingdom. *Sarcoma* vol 2014; article ID 481071, 14 pages.
71. National Institute for Health and Care Excellence (NICE). *Guidance on the use of temozolomide for the treatment of recurrent malignant glioma (brain cancer) [TA23]*. NICE; 2001. Available from: https://www.nice.org.uk/guidance/ta23 [accessed 23rd August 2019]
72. Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. *Health Qual Life Outcomes* 2008;**6**:84.
73. *Drugs and pharmaceutical electronic market information tool (eMIT)* Department of Health and Social Care; 2018. Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit [accessed 23rd August 2019]
74. The Royal Pharmaceutical Society. *British National Formulary*. 2018. Available from: https://about.medicinescomplete.com/ [accessed 23rd August 2019]
75. *Reference costs 2017-2018*. NHS Improvement; 2018. Available from: https://improvement.nhs.uk/resources/reference-costs/ [accessed 23rd August 2019]
76. *MIMS. Monthly Index of Medical Specialities. Database of prescription and generic drugs, clinical guidelines 2018*. Haymarket Media Group Ltd.; 2018. Available from: https://www.mims.co.uk/ [accessed 23rd August 2019]
77. Curtis L, Burns A. *Unit costs of health and social care 2018*. Canterbury: Personal Social Services Research Unit, University of Kent; 2018.
78. National Institute for Health and Care Excellence (NICE). *Blood transfusion [NG21]*. NICE; 2018. Available from: https://www.nice.org.uk/guidance/ng24 [accessed 23rd August 2019]
79. Round J, Jones L, Morris S. Estimating the cost of caring for people with cancer at the end of life: a modelling study. *Palliat Med* 2015;**29**:899-907.
80. National Institute for Health and Care Excellence (NICE). *Ipilimumab for previously treated advanced (unresectable or metastatic) melanoma [TA268]*. NICE; 2012. Available from: https://www.nice.org.uk/guidance/ta268 [accessed 23rd August 2019]
81. National Institute for Health and Care Excellence (NICE). *Nivolumab for treating squamous cell carcinoma of the head and neck after platinum-based chemotherapy [TA490]*. NICE; 2017. Available from: https://www.nice.org.uk/guidance/ta490 [accessed 23rd August 2019]
82. van der Graaf WT, Blay JY, Chawla SP, Kim DW, Bui-Nguyen B, Casali PG, et al. Pazopanib for metastatic soft-tissue sarcoma (PALETTE): a randomised, double-blind, placebo-controlled phase 3 trial. *Lancet* 2012;**379**:1879-86.

83. Centre for Bayesian Statistics in Health Economics (CHEBS). *Excel macros for sampling from multivariate Normal and Dirichlet distributions*. CHEBS, University of Sheffield; Available from: https://www.sheffield.ac.uk/chebs/software [accessed 23rd August 2019]
84. Briggs A, Sculpher M, Claxton K. *Decision modelling for health economic evaluation*. Oxford: Oxford Univeristy Press; 2006.
85. Latimer NR. Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data: inconsistencies, limitations, and a practical guide. *Med Decis Making* 2013;**33**:743-54.
86. Espinoza MA, Manca A, Claxton K, Sculpher MJ. The value of heterogeneity for cost-effectiveness subgroup analysis: conceptual framework and application. *Med Decis Making* 2014;**34**:951-64.
87. Coyle D, Buxton MJ, O'Brien BJ. Stratified cost-effectiveness analysis: a framework for establishing efficient limited use criteria. *Health Econ* 2003;**12**:421-7.
88. Albert CM, Davis JL, Federman N, Casanova M, Laetsch TW. TRK fusion cancers in children: a clinical review and recommendations for screening. *J Clin Oncol* 2018;**37**:513-24.
89. Office for National Statistics (ONS). *Cancer registration statistics, England: 2017*. ONS; 2019. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancerregistrationstatisticsengland/2017 [accessed 20th August 2019].
90. Brzezianska E, Karbownik M, Migdalska-Sek M, Pastuszak-Lewandoska D, Wloch J, Lewinski A. Molecular analysis of the RET and NTRK1 gene rearrangements in papillary thyroid carcinoma in the Polish population. *Mutat Res* 2006;**599**:26-35.
91. Feng J, Ebata K, Hansen F, Kivi L, Kriegshauser C, Morosini D, et al. 80P TRK wild-type and fusion protein expression in solid tumors: characterization by immunohistochemistry and in situ hybridization. *Ann Oncol* 2018;**29**: https://doi.org/10.1093/annonc/mdy314.021.
92. Morgan P, Woolacott N, Biswas M, Mebrahtu T, Harden M, Hodgson R. Crizotinib for untreated anaplastic lymphoma kinase-positive non-small-cell lung cancer: an Evidence Review Group perspective of a NICE Single Technology Appraisal. *Pharmacoeconomics* 2017;**35**:909-19.

10 Appendices

10.1 Philips checklist

Description of quality	Response (✓,✗ or NA)	Comments	Reference
Structure			
S1 Statement of decision problem objective			
Is there a clear statement of the decision problem?	✓		CS, Table 1, p14
Is the objective of the evaluation and model specified and consistent with the stated decision problem?	✓	The evaluation and model are consistent with the decision problem as outlined by the NICE Scope. However, the company does not give a clear definition of the lack of "satisfactory treatment options", which is one of the criteria which was used to identify relevant tumour types, over and above the presence of NTRK fusion. Additional areas of uncertainties are given by the identification of the number of patients at each line of therapy, the exact position of larotrectinib in the treatment pathway, and the testing methods used to identify eligible patients.	CS, Table 1, p14
Is the primary decision-maker specified?	Partly	Not specified, but implied as the decision problem is defined in terms of the NICE scope.	
S2 Statement of scope/perspective			
Is the perspective of the model clearly stated?	✓	NHS and Personal Social Services (NHS & PSS) perspective.	CS, Table 1, p16
Are the model inputs consistent with the stated perspective?	✓	Yes.	
Has the scope of the model been stated or justified?	✓	Yes. The scope of the model reflects the one set out by NICE and the expected marketing authorisation.	CS, Table 1, p17
Are the outcomes of the model consistent with the perspective, scope and overall objective of the model?	✓	The outcomes included in the model are life-years, quality adjusted life years based on EQ-5D and costs, and are consistent with both perspective, scope, and overall objective of the model, and the NICE Reference Case	

S3 Rationale for structure			
Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	✓	The company developed a cohort state transition model with a PartSA approach. The PartSA model contained three mutually exclusive health states: progression free (PF), progressed disease (PD) and death. The model is stated to be in line with the standard approaches followed in oncology modelling and to be consistent with previous economic evaluations submitted to NICE.	CS, Section B.3.2, p139-144
Are the sources of data used to develop the structure of the model specified?	✓	A review of the literature informed potential or previously used approaches in modelling histology independent treatments. A second review regarded the interpretation and analysis of basket trials. Given the lack of previous modelling experience for histology independent treatments, the model methodology was aligned to the NICE Reference Case using previously accepted modelling approached.	CS, p142 CS, Appendix M
Are the causal relationships described by the model structure justified appropriately?	Partly	The causal relationship was justified, but the lack of RCT data, the high number of tumour sites, and the low frequency of some tumour sites renders the causal relationship between larotrectinib and outcomes highly uncertain and likely to be biased.	
S4 Structural assumptions			
Are the structural assumptions transparent and justified?	✓	The company presents a table listing a set of relevant base-case model assumption, the mitigation strategy that was implemented, and its justification.	CS, Table 51, p194-195
Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Partly		
S5 Strategies/comparators			
Is there a clear definition of the options under evaluation?	✓	Yes. The CS details the two options under evaluation, that is Larotrectinib and a mixed basket of comparator treatments. The Company justifies the choice of comparators and that of using a mixed basked of pooled last-line comparator treatments.	CS, Section B.3.2
Have all feasible and practical options been evaluated?	✓		

Is there justification for the exclusion of feasible options?	✓		
S6 Model type			
Is the chosen model type appropriate given the decision problem and specified causal relationship within the model?	✓	Yes. The model is justified based on previous NICE Technology Appraisals and follows the recommendations of the NICE Reference Case	CS, Section B.3.2, p. 139
S7 Time horizon			
Is the time horizon of the model sufficient to reflect all important differences between options?	✓	The time horizon used in the model was 40 years for adult patients, and 80 years for paediatric and pooled (adult and paediatric) populations. This is assumed to represent a lifetime horizon able to reflect all important differences between Larotrectinib and the pooled comparator.	CS, Section B.3.2, p147
Are the time horizon of the model, the duration of treatment and the duration of treatment effect described and justified?		Time horizon: The time horizon is in line with NICE guidance, but not justified. Duration of treatment: The treatment stopping rules are not specified in the expected marketing authorisation. The model assumes treatment is discontinued at progression	CS, Section B.3.2, p. 147
S8 Disease states/pathways			
Do the disease states or the pathways reflect the underlying biological process of the disease in question and the impact of interventions?	✓	The model reflects the NICE Reference Case and previously accepted approaches to modelling oncological treatments.	CS, Section B.3.2, p. 140
S9 Cycle Length			
Is the cycle length defined and justified in terms of the natural history of disease?	Partly	The cycle length is defined but not justified in terms of the natural history of disease.	CS, Section B.3.2 p147
Data			
D1 Data identification			
Are the data identification methods transparent and appropriate given the objectives of the model?	✓		
Where choices have been made between data sources, are these justified appropriately?	✓	Appendix M in the CS details the various sources and the relevant choices made by the company when identifying data sources for the various tumour sites.	CS, Appendix M

Has particular attention been paid to identifying data for the important parameters in the model?	✓	Yes	
Has the quality of the data been assessed appropriately?	✓	The quality of effectiveness cost-effectiveness studies identified in the literature reviews was conducted with a relevant checklist.	CS, Appendices 2 (Response to Points for Clarification)
Where expert opinion has been used, are the methods described and justified?	NA	No formal elicitation methods applied in the submission, only validation of model parameters and outputs.	
D2a Baseline data			
Is the choice of baseline data described and justified?	Partly	The company undertook a comprehensive systematic review to inform the clinical and economic parameter for the comparators. However, the submission did not present enough information to allow a comparison between the comparator baseline characteristics or outcome data, by tumour site, with the corresponding data for larotrectinib.	CS, Appendix M
Has a half-cycle correction been applied to both cost and outcome?	✓	Yes	CS, Section B.3.2
D2b Treatment effects			
If the relative treatment effects have been derived from trial data, have they been synthesised using appropriate techniques?	NA	Relative treatment effectiveness was derived from non-randomised data. The main effectiveness inputs, PFS and OS, were derived from the pooled population (n=102) belonging to the three studies included in the Larotrectinib Clinical Trial Programme. Pooling of data across studies could introduce a source of bias, given the differences between the studies and, within the studies, across the various tumour sites. Comparator data were modelled independently for each tumour sites and was informed by historical data. The use of historical control is discussed by the ERG, as it can constitute an additional source of bias in the presence of hardly comparable populations. However, the direction and magnitude of the bias is unknown. Alternative approaches to modelling comparator data, such as using data from non-responders or considering data from the previous line of treatment, were also considered by the company and assessed by the ERG.	CS, Section B.3.3
Have the methods and assumptions used to extrapolate short-term results to final outcomes been documented and justified?	✓	The choice of parametric curve was informed through visual inspection, assessment of clinical plausibility, and metrics of statistical fit in line with NICE Decision Support Unit guidelines.	CS, Section B.3.3
Have assumptions regarding the continuing effect of treatment once treatment is complete been documented and justified?	Partly	The ERG notes how Overall Survival gains are likely to be driven by post-progression survival. In turn, this is likely to be biased in favour of Larotrectinib, as patients on Larotrectinib had access to a drug which is not currently available, LOXO-195, and to which patients in the comparator group had no access. The company did not clarify the role of LOXO-195 and did not provide a scenario adjusting treatment effect to reflect current NHS practice.	

Have alternative extrapolation assumptions been explored through sensitivity analysis?	✓	Yes, the company explored the use of alternative distribution when modelling survival data.	CS, Section B.3.8
Have alternative assumptions regarding the continuing effect of treatment been explored through sensitivity analysis.	✓	Yes. The company explored the impact of an alternative model specification using survival curves based on treatment to discontinuation as opposed to treatment until progression.	C.S., Section B.3.8
D2c Costs			
Are the costs incorporated into the model justified?	✓	Yes.	
Has the source of the costs been described?	✓	Unit costs were based on previous NICE TAs and available literature, the company's proposed list price for larotrectinib (as well as proposed PAS price), NHS Reference costs, Personal Social Services Research Unit (PSSRU) and the Department of Health's electronic market information tool (eMIT), and the British National Formulary (BNF). Where appropriate, unit costs were inflated to 2017/2018 prices. All sources were explicitly stated and described.	CS, Section B.3.5
Have the discount rates been described and justified given the target decision maker?	✓	Conventional 3.5% annual discount rates were presented for the base-case scenario. The discount rate was varied in a scenario to 1.5% for costs and effectiveness. The selection of discount rate was justified based on the NICE Reference Case	CS, Table 50, p192; Table 52, p. 196
D2d Quality of life weights			
Are the utilities incorporated into the model appropriate?	✓	The utilities incorporated into the model are in line with the NICE reference case.	CS, Section B.3.4
Is the source of the utility weights referenced?	✓	Yes. The company describes the relevant clinical trial sources and describes the sources for the mapping algorithm used to derive appropriate utility weights.	CS, Section B.3.4
Are the methods of derivation for the utility weights justified	✓	The company describes the process of data collection and derivation of EQ-5D-3L data using published algorithm sources.	CS, Section B.3.4
D3 Data incorporation			
Have all data incorporated into the model been described and referenced in sufficient detail?	✓	The data incorporated in the model and their sources are generally described with a sufficient level of detail.	
Has the use of mutually inconsistent data been justified (i.e. are assumptions and choices appropriate?)	NA		
Is the process of data incorporation transparent?	✓	Yes.	

If data have been incorporated as distributions, has the choice of distributions for each parameter been described and justified?	Partly	The choice of the probability distribution is not always justified.	CS, Appendix M
If data have been incorporated as distributions, is it clear that second order uncertainty is reflected?	Partly	The CS does not always justify in detail the source to inform the standard error of the parameters included in the model.	CS, Appendix M
D4 Assessment of uncertainty			
Have the four principle types of uncertainty been addressed? If not, has the omission of particular forms of uncertainty been justified?	✓		
D4a Methodological			
Have methodological uncertainties been addressed by running alternative versions of the model with different methodological assumptions?	✓	The company considers alternative modelling methods for the counterfactual, namely non-responder control analysis and a naïve comparison with the previous line of treatment.	CS, Section B.3.8, p. 213-217
D4b Structural			
Is there evidence that structural uncertainties have been addressed via sensitivity analysis?	×	Key structural uncertainties in terms of cure timing and OS extrapolation were not sufficiently explored. A set of issues, such as the adjustment of larotrectinib treatment effect to reflect current practice or the inclusion of the cost of testing, were left unexplored despite ERG requests.	CS, Section B.3.8
D4c Heterogeneity			
Has heterogeneity been dealt with by running the model separately for different subgroups?	×	The model was not run separately by subgroups. The CS states that this was due to the histology independent nature of the intervention, and because it was not possible to identify subgroups.	CS, Section B.3.9, p. 220
D4d Parameter			
Are the methods of assessment of parameter uncertainty appropriate?	✓	In line with the NICE reference case deterministic sensitivity analyses were performed on a series of model parameters. Probabilistic sensitivity analyses were also performed.	CS, Section B.3.8
If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?	Partly	The company presents deterministic and probabilistic sensitivity and scenario analysis. However, the one-way deterministic sensitivity analysis, consisting in varying model inputs over +/- 20% of the point estimate, were not justified. The probability distribution used in the PSA is not justified for each type of parameter. Also, the selection of parameters which are varied in the PSA was not justified.	CS, Section B.3.6.3, Table 52; Section B.3.8
Consistency			
C1 Internal consistency			

Is there any evidence that the mathematical logic of the model has been tested thoroughly before use?	✓	The company states (Section B.3.2, Page 143) that the model was designed in accordance with the requirements of NICE guidance and ISPOR-SMDM Guidelines. The CS states an initial model validation exercise was undertaken by health economist who had not developed the model and by an independent health economic and outcomes research consultancy.	CS, Section B.3.10
C2 External consistency			
Are any counterintuitive results from the model explained and justified?	NA		
If the model has been calibrated against independent data, have any differences been explained and justified?	NA		
Have the results of the model been compared with those of previous models and any differences in results explained?	NA	The company did not retrieve any other economic evaluation or cost-effectiveness studies investigating treatments in a TRK-fusion population. The company did not identify any other economic model considering multiple tumour sites from a single-arm trial.	CS, section B.3.2, page 137

10.2 Company base-case analysis with PAS price

The results in this section reflect the outcome of analyses when the patient access schemes (PAS) discount for larotrectinib is applied. The PAS price consists of [redacted] (as submitted on the 5th August 2019) for the three larotrectinib formulations resulting in a cost per mg reduction from [redacted] to [redacted].

The company's base-case results are summarised in Table 53 and Table 54 for the deterministic and probabilistic analyses, respectively.

Table 53 Company base-case deterministic cost-effectiveness results with PAS price

	Larotrectinib	Comparators	Incremental
Total costs	[redacted]	[redacted]	[redacted]
Total life years	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]
ICER			[redacted]
Abbreviations: ICER, incremental-cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years			

The use of a PAS price reduces the total mean costs for larotrectinib in the company's base-case with the proposed NHS list (as submitted on the 5th August 2019) from [redacted] to [redacted], resulting in an ICER reduction from [redacted] to [redacted] per additional QALY).

Table 54 Company base-case probabilistic cost-effectiveness results with PAS price

	Larotrectinib	Comparators	Incremental
Total costs	[redacted]	[redacted]	[redacted]
Total life years	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]
ICER			[redacted]
Abbreviations: ICER, incremental-cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years			

The probabilistic and deterministic results are comparable. The company also submitted the PAS price results for their scenario analyses (see ID1299 larotrectinib PAS template 05082019KM). The ERG did not undertake formal validation of these results as they were received less than three weeks before the ERG report was due.

10.3 Response-based survival models

Figure 20 Generic functional forms for the response-based survival models (from the model)

Functions in R
p: probability of survival
t: time
A: intermediate calcuation group
i: response status, i= 1 for responders and i=0 for non-responders

Exponential: p = exp(-t/A)
A=exp[-ln(Rate) - coefficient*i]

Gompertz: p=exp{(A/shape)*[1-exp(shape*t)}}
A=exp(ln(Rate)+coefficient*i)

Log logistic: p=1/{exp[shape*ln(t/A)]+1}
A=exp[ln(Rate)+coefficient*i]

Log normal: p = {1-ERF[ln(t)/((2^0.5)*sd)-A]}/2
A= (mean+coefficient*i)/((2^0.5)*sd)

Weibull: p = exp[-exp(shape*ln(t/A))]
A=exp[ln(scale)+coefficient*i]

Gen gamma: p = 1- pgengamma(t, mu + coefficient*i, sigma, Q)
Use base case function

Figure 21 R regression output for PFS response-based survival models (from the model)

Figure 22 R regression output for OS response-based survival models (from the model)

10.4 NTRK fusion testing costs

Testing Strategy

The testing strategy per tumour site was informed by discussions with Helene Schlect (personal communication July 2019), clinical advisor to the ERG and one published testing algorithm [31] described in Section 2.2.2.2. NTRK fusion testing strategies will vary by tumour site depending on the NTRK fusion rate and whether the current diagnostic pathway for that tumour site already involves NGS testing. When a tumour histology has a high NTRK fusion rate, patients should be tested with a RNA-based NGS panel for the three NTRK fusion types. For tumour histologies with lower fusion rates, a cheaper screening test, IHC, can be used with a confirmatory RNA-based NGS offered only to those who screen positive. However, if the current diagnostic pathway for a particular tumour site with low NTRK fusion rates already includes a RNA-NGS panel, the existing panel can be expanded to also detect the NTRK fusions (at a negligible additional costs) and screening with IHC is not necessary. The ERG assumed as a starting point that that NTRK fusion testing will be performed with:

- RNA-based NGS for tumour sites with high NTRK fusion rates
- IHC followed by RNA-based NGS (henceforth, referred to as NGS) testing for tumour sites with low NTRK fusion rates

As stated in Section 2.2.2.1, WGS is currently recommended in the National Genomic Test Directory for sarcomas, melanomas, and paediatric cancers. However, the accuracy of WGS to detect NTRK fusion types is not perfect (ref), and confirmatory NGS of patients who screen positive would still be required before patients are offered larotrectinib. Therefore, ERG further assumed that for these tumour sites and if NTRK fusion rates are low, screening with WGS (instead of IHC) is conducted in all patients followed by NGS testing for those who screen positive with WGS. WGS is assumed to impose no incremental cost as it would be offered to the full tumour site population (i.e. patients in the comparator treatment would also have received WGS). Screening with WGS is assumed not to be necessary for IFS patients, given the high NTRK fusion rate for this tumour site. All IFS patients are, therefore, assumed to be tested with NGS.

The clinical advisor to the ERG also stated that some laboratories in the UK already test NSCLC with a RNA-based NGS panel for lung cancer that can be expanded to include all NTRK fusion types at a negligible incremental cost. The ERG assumed that NSCLC patients would be tested with an expansion of the currently used RNA-based NGS panel for lung cancer.

Finally MASC patients are already routinely tested for NTRK fusion in the NHS [11], so that no testing costs are attributed to these tumour sites.

Table 55 summarises the testing strategies assumed in this scenario for each tumour site.

Table 55 NTRK testing scenario – testing strategies per tumour site

Tumour site	NTRK Fusion		Testing Strategy
	Rate	Source	
NSCLC	0.09%	Larotrectinib FDA submission [5]	NGS+
Salivary (non-MASC)	1.72%	Larotrectinib FDA submission [5]	IHC + NGS
MASC	100%	Larotrectinib FDA submission [5]	NTRK fusion testing
Melanoma	0.21%	Okamura et al. (2018) [26]	WGS + NGS
Colorectal	0.12%	Larotrectinib FDA submission [5]	IHC + NGS
Appendix	4.00%	Amatu et al. (2016) [1] (assumed same as pancreatic)	IHC + NGS
GIST	1.28%	Larotrectinib FDA submission [5]	WGS + NGS
STS adults (non-GIST)	0.56%	Larotrectinib FDA submission [5]	WGS + NGS
Bone Sarcoma	1.00%	ERG Assumption, based on STS rate from Figure 1 (CS)	WGS + NGS
STS paediatric	0.56%	Larotrectinib FDA submission [5]	WGS + NGS
IFS	90.90%	Larotrectinib FDA submission [5]	NGS
Breast (non-secretory)	0.07%	Larotrectinib FDA submission [5]	IHC + NGS
Cholangiocarcinoma	0.10%	Larotrectinib FDA submission [5]	IHC + NGS
CNS	0.05%	Larotrectinib FDA submission [5]	IHC + NGS
CNS paediatric	5.30%	Okamura et al. (2018) [26]	WGS + NGS
Pancreas	0.26%	Okamura et al. (2018) [26]	IHC + NGS
Thyroid	4.94%	Assumption, weighted average of fusion rates for papillary and generic thyroid tumours in larotrectinib FDA submission [18](weighted by UK cancer incidence statistics for all thyroid tumors [89] and papillary thyroid tumours [90])	IHC + NGS

Testing costs

The costs per testing strategy includes the costs of screening (for strategies including this component) and the cost of testing with NGS. The cost of screening is calculated by multiplying the number of patients needed to screen (inverse of the fusion rate for each tumour site) by the unit cost of the screening test (WGS or NGS). The cost of NTRK fusion testing is incurred by all patients for tumour sites where there is no screening stage, and by 9% of patients who undergo screening[91]. This assumes that 9% of patients in the low NTRK fusion rate tumour site specific populations will screen positive with IHC. This IHC positive rate estimate was sourced from a study assessing the diagnostic accuracy of an IHC staining protocol across 18 tumour sites and assuming a screening threshold of at least 1% of staining (to increase test sensitivity) [91]. The screening positivity rate is assumed to be the same for WGS. The unit costs for each type of test are shown on Table 56, and the numbers needed to screen and full testing strategy cost by tumour site, alongside the tumour site weights as per the larotrectinib integrated efficacy analysis in Table 57.

Table 56 Unit costs of NTRK fusion testing

Test	Cost per test	Source
IHC	£75	TA406 [92] Assumes midpoint of range considered by the Committee
NGS	£350*	Cost for RNA-based panel for all types of NTRK fusion Dr Helene Schlecht, personal communication
NGS+	£0	Dr Helene Schlecht, personal communication.
WGS	£0	Assumption
*Cost for RNA-based panel for all types of NTRK fusion; NGS+, Expansion of existing RNA-based NGS panel to detect NTRK fusions		

Table 57 Cost per strategy

Tumour site	Number needed to screen	Strategy	Cost per patient	Tumour site weight
NSCLC	[redacted]	NGS+	[redacted]	7%
Salivary (non-MASC)	[redacted]	IHC + NGS	[redacted]	7%
MASC	[redacted]	NTRK fusion testing	[redacted]	10%
Melanoma	[redacted]	WGS + NGS	[redacted]	7%
Colorectal	[redacted]	IHC + NGS	[redacted]	6%
Appendix	[redacted]	IHC + NGS	[redacted]	1%
GIST	[redacted]	WGS + NGS	[redacted]	5%
STS adults (non-GIST)	[redacted]	WGS + NGS	[redacted]	9%
Bone Sarcoma	[redacted]	IHC + NGS	[redacted]	2%
STS paediatric	[redacted]	WGS + NGS	[redacted]	12%
IFS	[redacted]	NGS	[redacted]	13%
Breast (non-secretory)	[redacted]	IHC + NGS	[redacted]	1%
Cholangiocarcinoma	[redacted]	IHC + NGS	[redacted]	2%
CNS adults	[redacted]	IHC + NGS	[redacted]	3%
CNS paediatric	[redacted]	WGS + NGS	[redacted]	6%
Pancreas	[redacted]	IHC + NGS	[redacted]	1%
Thyroid	[redacted]	IHC + NGS	[redacted]	10%

The weighted overall cost of testing for larotrectinib applied in the model is £18,618.

10.5 ERG's exploratory analyses: PAS price

The section presents the results with the PAS discount of the company's adjusted base-case, and the ERG scenario analyses 1 to 14 reported in Section 6 of the ERG main report. Scenario 12.1 and 12.2 correspond to the ERG's base-case. All results are presented in **Error! Reference source not f ound.**1.

The ERG base case cost-effectiveness results generate an ICER for larotrectinib of [redacted] and [redacted] per QALY gained for the ePAS2 population and the full integrated efficacy analysis respectively. Considering the cost of NTRK gene fusion testing increases the ERG base case ICER to [redacted] per QALY gained (ePAS2 population), and to [redacted] per QALY gained (integrated efficacy analysis population).

Table 1 ERG's scenario analyses deterministic cost-effectiveness results with PAS price

Scenario	**Larotrectinib**			**Comparator**			**Incremental**			**ICER (/QALY)**
	Costs	**QALYs**	**LYG**	**Costs**	**QALYs**	**LYG**	**Costs**	**QALYs**	**LYG**	
Adjusted Company's Base Case	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
1. Response-based survival approach, Weibull for OS and PFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2. Response-based survival approach, Exponential for OS and PFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
3. Response-based survival approach, Gompertz for OS and PFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
4. Response-based survival approach, Gompertz for OS and Weibull for PFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
5. Company survival approach, Weibull for OS and PFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
6. Company survival approach, Exponential for OS and PFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
7. Company survival approach, Gompertz for OS and PFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
8. Company survival approach, Gompertz for OS and Weibull for PFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
9. Scenario 4 + post-progression survival equal for larotrectinib and comparator	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
10. Scenario 4 + post-progression survival for larotrectinib same as comparator overall survival	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
11.1. Scenario 4 + same post-progression utility for larotrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
11.2. Scenario 10 + same post-progression utility for larotrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

12.1. Scenario 4 + response rate from ePAS2 (ERG base-case)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12.2. Scenario 4 + full integrated efficacy population (ERG base-case)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
13.1. Tumour-specific response rate – IFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
13.2. Tumour-specific response rate - Colorectal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
14.1 Inclusion of NTRK Testing Cost + Scenario 12.1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
14.2 Inclusion of NTRK Testing Cost + Scenario 12.2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

National Institute for Health and Care Excellence
Centre for Health Technology Evaluation

ERG report – factual accuracy check

Larotrectinib for treating NTRK fusion-positive advanced solid tumours [ID1299]

You are asked to check the ERG report to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **5pm on 4 September 2019** using the below comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the committee papers.

The factual accuracy check form should act as a method of detailing any inaccuracies found and how and why they should be corrected.

ERG Responses

No.	Description of problem	Description of proposed amendment	Justification for amendment	ERG response
1	This is not an inaccuracy. However, CIC marking is no longer needed for the proposed marketing authorisation. Page 12, section 1.1; page 29 section 2.2.1; page 36, section 3.1. Also, page 20, section 1.7.3 'suitable' no longer needs to be CIC. Also, unsatisfactory on page 29, section 2.2.1, the last sentence of paragraph 2 no longer needs to be CIC. Reference to the anticipated marketing authorisation on page 93 also no longer needs to remain CIC.	The wording can now be included as follows throughout the document without the CIC marking as per the EMA website: "Vitrakvi as monotherapy is indicated for the treatment of adult and paediatric patients with solid tumours that display a Neurotrophic Tyrosine Receptor Kinase (NTRK) gene fusion, • who have a disease that is locally advanced, metastatic or where surgical resection is likely to result in severe morbidity, and • who have no satisfactory treatment options." https://www.ema.europa.eu/en/medicines/human/summaries-opinion/vitrakvi	As we now have a positive CHMP opinion, the wording of the proposed marketing authorisation no longer needs to retain the CIC marking	Updated markings made as proposed.
2	Page 13, section 1.2. The report states: 'Fourteen different cancers were represented…'. This is incorrect.	Please replace the wording with: 'Fifteen different cancers were represented…'	For factual accuracy. Fourteen different cancers are represented in the ePAS2 dataset. Additionally, Primary CNS tumours are represented in the SAS3 dataset, making a total of 15 different	Correction made as proposed

			cancers. Also this number should be marked AIC.	
3	The report refers in a number of places and indeed, identifies as a potential area of uncertainty, that [redacted] post-progression patients received LOXO-195. This is a factual inaccuracy introduced by Bayer and we apologise for causing confusion. The figure of [redacted] patients actually relates to the number of patients as of June 2019 who had previously received larotrectinib as part of the trial programme who had gone on to enter the LOXO-195 study or receive LOXO-195 through a compassionate use programme. It was not the number of patients in the July 30th 2018 data cut off of the ePAS2 and SAS3 dataset who had received LOXO-195 at the point of that data cut-off. This number is in fact [redacted] and we have verified this with our global colleagues. Page 13, section 1.2; Page 14, section 1.3; Page 54, section 4.2.3.3; Page 73, section 4.7.1; Page 97, section 5.2.4.2; Page 171, section 6.3.4	We suggest that an erratum is added or a footnote throughout highlighting this unintended error in the information Bayer shared at the point of ERG clarification questions.	We recognise that this factual inaccuracy has been introduced by Bayer, but it is appropriate to correct this as it is a particular issue drawn out by the ERG. We apologise for causing this confusion.	This has been updated directly in the amended report.
4	The report refers in a number of places to the number of patients who received larotrectinib post-progression being [redacted]. This is another factual inaccuracy introduced by Bayer, as we have been advised that [redacted], not [redacted] patients continued larotrectinib post progression. Apologies for any confusion caused. Page 15, section 1.3; Page 54, section 4.2.3.3; Page 73, section 4.7.1; Page 97, section 5.2.4.2; Page 135, section 5.2.8.2; Page 171; section 6.3.4	The total number of patients with post progression larotrectinib was [redacted]. Please update this is all sections as highlighted in the column to the left.	We recognise that this factual inaccuracy has been introduced by Bayer, but it is appropriate to correct this as it is a particular issue drawn out by the ERG. We apologise for causing this confusion.	This has been updated directly in the amended report.

	We have also received up to date data on those who continued post progression larotrectinib at the data cut off (not). Page 54, section 4.2.3.3.	Please change the sentence to: of the 93 patients continued to receive larotrectinib post-progression with the duration of treatment ranging from days (patients continuing to receive treatment)		
5	The problem is as referred to in issues 3 and 4 above. Page 129, section 5.2.7.2 also reports the total number of patients with post progression larotrectinib and LOXO-195 and this number should be (received LOXO-195 and of these , had also continued larotrectinib post progression (out of who received larotrectinib post-progression)	The total number of patients with post progression larotrectinib and LOXO-195 was . The sentence should be updated as follows: '……the use of post-progression treatments such as larotrectinib and LOXO-195 in of the patients who had progressed in the ePAS2 dataset…….'	We recognise that this factual inaccuracy has been introduced by Bayer, but it is appropriate to correct this as it is a particular issue drawn out by the ERG. We apologise for causing this confusion.	This has been updated directly in the amended report.
6	The problem is as referred to in Issue 3 above. Specifically, page 144, section 5.2.9.1 refers to a large proportion of patients in the larotrectinib integrated efficacy analysis receiving post-progression treatments such as LOXO-195 and larotrectinib.	The ERG may now wish to reconsider the wording 'large proportion'.	We recognise that this factual inaccuracy has been introduced by Bayer, but it is appropriate to correct this as it is a particular issue drawn out by the ERG. We apologise for causing this confusion.	The word large has been removed from this sentence
7	This may be a matter of interpretation, but the ERG report states on page 14. Section 1.3: The division between the submission's efficacy datasets	The datasets used in the submission are clearly described on pages 43 and 62-63 of the company submission. The datasets are	The ERG report implies that Bayer have not been clear	Not a factual inaccuracy

	('ePAS2', and 'SAS') appears quite arbitrary and was not clearly justified in the submission. Bayer believe the populations were clearly defined in the submission. This comment also applies to the same text found on page 73, section 4.7.1.	also referred to descriptively in the response to ERG clarification question A4. Bayer would therefore suggest that this sentence is deleted from the ERG report (in all places where it appears).	over the datasets used which is incorrect.	
8	Page 15, section 1.4, of the ERG report states: 'In the base case analysis, established practice consisted of a composite comparator represented by a weighted average of comparators from the tumour types represented in the integrated efficacy analysis for larotrectinib, this is referred to as the historical comparator.' The company submission did not explicitly call the weighted average comparator arm 'historical comparator' and wanted to be careful not to mix the approach with matched control from EMR data.	Please remove the word 'historical'. The sentence should read: 'In the base case analysis, established practice consisted of a composite comparator represented by a weighted average of comparators from the tumour types represented in the integrated efficacy analysis for larotrectinib, this is referred to as the comparator.' **Or**, if ERG prefer to refer to the comparator this way, this sentence can be modified to – 'In the base case analysis, established practice consisted of a composite comparator represented by a weighted average of comparators from the tumour types represented in the integrated efficacy analysis for larotrectinib, this is referred to as the historical comparator in this ERG report.'	Historical comparator is usually interpreted as the matched control approach, where EMR database is reviewed to identify a group of patients having the same baseline characteristics with those in the intervention arm. However, the company submission, while still aiming to use studies with comparable populations, leveraged published studies of solid tumours included in the larotrectinib clinical trial programme. Individual characteristic match and adjustment were not performed. The CS did not explicitly refer to the comparator arm as 'historical comparator'.	This has been updated to "this is referred to as the historical comparator in this ERG report.'

9	Page 16, Section 1.4, of the ERG report states: 'Within the PFS and PD health states, the model distinguished between patients who are receiving treatment and those who are not.' This is not the case for situations where treat-to-progression is applied.	Please revise this sentence to: 'The base case assumed treat-to-progression for the larotrectinib arm and followed the treatment duration for comparator treatments in the source documents. A scenario was available to apply the time to treatment discontinuation curve from the larotrectinib clinical trial programme.'	There is no explicit separation of on-treatment versus off-treatment within the PFS and PD health states. It is true that a patient can be on or off treatment within PFS or PD states, but treatment duration is an independent indicator in the model for calculating treatment cost.	Correction made as proposed
10	Page 16, Section 1.4 of the ERG report, there is CIC information not highlighted (selected model for OS and PFS).	Please revise the sentence to – The models selected for the company's base-case analysis were extrapolated [redacted] OS and PFS survival functions.	To preserve CIC marking.	Correction made as proposed
11	Page 19, section 1.7.1 states that the primary analysis included only 93 patients. This is incorrect. This paragraph also refers to a maximum of 17 patients in any tumour site group. This is incorrect.	The primary population for analysis in the submission and used within the economic modelling was 102 patients (93 patients in the ePAS2 set and 9 from the SAS3 dataset). The maximum number of patients in any tumour site group is 25 (please see table 32 in the submission).	For factual accuracy.	The first issue is not a factual inaccuracy. The submission uses the terminology 'primary analysis set' to refer to the 93 patients. For the second issue, we considered adult and child STS cases as distinct sites. This has been clarified in the report.
12	Page 19, section 1.7.2 of the ERG report AND Page	Bayer suggest the wording is changed as	To ensure an accurate description of the	Correction made

	188, section 8.1.2 states that 'Survival curves were estimated for each tumour site using data from past NICE TAs.' This is an incomplete description of the source data as, in some cases, no TA was identified. This paragraph also states '…..; there were some tumour sites with no TA, so assumptions were made by equating survival across similar tumour sites….' This is also an incomplete description of the source data. Lastly, this paragraph refers to the 'somewhat arbitrary selection of previous TAs'. Bayer disagree with this assessment based on our description of source selection criteria in the company submission and also response to the ERG clarification question A14. • This comment also applies to the word 'arbitrary' on page 58, section 4.4.1	follows: 'Survival curves were estimated for each tumour site using data from past NICE TAs where identified.' Bayer suggest the wording is changed as follows: ….; there were some tumour sites with no TA, and either a source from the literature review was used, or where no suitable data could be found, assumptions were made by equating survival across similar tumour sites….' Bayer respectfully request that the words 'somewhat arbitrary selection of previous TAs' are removed from the report as this is not fair or accurate.	source, and selection of the source, for the survival curves.	broadly as proposed We note that in section 8.1.2 the clarification requested was already present. Not a factual inaccuracy
13	Section 1.7.4 on page 21 of the report states: The ERG considers that, by only exploring alternative parametric distribution for the historical comparator analysis, the company do not fully explore the implications of the large variation in survival times for larotrectinib and the ICER according to different parametric distributions. This is incorrect.	Bayer do explore alternative parametric fits for larotrectinib and all comparators. Please refer to table 57 of the company submission and delete this wording.	The statement is factually inaccurate.	This has been amended in the updated report to clarify that this refers to the lack of exploration with the response based model. Text added "and not the response based model"

14	Page 23, section 1.8.2 of the ERG report: Subgroup level overall response rates (ORR) are reported incorrectly. AIC highlighting is missing if this is the ORR from the company submission. This also applies to Page 180, section 6.5.3.	Please correct the ORR rates for infantile sarcoma and colon cancer. They should be [redacted]% and [redacted]%.	For factual accuracy and to preserve AIC marking. Please see table 82 Appendix E.	These ORRs are from the BHM. “with ORR estimated from the BHM” has been added to the updated report These have been marked as AIC
15	Page 24, section 1.8.2. The ORRs mentioned do not specify that they are based on the ERG’s Bayesian Hierarchical model (BHM) and could be misinterpreted to be based on the company submission. ORRs mentioned on page 165, 176, 181, 182 and 184 also do not specify that they are based on the ERG’s BHM. They should also be marked AIC.	Please indicate in the text that these ORRs are based on the BHM.	For clarity in presentation of the clinical data. While these values represent the output of the ERG’s BHM, they are representative of proprietary data and in the context of the report, they provide an indication of the base case values from the manufacturer. As such they should be marked AIC.	See above. This has been noted as generated for the BHM and marked AIC
16	Section 2.1.1 on page 25 of the report refers to the estimated eligible population being [redacted] in 2020. Although this is a minor point, the number we estimate is <[redacted].	Please add a ‘<’ symbol before the estimate of [redacted].	This is a minor point but improves the accuracy.	Correction made as proposed
17	Section 2.1.1, page 28, Table 2 contains numbers which indicate the estimated eligible population, which have been marked elsewhere as CIC e.g. page 12.	The last 4 four columns in table 2 should be marked as CIC.	To maintain CIC on patient estimates.	Correction made as proposed

	The same comment applies to Table 4on page 35.	All columns with numbers in table 4 should be marked as CIC.		
18	Page 37, section 3.1 of the report states. 'Around one third (28/102) (27%), (28/93 (30%) if excluding CNS patients))…' This is incorrect.	Please replace with the following text: 'Around one third (34/102) (33%), (28/93 (30%) if excluding CNS patients))…'	For factual accuracy. There were 28 patients in ePAS2 <18 years of age and 6 patients in SAS3, as per Table 8 of the company submission.	Correction made as proposed
19	Page 37, section 3.1 of the report states: 'The number of patients with NTRK2 fusions is low 10/102 (9.8%), primarily in patients with a primary CNS tumour (9/10)'. This is incorrect.	Please replace with the following text: 'The number of patients with NTRK2 fusions is low 10/102 (9.8%), primarily in patients with a primary CNS tumour (7).'	For factual accuracy. As per Table 8 of the company submission.	Correction made as proposed
20	Page 37, section 3.1 of the report states: 'The ERG also note that [redacted]/102 ([redacted]) of patients in the clinical trial evidence have an Eastern Cooperative Oncology Group (ECOG) performance status of 1 or 0….' This is incorrect.	Please replace with the following text: 'The ERG also note that [redacted]/102 ([redacted]) of patients in the clinical trial evidence have an Eastern Cooperative Oncology Group (ECOG) performance status of 1 or 0….'	For factual accuracy. As per Table 8 of the company submission.	Correction made as proposed
21	Page 48, section 4.2.2 of the ERG report states: 'Most patients had an ECOG performance status of 0 or 1 (89%) or a Karnofsky / Lansky score of >90 (78%) and all stages of diagnosis (I-IV) were represented, though this was unknown in 18% of patients.'	Please replace with the following text: 'Most patients had an ECOG performance status of 0 or 1 (89%) and all stages of diagnosis (I-IV) were represented, though this was unknown in 18% (17/93) of patients in ePAS2 and 44% (4/9) patients in SAS3'.	Performance status scores were equated to ECOG scores for the pooled analysis. The Karnofsky / Lansky score % refers to the NTRK+ patients in the SCOUT study	Correction made as proposed

	This is incorrect.		and not all of these patients are included in the efficacy pooled analysis sets. In the 102 patients available for efficacy analysis, stage at diagnosis was not reported or unknown for 21 patients (21%) or put another way unknown in 18% (17/93) of patients in ePAS2 and 44% (4/9) patients in SAS3.	
22	Page 50, section 4.2.3.1, of the ERG report states: '…..Whereas the ePAS population included a tiny minority of NTRK-2 patients (3%), 7 out of the 9 SAS3 patients (78%) were NTRK 2.' The description of the dataset is incorrect and there is an absence of AIC marking,	Please replace with the following text: '…..Whereas the ePAS2 population included a tiny minority of NTRK-2 patients (3%), 7 out of the 9 SAS3 patients (78%) were NTRK-2.'	Correction of the dataset referred to plus AIC highlighting.	Correction and AIC highlighting made as proposed
23	Page 51, section 4.2.3.1 of the ERG report states: '…In response, the company provided a conference abstract which had a data cut-off date of 19 February 2019 and which reported results for a total of 159 patients….No subgroup results were reported and it was unclear whether or not the cohort of 159 patients included primary CNS patients. Moreover, the results reported were based on investigator assessments – as noted earlier…' Further clarity on the evaluable patient numbers and	Please replace with the following text: '…In response, the company provided a conference abstract which had a data cut-off date of 19 February 2019 and which reported results for a total of 159 patients (153/159 patients evaluable for efficacy)…… No subgroup results were reported and it was unclear whether or not the cohort of 159 patients included primary CNS patients. Moreover, the results reported were based on investigator	The response data were evaluated from 153 patients not the total 159 patients. The conference abstract was marked AIC, therefore information contained within or comments on what it didn't contain (which can elude to its	Additions made as proposed.

	AIC marking to be added.	assessments – as noted earlier…'	contents) should be AIC.	
24	Page 51-52, section 4.2.3.1 of the ERG report states: '…Of these [redacted] patients, [redacted] ([redacted]) were still in response (at the last data cut) and [redacted] had progressed disease; the median duration of response [redacted]' This should be marked as AIC.	Please replace with the following text: '…Of these [redacted] patients, [redacted] were still in response (at the last data cut) and [redacted] had progressed disease; the median duration of response [redacted])….'	This data is marked AIC in the company submission documents.	AIC highlighting made as proposed
25	Page 57, section 4.4.1 of the report states: 'For each tumour site Kaplan-Meier data were extracted from the selected TA and digitised.' This is not completely factually accurate.	The sentence should be replaced as follows: For each tumour site Kaplan-Meier data were extracted from the selected data source and digitised.	For factual accuracy.	Correction made as proposed
26	Page 63, section 4.5.3 states that no grade 3 or 4 SAE data were presented for larotrectinib. This is not correct. Table 28 of the company submission presents Grade 3 or 4 TEAEs (all and drug-related) occurring in ≥2% of patients with NTRK fusion-positive cancer in the pooled analysis of larotrectinib clinical trials (safety analysis set).	Bayer request the last sentence of 4.5.3 be removed.	For factual accuracy.	This relates to the comparator therapies, not larotrectinib. The text has been amended for clarity to "…since no data were presented on grade 3 or 4 SAEs for comparator therapies"
27	Page 72, section 4.7.1 refers to the largest number of patients at any one site (17) being for MASC. This is not correct. There were 17 salivary gland cancers (10 of which were categorised as MASC) – see Table 1, on page 34 of the response to the ERG	The change depends on what the ERG want to convey here. The largest single group is soft tissue sarcoma (21 patients).	For factual accuracy.	Correction made to more clearly express our intention

	clarification questions (Question A9).			
28	Page 78, Section 5.1.2 of the ERG report states: 'A date limit was applied to restrict retrieval of articles published from 2008 onwards and retrieval of conference abstracts in some of the search strategies for EMBASE were limited to 2015 onwards.' There was no limitation to the searches. Date limitation was applied when determining study inclusion/exclusion, and this limitation was only applied to NSCLC searches and the limitation to conference abstracts was 2017 onwards.	Please revise the sentence to: 'The limits for human studies and specific publication types were applied to the EMBASE searches. No other limits were applied to the searches. A date limit was applied during study inclusion/exclusion phase and was applied to NSCLC only. It restricted retrieval of articles to those published from 2008 onwards and retrieval of conference abstracts to those from 2017 onwards.'	The revision ensures alignment with Table 89 of Appendix G of the company submission and Table 8-2 of Appendix 2 of the response to the ERG clarification questions.	Correction made as proposed
29	Section 5.1.4, page 79 of the ERG report states: 'For the second search, the results presented by tumour site in Appendix 2 (company's response to points for clarification) are in tabular form without any accompanying narrative.' Narrative is available in summary format on page 15-21 of Appendix 2 and in more detail by tumour site and by endpoint on page 31-644 of Appendix 2.	Please revise the sentence to – 'For the second search, the results by tumour site and by endpoint of interest are presented in both tabular format and in narrative in Appendix 2 (company's response to points for clarification).'	This revision ensures alignment with contents of Appendix 2 from company's response to points for clarification.	Correction made as proposed
30	Page 98, section 5.2.4.2 of the ERG report presents ICER values for scenarios the incorrect way around.	ICER values should be updated as follows: 'The first scenario reduces the ICER to £[redacted] per additional QALY, while the second has greater impact reducing the ICER to £[redacted] per additional QALY (see Table 38).'	ICERs reported the wrong way around.	Correction made as proposed
31	Page 101, section 5.2.6.1 states: 'The ERG notes that there is lack of consistency between the estimates of median PFS for the ePAS2	Bayer actually stated in our response: 'Regarding the difference in PFS highlighted by the ERG between that	For factual accuracy.	This has been amended to "the company stated that the different data-cuts,

	population (n=93) in the integrated efficacy analysis (i.e. excluding CNS patients), 27.4 months, and for an earlier data cut (n=47, 17th July 2017) of the NAVIGATE trial (p78 of the corresponding clinical study report), [redacted] months. In response to the request by the ERG for further explanation, the company stated that this difference in median PFS is solely due to the different data-cuts from which these estimates were sourced (see company's clarification additional response),……' This misrepresents the response from Bayer. Please address this.	reported in the model and that in the NAVIGATE CSR, we would like to direct the ERG to consider the date of the interim CSR that was provided in response to question A13. The interim CSRs were dated December 2017 - January 2018, reflecting the data cut-off of 17th July 2017, whereas the pooled data we presented in our submission, which formed the basis of the EMA regulatory submission, was based on a data-cut off of July 2018. The difference in the time point of data cut off therefore is ***a factor in the difference*** highlighted by the ERG'. We did not state that the different data-cuts were the sole reason for any difference.		from which these estimates were sourced, is a factor in the difference in median PFS"
32	This was an omission by Bayer – apologies, but Figure 22 on page 120 (Figure 111, Appendix L) should be marked as CIC	Please mark this figure as CIC.	CIC data to be marked.	Correction made as proposed
33	Page 134, section 5.2.8.2 refers to 'the expected price list (June 2018). This should read 2019.	Please change 2018 to 2019.	For factual accuracy.	Correction made as proposed
34	Page 143, section 5.2.9.1. The ERG report states that: 'The majority of QALY gains for larotrectinib were generated within the 'Progressed' state and then in brackets quote a figure (percentage). This figure should be CIC.	Please mark this figure as CIC.	CIC data to be marked.	Correction made as proposed
35	Page 146, section 5.2.9.2, table 34 has data which is CIC but this is not marked. Relating to drug costs, the figure of [redacted]mg should be	Please mark this figure as CIC.	CIC data to be marked.	Correction made as proposed

	CIC.			
36	Page 161, section 6.3.1 of the ERG report states that: 'In order to explore a number of uncertainties highlighted in Section 5.3, the ERG further adapted the company's response-based scenario model, which only included alternative survival distributions for the non-responder (proxy for comparator survival in one of the company's scenario).' This is incorrect as the model actually allows for larotrectinib survival curves to change within the responder/non-responder analysis.	Please amend the text as follows: 'In order to explore a number of uncertainties highlighted in Section 5.3, the ERG further adapted the company's response-based scenario model' ~~, which only included alternative survival distributions for the non-responder (proxy for comparator survival in one of the company's scenario)~~.	The model does include this functionality – changing the PFS selection on the "Settings" page will change the larotrecinib PFS distribution while changing the OS model selection in cell QA9 on the "Survival Curves" page will change the OS distribution.	Not a factual inaccuracy The curves for only larotrectinib responders were not parameterised. Switching the distribution in QA9 did not allow choosing between curves that were not there.
37	The following pages of the ERG report report the response rates from the BHM: • Page 163, paragraph 2 • The table on page 164, row 5, the bottom of this table which extends to page 165, • Page 175, paragraph 3 and the headers of table 47 on this page, • Page 176, paragraphs 2 and 3, along with table 48 on this page These values should be marked as AIC.	Please mark response rates as AIC.	While these values represent the output of the ERG's BHM, they are representative of proprietary data and in the context of the report, they provide an indication of the base case values from the manufacturer.	Correction made as proposed
38	Page 178, section 6.5.1 of the ERG report makes an incorrect statement that heterogeneity in survival outcomes was not explored for larotrectinib,	Please remove this statement and clarify with the following: Heterogeneity was explored in the CS by means of univariate OS-Cox models using the ePAS dataset and that nearly all	Heterogeneity was explored by means of univariate OS-Cox models, see Table 103 of Appendix L	This has been amended to "the CS does not fully explore heterogeneity in survival outcomes for

		univariate models generated extremely large hazard ratios and wide 95% Cis. These findings stress the large uncertainty given the number of patients and events in the sample. Consequently, no further steps were undertaken to run a multivariate-adjusted Cox model.		larotrectinib and importantly the consequences of any differences for cost-effectiveness"
39	Page 185, section 7 of the ERG report refers to 'exponential modelling of the comparator data (provided in the supplied economic model)' and this seems to be reported in Table 52, column 2. Bayer was unable in the time available to validate from where in the model the ERG had taken this data.	Please provide further clarity on where this data has been sourced from.	For clarity.	This was from the supplied economic model excel file: survival inputs worksheet, taking the exponential shape parameters for each tumour site (where reported). For simplicity, this has not been added to the report
40	Page 187, section 8.1.1 states: 'The primary analysis included 93 patients, with a further patients with primary CNS tumours also considered.' There appears to be text missing. Also……'; with a maximum of 17 patients in any tumour site group'. As per issue 27 above, this is not correct. AIC marking is missing.	Regarding the first point, the sentence should be amended as follows: 'The primary analysis included 93 patients, with a further 9 patients with primary CNS tumours also considered.' The change depends on what the ERG want to convey here. The largest single group is soft tissue sarcoma (21 patients).	For factual accuracy. To preserve AIC marking.	Correction made as proposed

41	Page 188, section 8.1.3 states: The approval is for whenever no “suitable” alternative exists’. This is not factually accurate.	The sentence should be reworded as follows: The approval is for …….. and who have no satisfactory treatment options.	For factual accuracy.	Correction made as proposed
42	Page 190-292, section 8.2. Not a factual accuracy, but Bayer suggest further clarity over the price used in the ERG generated ICERs. The Report states: The ERG base case cost-effectiveness results generate an ICER for larotrectinib of [redacted] and [redacted] per QALY gained for the ePAS2 population and the full integrated efficacy analysis respectively.	If correct and appropriate, please add the following in brackets after these sentences: ‘(exclusive of the confidential PAS)’.	For factual accuracy.	Correction made as proposed
43	Throughout it is not clear to the reader which price the ICERs relate to. There is a mixture of reporting of those relating to the original price, those reflecting the new price and ICERs generated by the ERG. It is not clear throughout which price has been used by the ERG and we have not had the opportunity to cross check all of these in the time available. Further, there is reference on page 142 to a confidential PAS addendum which has not been shared with Bayer so has not been checked for factual accuracy.	The ICERs should all be updated using the updated price and all of the updated analysis that was provided by Bayer in the timelines agreed with NICE. Further, the confidential PAS addendum should be shared with Bayer for a factual accuracy check.	For factual accuracy.	The PAS addendum has been moved to the main report appendix and is referred to in section 5, page 146 In section 5 we have added throughout that results reported I the main text are exclusive of the confidential PAS

Evidence Review Group's Report

Larotrectinib for treating NTRK fusion-positive advanced solid tumours (addendum)

Additional analyses requested by the NICE technical team

Produced by CRD and CHE Technology Assessment Group, University of York, Heslington, York YO10 5DD

This document reports the results of additional analyses requested by the NICE technical team prior to the technical engagement meeting. The additional scenarios replicate the ERG alternative base-case analysis for the full integrated efficacy analysis (objective response rate [ORR] estimated by the Bayesian hierarchical model of 57%) presented in Section 6.4 of the ERG report (Table 48) and combine it with the following assumptions:

- Same post-progression survival for larotrectinib and comparator
- Post-progression utility independent of treatment (for larotrectinib and comparator)
- NTRK fusion testing costs (with and without)

Results are presented for the updated pricing structure for larotrectinib submitted by the company on the 23rd July 2019, both with and without a confidential patient access scheme (PAS) discount, consisting of (Table 1 and The scenario combining the NICE technical team assumptions, generates an ICER for larotrectinib of per QALY gained when excluding the costs of NTRK fusion testing (without PAS) and £256,957 when including testing costs (without PAS).

Table 2 Additional scenario (with PAS price)

	Larotrectinib	**Comparator**	**Incremental**
Total LYG			
Total QALYs			
Total costs w/o testing costs			
Total costs with testing costs			
ICER (per QALY)	**w/o testing costs**		
	with testing costs		

, respectively). A detailed breakdown of costs and QALYs are presented for Table 1. Due to the model inflexibility, data limitations and time constraints, all results are based on the deterministic analysis.

Table 1 Additional scenario (without PAS price)

	Larotrectinib	**Comparator**	**Incremental**
LYG			
Progression-free			
Progressed disease			
Total LYG			
QALYs			
Progression-free			

Progressed disease	[redacted]	[redacted]	[redacted]
AEs	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]
Costs			
Progression-free	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]
End of life care	[redacted]	[redacted]	[redacted]
AEs	[redacted]	[redacted]	[redacted]
Treatment	[redacted]	[redacted]	[redacted]
Testing costs	[redacted]	[redacted]	[redacted]
Total costs w/o testing costs	[redacted]	[redacted]	[redacted]
Total costs with testing costs	[redacted]	[redacted]	[redacted]
ICER (per QALY)	**w/o testing costs**		[redacted]
	with testing costs		[redacted]

The scenario combining the NICE technical team assumptions, generates an ICER for larotrectinib of [redacted] per QALY gained when excluding the costs of NTRK fusion testing (without PAS) and £256,957 when including testing costs (without PAS).

Table 2 Additional scenario (with PAS price)

	Larotrectinib	Comparator	Incremental
Total LYG	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]
Total costs w/o testing costs	[redacted]	[redacted]	[redacted]
Total costs with testing costs	[redacted]	[redacted]	[redacted]
ICER (per QALY)	**w/o testing costs**		[redacted]
	with testing costs		[redacted]

Including the confidential PAS reduces the ICER [redacted] per QALY gained, without testing costs and to [redacted] per QALY gained, with testing costs included.

NICE National Institute for Health and Care Excellence

Technical engagement response form

Larotrectinib for treating NTRK fusion-positive advanced solid tumours [ID1299]

As a stakeholder you have been invited to comment on the technical report for this appraisal. The technical report and stakeholders' responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

We need your comments and feedback on the questions below. You do not have to answer every question. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be summarised and used by the technical team to amend or update the scientific judgement and rationale in the technical report.

Deadline for comments: **5pm Friday 18 October 2019**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the technical report which summarises the background and submitted evidence. This will provide context and describe the questions below in greater detail.
- Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.
- Do not include medical information about yourself or another person that could identify you or the other person.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow. If confidential information is submitted, please also send a second, fully

redacted, version of your comments (AIC/CIC shown as [redacted]). See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	**Lesley Gilmour**
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**Bayer plc**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**Current Situation** • Bayer does not have direct or indirect links with, or funding from, manufacturers, distributors or sellers of smoking products but Bayer provides pesticides for crops, which would therefore include tobacco crops. • Bayer is a member of the Cooperation Centre for Scientific Research Relative to Tobacco (CORESTA) (http://www.coresta.org/) within the scope of recommendations of pesticides used for protection of tobacco plants. • It is also a member of country and EU business federations such as the Confederation of British Industry (CBI) and 'Business Europe', which include tobacco companies. **Past Situation** In 2006, Bayer and its subsidiary Icon Genetics piloted a new process for producing biotech drugs in tobacco plants. Icon Genetics was acquired by Nomad Bioscience GmbH from Bayer in 2012.

Questions for engagement

Issue 1: Prevalence of NTRK gene fusions	
Is the distribution of patients in the pooled analysis generalisable to NHS clinical practice?	NTRK fusion cancer is a rare disease and not all tumour types have yet been captured in the study programme however, patients were recruited sequentially as they presented and no solid tumour type was excluded from the larotrectinib trials. Therefore, in terms of patient identification, we would expect study recruitment to be generalisable to patients identified in clinical practice. The company acknowledge that the distribution of tumour types seen in general practice may vary. Indeed, given that NTRK fusion cancer was not well characterised prior to the development and availability of TRK inhibitors such as larotrectinib, screening for NTRK gene fusions was not widely conducted. As genomic testing becomes more widely adopted across the globe, additional tumour types may be identified where NTRK gene fusions are found. However, due to the mode of action, i.e. specifically targeting the protein product of the NTRK fusion genes (i.e. TRK fusion proteins), irrespective of the location or histology of the tumour, there is no reason to expect these currently unidentified tumour types to behave differently. The NAVIGATE and SCOUT studies are still open for enrollment and it is likely that additional tumour types will be identified and studied. *The company is committed to making this data available should larotrectinib be accepted for use via the Cancer Drugs Fund, thereby attempting to address this aspect of uncertainty.* The company conducted the cost-effectiveness analysis using weights derived from the distribution of NTRK gene fusions in each tumour site in patients enrolled in the larotrectinib clinical trials. In order to assess the impact on the ICER, a sensitivity analysis has been conducted applying in the model NTRK gene fusions weights by tumour site (for those tumour sites where The company has study data) obtained from a systematic literature review to the economic model (Table 1). This had little impact on the ICER.

Table 1: NTRK incidence per tumour distribution comparison

Tumour type	SLR proportions	Clinical trial proportions
STSp	[redacted]	[redacted]
Salivary	[redacted]	[redacted]
Cholangio	[redacted]	[redacted]
GIST	[redacted]	[redacted]
nGIST/bone	[redacted]	[redacted]
Thyroid	[redacted]	[redacted]
CRC/App	[redacted]	[redacted]
NSCLC	[redacted]	[redacted]
Melanoma	[redacted]	[redacted]
Pancreas	[redacted]	[redacted]
CNS/glioma	[redacted]	[redacted]
Breast	[redacted]	[redacted]

When comparing the base case analysis results (Table 2: ICER of £[redacted]) with results of the sensitivity analysis (Table 3: ICER of £[redacted]), it appears that the impact of the variation in the distribution of the NTRK gene fusions by tumour site has little impact on larotrectinib's cost-effectiveness.

An additional responder/non-responder analysis has been run using the distribution of the NTRK gene fusions from the systematic literature review (Table 3: ICER of £[redacted]). The results seem to confirm the limited impact of the variation in distribution on larotrectinib's cost-effectiveness.

Table 2: Base case

	Costs	QALYs	ICER
Comparators	[redacted]	[redacted]	
Larotrectinib	[redacted]	[redacted]	
Incremental	[redacted]	[redacted]	[redacted]

Table 3: Assumptions with epi SLR data distribution

	Costs	QALYs	ICER
Comparators	[redacted]	[redacted]	
Larotrectinib	[redacted]	[redacted]	
Incremental	[redacted]	[redacted]	[redacted]

Table 4: Assumptions with epi data distribution – responder/non-responder analysis

	Costs	QALYs	ICER
Comparators	[redacted]	[redacted]	
Larotrectinib	[redacted]	[redacted]	
Incremental	[redacted]	[redacted]	[redacted]

There is no reason to suggest that the tumour site efficacy (and safety) demonstrated in the larotrectinib studies would not be generalisable to the population found in clinical practice in England.

What is the total number of patients who would receive larotrectinib?

The company acknowledge that the number of patients who may receive larotrectinib is uncertain. A recent systematic literature review of the real world studies reporting on the frequency of NTRK gene fusion has recently been completed. Bayer are updating the epidemiological data presented in the submission and will submit as a piece of additional evidence.

With reference to identifying patients eligible for larotrectinib, please refer to our responses to questions 3, 4, 6 and 7.

Issue 2: Treatment pathway and positioning	
How will the term 'satisfactory' be defined in clinical practice?	The marketing authorisation for larotrectinib places the product in a setting where no satisfactory treatment options remain since patients will have failed to respond to standard of care, did not tolerate it or do not have any standard of care for treatment. According to the draft EPAR: [redacted] Some patients presenting with a disease in which cure through surgery is the therapeutic goal, could have a better outcome with cytoreduction of the tumour with larotrectinib followed by surgical resection, thus avoiding disfiguring amputation *['not a satisfactory treatment option']* and permitting limb salvage. In the less frequent tumour sites such as appendix, salivary gland, and secretory breast carcinoma, there are limited or no treatment guidelines or recommendations due to scarcity of evidence supporting systemic therapy i.e. '*no satisfactory treatment options'*. The licensed population represents a small yet diverse group, ranging from infants to adults with multiple tumour sites / histologies but with a commonality of a high unmet medical need. The license is reflective of the trial population in that according to inclusion criteria, patients: • must have progressed or be nonresponsive to available therapies, be unfit for standard chemotherapy or for which no standard or available curative therapy exists [study 14001], • must have received prior standard therapy appropriate for their tumour type and stage of disease, or in the opinion of the investigator, would be unlikely to tolerate or derive clinically meaningful benefit from appropriate standard of care therapy [study 15002] • have relapsed, progressed or nonresponsive to available therapies and for which no standard or available systemic curative therapy exists OR with locally advanced IFS who would require, in the

	opinion of the Investigator, disfiguring surgery or limb amputation to achieve a complete surgical resection [study 15003] Accordingly, patients enrolled in the trial programme were considered to have no satisfactory treatment options based on clinical judgement, and as such, the trial population should be reflective of how patients would be identified in clinical practice. Indeed, for those patients where response to prior systemic therapy was reported, the ORR to that line of therapy was [redacted]%. *The real world interpretation of 'satisfactory' could be collected within the CDF.*
For each tumour type, at what point(s) in the respective treatment pathways will larotrectinib be used in clinical practice?	In line with the trial inclusion criteria *(see response to the question above)* and license, The company anticipate that larotrectinib will be used in patients who have solid tumours that display a NTRK gene fusion who have a disease that is locally advanced, metastatic or where surgical resection is likely to result in severe morbidity, and who have no satisfactory treatment options. Clinical judgement was used in the trials to identify patients who have no satisfactory treatment options. It is the company's expectation that in practice, the place in therapy will also be subject to some clinical judgement. However it is the company's interpretation of the license that patients who are eligible for other licensed therapies would be offered these therapies before receiving larotrectinib. The licensed therapies available to an individual patient will vary according to the tumour location. So the place in therapy and sequence of prior treatments will vary according to an individual's tumour location. Consistent with this, in the cost-effectiveness model, the company placed larotrectinib as 'last-line' therapy, after patients have exhausted licensed treatment options. The weighted comparator is therefore intended to represent prognosis in patients after they have experienced all licensed treatment options. *The real world positioning in the treatment pathway could be collected within the CDF.*

Is the clinical evidence for larotrectinib generalisable to the positioning in the treatment pathways in clinical practice in England?	There is no reason to believe that the clinical evidence for larotrectinib is not generalisable to how it would be used in clinical practice. The license is reflective of the clinical evidence in that according to inclusion criteria, patients: • must have progressed or be nonresponsive to available therapies, be unfit for standard chemotherapy or for which no standard or available curative therapy exists [study 14001], • must have received prior standard therapy appropriate for their tumour type and stage of disease, or in the opinion of the investigator, would be unlikely to tolerate or derive clinically meaningful benefit from appropriate standard of care therapy [study 15002] • have relapsed, progressed or nonresponsive to available therapies and for which no standard or available systemic curative therapy exists OR with locally advanced IFS who would require, in the opinion of the Investigator, disfiguring surgery or limb amputation to achieve a complete surgical resection [study 15003] Any patients enrolled in the clinical trials without receiving standard of care (i.e. those who received larotrectinib as first line) were determined by the treating physician as being unlikely to tolerate or derive clinically meaningful benefit from appropriate standard of care therapy. The company would expect that this is how larotrectinib would be used in clinical practice in England and, as such, the clinical evidence is generalisable to the positioning in the treatment pathways in clinical practice in England. Accordingly, the trial data and license reflect how these patients would be treated in clinical practice.
Issue 3: NTRK gene fusion testing	
What is the likely screening pathway to identify NTRK fusion positive solid tumours?	Through the National Genomic Medicine Service, NHS England is implementing WGS for all paediatric and sarcoma cancers. The results of this test will indicate the NTRK status of these patients. For all other solid tumours, NHS England is implementing broad panel NGS testing which will include the capability to test NTRK1, 2 and 3.

At what point in the treatment pathway for each tumour type will NTRK gene fusion testing be carried out?	WGS and broad NGS panels will be implemented as part of the broader initial diagnostic processes for patients with cancer.
What proportion of the screening costs should be included in the economic model for this appraisal? Or is it appropriate to exclude all testing costs from the analyses?	The company do not consider that screening costs should be included in the economic model as to do so would not be in line with the NICE methods. Building on the *NHS Five Year Forward View Next Steps* the Board of NHS England initiated the creation of the National NHS Genomic Medicine Service in March 2017. It is clear from all of the communication on the National Genomic Medicine Service from NHS England that the service is not a screening service for individual medicines. The service's intention and design is to provide the necessary genomic profile for all types of cancer to inform diagnosis, staging and treatment. The NICE methods guide states: "*The use of a technology may be conditional on the presence or absence of a particular biomarker (for example a gene or a protein). If a diagnostic test to establish the presence or absence of this biomarker is carried out* ***solely*** *to support the treatment decision for the specific technology, the associated costs of the diagnostic test should be incorporated into the assessments of clinical and cost effectiveness."* Whether it is large panel testing or WGS carried out under the NHS Genomic Medicine Service, there is the potential for a number of actionable targets to be identified. As such, it is not appropriate to assign cost of testing to the assessment of cost-effectiveness of larotrectinib. Indeed, whole genome sequencing will be funded nationally by NHS England (1). Whilst we can see from an academic health economic point of view the rationale for including testing costs, the company do not believe the cost of implementing a National Genomic Service should be considered within NICE's cost effectiveness assessment, according to NICE's own methods guide. However, the company welcome discussions on supporting the implementation of the National Genomic Medicine Service, and in our initial meeting with the Accelerated Access Collaborative we focused on how industry could work with the NHS and other system partners to accelerate the implementation of the

National service. It is in these other forums that industry support for the NHS' investment in the National Genomic Medicine Service rightly sit.

The company acknowledge that equitable testing is currently being implemented but will not be fully operational for a few years. So, despite this not being in line with the very clear statement in the methods guide, we have included a scenario, varying patient numbers, which could be envisaged to be necessary over the short term interim period, where last line patients only are tested. The figure of [redacted] relates back to the originally submitted budget impact model and represented the maximum estimated numbers who may be eligible for testing; i.e. the patients who had completed last line therapy. During the clinical validation interviews, we were advised that [redacted] % of patients with advanced disease would be potentially fit for further therapy, hence the number of [redacted] patients is included in the analysis below. By including the costs (based on the ERG estimated cost per test) for testing [redacted] patients, the revised ICER is £[redacted] using the company base case and £[redacted] using the ERG base case. Alternative costs for testing [redacted] and [redacted] patients have been included to allow for scenarios where it takes longer than one year to introduce NHS England's testing strategy. The company consider this would be the maximum impact and would be short-lived given the cohort of untested late stage patients would reduce over time.

Submitted base case

Number of patients to test	Cost per test	Total cost of NTRK testing	Larotrectinib			Pooled comparator			ICER
			Costs	QALYs	LYs	Costs	QALYs	LYs	
Company base case	-	£0	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ERG base case

Number of patients to test	Cost per test	Total cost of NTRK testing	Larotrectinib			Pooled comparator			ICER
			Costs	QALYs	LYs	Costs	QALYs	LYs	
ERG base case	-	£0	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

1. 2018/2019 final draft National Genomic Test Directory FAQ.

Issue 4: Identification of NTRK gene fusions – diagnostic accuracy	
What is the expected diagnostic accuracy of NGS testing?	WGS and broad NGS panels are used to detect a range of potential genomic alterations therefore they should not be considered as a screening test for NTRK. The intention of the National Genomic Medicine Service is to generate the relevant genomic profile for the cancer patient to inform diagnosis, staging and treatment – The National Genomic Medicine Service is not a 'screening service' for precision medicines.
Is there a testing sequence that could avoid a substantial number false positive results in low NTRK fusion-positive tumour types?	WGS and broad NGS panels are used to detect a range of potential genomic alterations leading to carciogenesis for any given solid tumour and part of NHS England's strategic rationale is that by implementing these broad single tests there will be efficiencies and economies of scale from avoiding multiple single gene tests.

	The intention of the National Genomic Medicine Service is to generate the relevant genomic profile for the cancer patient to inform diagnosis, staging and treatment – The National Genomic Medicine Service is not a 'screening service' for precision medicines.
Is it appropriate to limit testing to avoid false positive results and the associated costs?	The company considers that it does not seem appropriate or equitable to limit testing to particular tumour types when larotrectinib has been designed and is licensed as a histology independent treatment for all patients with solid tumours harbouring the NTRK gene fusion.
Issue 5: Primary CNS tumours	
Should patients with primary CNS tumours be included in the analysis?	The company included the patients with primary CNS tumours in the cost-effectiveness analysis, according to the post-hoc addition of this cohort to the pooled analysis. For further clarity, the primary analysis of ORR excluded primary CNS patients. Exclusion of CNS tumours in the overall efficacy estimates was based on the following rationale: • Data from patients with a primary CNS tumour were evaluated using either RANO or RECIST v 1.1 criteria, whereas other solid tumors were evaluated using RECIST v1.1 only; • Surgery and radiation treatments can lead to varying amount of oedema/inflammation/scarring, which can impact the radiological assessment in these patients; and • The data for patients with primary CNS tumours were not Independent Review Committee (IRC) verified. In addition, lack of progression rather than ORR by RECIST 1.1 or RANO may be a better parameter to describe efficacy of an agent in CNS tumours. As such, the efficacy of larotrectinib on primary CNS tumours was analysed separately.
Issue 6: Trial study design	

Is it appropriate to consider the 'basket' trial design for statistical evidence of heterogeneity?	We have been advised by our statistical experts that the studies in the development programme for larotrectinib were not designed to analyse each cohort separately, given the rarity of the disease. As such, assessing efficacy per tumour type or 'basket' is statistically inappropriate. The rarity of disease prevents a standard approach to assess heterogeneity. Early in the development programme with advice from global regulators, the decision was made to pool efficacy data across all 3 studies from patients with a solid tumour harbouring an NTRK gene fusion. This was possible due to the consistency of treatment response, safety, and tolerability across tumours and age groups for larotrectinib, and the common eligibility criteria and study procedures. The pooled analysis approach provides a more robust estimate of the responses in patients with NTRK fusion cancer and was agreed with regulatory agencies. The pooled analysis was used for both the US Food and Drug Administration (FDA) and the EMA regulatory submissions. As acknowledged by the technical team, NAVIGATE is ongoing and more robust data will be generated with larger patient numbers and longer follow up. *Entry to the CDF would allow the data to mature giving rise to more meaningful analyses.*
Issue 7: Heterogeneity of response across different solid tumour types	
Is a homogeneous response to larotrectinib across different tumour types a reasonable assumption?	The company understands that the histology-independent nature of larotrectinib and the rarity of the NTRK gene fusion cancers targeted by larotrectinib present unorthodox challenges to the traditional technology assessment process. Larotrectinib is an innovative technology that specifically targets the protein product of the NTRK fusion genes (i.e. TRK fusion proteins), irrespective of the location or histology of the tumour, turning off signalling pathways that usually allow NTRK fusion-positive cancers to grow. Larotrectinib is effective across a broad range of tumours including rare tumours and rare subsets of more common tumours, and in paediatric and adult patients ranging in age from [redacted] years.

	The common anomaly of NTRK gene fusion, the consistency in treatment response, safety, tolerability across tumours and age groups, and the common eligibility criteria and study procedures, was the basis of pooling of data on larotrectinib in support of global regulatory submissions. In 2018 guidance, using the ESMO Scale for Clinical Actionability of molecular Targets (ESCAT) grading algorithm, the committee ranked NTRK gene fusions as a Tier 1c actionable driver across tumours. Targets are designated 'tier I-C' if clinical trials in multiple tumour types, or basket clinical trials, have demonstrated a clinically meaningful benefit for the target–drug pair with similar magnitude of benefit across the different tumour types. In this scenario, the clinical value of a target–drug match can be accepted across cancers that harbour the target abnormality. Larotrectinib was provided as an example as it showed substantial antitumour activity in cancers of diverse histological tumour type sharing activating fusions in TRK genes (1). We believe that consideration of response by tumour location only serves as a distraction and introduces the potential for decision-making to be based on chance findings. The totality of the clinical and non-clinical body of evidence supports a histology-independent indication since larotrectinib has demonstrated a large magnitude of effect irrespective of tumour site. We do not believe the uncertainty inherent to small datasets is improved by cutting the data further. Giving consideration to any other therapy e.g. an antihypertensive, natural variation in response between patients would be expected if there was a small sample, but this would not warrant a test for heterogeneity in effect. 1. Mateo et al. A framework to rank genomic alterations as targets for cancer precision medicine: the ESMO Scale for Clinical Actionability of molecular Targets (ESCAT). 2018. Annals of Oncology 29: 1895–1902.
Is the Bayesian Hierarchical Modelling (BHM) framework an appropriate method to capture the heterogeneity of response across different solid tumour types for this appraisal?	Larotrectinib has been developed to treat patients with NTRK fusion positive solid tumours regardless of tumour type and as such, the company do not consider it appropriate to analyse the data by the site of origin of the tumour. A senior team within the company with representation from clinical development, medical, statistics, regulatory, epidemiology, marketing and market access considered the request and decided that is was

	not appropriate to provide the data that the ERG requested for the purposes of BHM to try to characterise uncertainty. The reasons for this are set out below: • The company considers that currently the data are too limited in patient numbers per tumour histology and lacking sufficient follow-up to support analyses of tumour heterogeneity. In particular, the extremely low rates of death events ([redacted] events observed across [redacted] tumour types, and no events in [redacted] tumour types) and the low number of patients per tumour (only [redacted] tumour types have a least [redacted] patients) represent significant challenges to more advanced statistical analyses at the present time. • Whilst developing our approach to HTA, the company discussed Bayesian methods with external HTA experts and opinions were mixed given the current status of the evidence. Experts acknowledged application of the decision problem to a Bayesian framework was currently theoretical and would require an extensive amount of external data to populate which is currently not available. Upon evaluation, whilst theoretically possible, a Bayesian approach was not considered feasible at this time. • The company conducted a targeted literature review to explore published approaches designed to analyse basket trials; no evidence was found of published methods or examples of any application of the Bayesian framework to time to event outcomes. In light of this the company considers that any exploration of a Bayesian framework would be academic and not appropriate as part of this technology appraisal. • The company is discussing post-marketing commitments with the EMA that will provide a much more substantial basis to assess tumour heterogeneity. The company will share these commitments with NICE once a final agreement has been reached. The company is open to reconsider Bayesian analysis at a time when the recruitment of patients and follow-up duration is sufficient to generate meaningful results. The BHM methodology applied by the ERG does not seem to be able to replicate the ORR based on the trial data and is consistently lower irrespective of the chosen prior. Either the methodology is inappropriate, or the information informing this approach is incomplete or incorrect.

	An alternative is to change the prior for mu to Normal(19.3472979, precision=1/10), which would essentially replicate the actual trial results. This seems to be the more valid assumption as there are no alternative datasets available to refute this estimate and all of the data to-date suggests high ORRs with larotrectinib treatment in NTRK fusion positive patients. To assess whether the ERG's BHM is appropriate, the response rates predicted for unknown tumour types should be validated with external sources of information such as clinical opinion.
Would it be appropriate to apply the BHM framework to explore the heterogeneity in the time to event outcomes?	In the ERG's dual-partitioned response-based survival model, surrogacy between response and survival is implied by the model structure. However, no evidence is available to suggest that response is a suitable surrogate for PFS and OS. Therefore, without evidence for surrogacy, no model should imply it's clinically viable. If there were evidence to suggest it is clinically viable and the BHM framework would be applied appropriately as suggested in question 15, ideally the BHM framework would be consistently applied for time to event outcomes in addition to ORR. However, given that there is no evidence available to suggest surrogacy relationship exists and the small numbers of events at the tumour level, the company believe it is not currently meaningful to explore heterogeneity in any outcomes at the tumour level.
Issue 8: Constructing a comparator arm	
Is the company's comparator arm suitable for decision making?	The comparator arm selected is in line with the final scope issued by NICE and our approach to comparator selection was validated with clinicians. Currently, there are no approved treatment options in the UK specifically for patients with NTRK fusion-positive solid tumours and, to date, treatment recommendations regarding NTRK fusion-positive cancer have not been included within any UK guidelines. The approach taken to identifying the comparator arm was to consider standard of care after patients have exhausted all satisfactory treatment options, in line with the marketing authorisation.

	Where possible, relevant NICE TAs were used to source the data for the comparator arm and placebo arms of trial data were selected to represent standard of care in a last line setting where no active treatment option remains. Given the process and scrutiny undertaken in each technology appraisal to select the Committee's preferred inputs and assumptions, these sources were determined to be most suitable for decision making in England, and allowed the data and assumptions used in the model to reflect the Committee's preferred assumptions. This minimises uncertainty, and allows incorporation of input from the wide range of stakeholders who contributed to previous appraisals. Where this was not possible, the proxy standard of care was active treatments not deemed satisfactory (e.g. not approved by NICE and/or not in guidelines or where clinicians have advised may be used in clinical practice but would be considered unsatisfactory) and that are used once all other lines of active treatments have been exhausted. The comparators identified were weighted by patient enrolment per tumour location in the clinical trials and our approach to comparator selection validated with clinical experts. Where alternative comparators were suggested by experts, this was then tested in scenario analysis with minimal impact on the ICER.
Is it appropriate to use non-responders as a proxy for patients not having an active treatment or previous line of therapy to generate a comparator arm for this appraisal?	In response to discussions held during the NICE scoping phase of the appraisal, alternative approaches for controlling for the larotrectinib clinical trial data have been explored: (1) non-responder control analysis and (2) comparison to previous line of therapy. Appreciating the limitations of the different approaches, they all produced very similar ICERs, lending credibility to the results.
Issue 9: Comparator treatments	
Are the comparators identified representative of where larotrectinib would be used in the treatment pathways?	Currently, there are no approved treatment options in the UK specifically for patients with NTRK fusion-positive solid tumours and, to date, treatment recommendations regarding NTRK fusion-positive cancer have not been included within any UK guidelines.

	As referred to in Issue 8, the approach taken to identifying the comparator arm was to consider standard of care after patients have exhausted all satisfactory treatment options, in line with the marketing authorisation. Where possible relevant NICE TAs were chosen as source data at the appropriate point in the treatment pathway. All data chosen for the weighted comparator arm reported that patients had received previous lines of therapy, varying from second line to fifth line. Where, in one of the clinical validation interviews, an alternative treatment was suggested (STS non-GIST, pazopanib), data was sought and then this tested in scenario analysis, with minimal impact on the ICER.
Issue 10: Subsequent therapies	
What subsequent treatments would be expected in clinical practice after larotrectinib?	In line with the marketing authorisation, in clinical practice, larotrectinib is to be used when there are no satisfactory treatment options remaining. As such, the only potentially active treatments a patient may receive in clinical practice would be via clinical trials, compassionate use programmes or individual funding requests i.e. all of which would not be approved for use in clinical practice. This is in line with the trial data and further, this would apply to both arms of the model and is a plausible assumption in the real world. Some patients in the trial programme were able to proceed to limb sparing surgery with curative intent after larotrectinib treatment. This has not been captured within the cost effectiveness model. *Post-progression management could feasibly be collected under the CDF.*
Should experimental treatments be adjusted for in this analysis?	The company do not consider this adjustment should be made. Whilst recognising that the data is immature, the company are asking NICE to consider the plausible benefit which can be achieved through use of larotrectinib in patients with no satisfactory treatment options. Indeed, the draft EPAR states: [redacted] It seems completely unreasonable to assume that any additional post-progression benefit comes entirely from experimental treatments for which there is no data. To our knowledge there is no evidence of any drug (including the experimental drug LOXO-195) demonstrating effectiveness in the post progression larotrectinib or entrectinib setting. Equalising benefit in

	the post-progression setting would therefore appear to be an overly conservative and unreasonable approach, without any basis. Further, after an extensive review of previous NICE technology appraisals since October 2016, there were 48 oncology appraisals in the 'last line' setting and we could find no evidence of the company being criticised for not including experimental therapies in the modelling that may have been received post-progression. As such, we are perplexed as to why this has been requested in the current appraisal. Any experimental treatments used in clinical practice could equally be offered to patients who had received standard of care or larotrectinib, and as such, this would be expected to have a minimal impact on the results of the cost effectiveness analysis. The UK Company have requested data from our global colleagues to further explore the potential impact of post-progression treatments (larotrectinib, LOXO-195, surgery) on the efficacy data. *We would hope this data will be available to us by shortly and we would then submit further analyses to NICE by 25th October.* Further to the comment on the teleconference regarding potential dose escalation again, we are seeking urgent further clarification on this point and *will provide you with detailed analysis by 25th October*. In the meantime, we can assure you that the average dose used in the economic model included post progression exposure and any dose escalation.
Issue 11: Model structure	
What is the most appropriate model structure for this appraisal?	This is the first appraisal of a histology independent treatment. There is no precedence or guidance for evaluating the cost effectiveness of histology independent treatments where activity and clinical evidence is not confined to a particular tumour location. A number of steps were taken to validate the approach taken for the economic evaluation. In order to ensure the scientific rigor of this appraisal, the company partnered with a number of Health Economic advisors. As referred to under Issue 8 above, three different modelling approaches for controlling for the larotrectinib clinical trial data have been explored.

	Appreciating the limitations of the different approaches, they all produced very similar ICERs, lending credibility to the results. The company believe that a simple partitioned survival model that does not require the assumption of a surrogacy relationship between response and survival to be a more appropriate approach for the current appraisal. As stated in the company submission, there are inherent limitations in a response-based partitioned survival approach, and strong assumptions needed to be made to incorporate the analysis into the model: • Low numbers of events, especially important for non-responders, as this substantially reduces the confidence in the overall survival analysis; ○ There are very few non-responders (n=[redacted] patients depending on assessment method); overall response rate was [redacted] % (95% confidence interval [CI], [redacted]) according to independent review and [redacted] % (95% CI, [redacted]) according to investigator assessment. • Uncertainty in the projected survival curves given the relatively short, variable follow-up in the larotrectinib clinical trial programme; • The differences in the distribution of tumour sites/disease severity between responders and non-responders are not accounted for; • The assumption that the non-responders would represent a control arm. Patients on larotrectinib may not respond for a variety of reasons, and may be inherently different to those patients that do respond. The company also believe the methodology implemented to adjust the ratio between progression-free and post-progression survival to be inappropriate. We refer the technical team to a review of studies incorporating relationships between PFS and OS conducted by the NICE DSU team, which states (1): *We have found that the level of evidence available supporting a relationship between PFS/TTP and OS varies considerably by cancer type and is not always consistent even within one specific cancer type. Furthermore, even where robust consistent evidence supporting a correlation between the treatment effects (i.e level 1 evidence according to Elston and Taylor) is available, it is unclear how that should be*

	converted into a quantified relationship between PFS and OS treatment effects within a cost-effectiveness model. Therefore, any cost-effectiveness analysis which makes a strong assumption regarding the relationship between PFS and OS should be treated with caution. Whether OS is confounded by post-progression treatment is dealt with in question 21. 1. http://nicedsu.org.uk/wp-content/uploads/2016/03/PFSOS-Report.FINAL_.06.08.12.pdf
Issue 12: Extrapolation of overall and progression-free survival	
Which parametric curve is most appropriate for extrapolating overall and progression-free survival in the company base case?	The company acknowledge that there is uncertainty around the long-term efficacy and survival profile for larotrectinib. We also acknowledge that the uncertainty is driven by the immaturity of the data, with ongoing data collection providing more certainty of the outcomes over time. With the company's intention for larotrectinib to be made available under the CDF, it is reasonable to consider clinical plausibility rather than seeking definitive answers or taking a highly conservative approach. However, the base case methodology to explore the efficacy and survival profile for larotrectinib into the extrapolated period beyond the trial-based Kaplan-Meier follows the DSU preferred fitting of standard parametric curves, resulting in transparent projections for longer term outcomes. The submitted base case curves for progression-free and overall survival incorporate the most conservative clinically plausible extrapolations, [redacted] for PFS and [redacted] for OS, whilst considering the statistical fit to the Kaplan-Meier through calculated AIC and BIC. All parametric fits were tested through scenario analysis, using national life table mortality to moderate clinically implausible curves. In considering the base case, and immaturity of the OS data, a conservative approach was adopted in considering only the [redacted] and [redacted] model. Both of these models underestimate PFS and OS versus the observed trial data. Applying the [redacted] is a more simplistic approach as it relies on one parameter rather than two; however, it assumes a constant hazard throughout lifetime as it does not account for the change in survival hazards with aging.

	The [redacted] distribution provides closer estimates to the later points of the KM than the [redacted] (whilst still slightly underestimating the observed data) and also tends to be cited as more appropriate for modelling the change in hazards with aging. The [redacted] was therefore selected as the base case for modelling survival of larotrectinib patients for PFS and OS. An assessment of clinical acceptability determined that when using the lognormal, log logistic, Gompertz and generalised gamma distributions, patients overall survival exceeded current UK life expectancy (based on published all-cause mortality rates).

Which parametric curve is most appropriate for extrapolating overall and progression-free survival in the response-based analysis?	Uncertainty around the extrapolated progression-free and overall survival curves applied in the company base case was explored extensively within the submission; including probabilistic sensitivity analysis for the parametric curve parameters using the DSU preferred CHEBS function. Additional scenarios exploring alternative methods for both larotrectinib and comparator survival were performed, including responder/non-responder analysis and application of a naïve Growth Modulation Index to represent a positive hazard ratio for efficacy and survival profiles of larotrectinib. For the company submitted responder/non-responder scenario analysis, a [redacted] distribution was recommended for PFS and OS because it is the most clinically plausible distribution for both PFS and OS and to be consistent with the base case analysis as this scenario was meant to test uncertainty in comparator survival projections. The rationale for the ERG's choice for Gompertz for OS curve is not clear in the ERG report. It seems this distribution was chosen simply based on conservatism. Below we explain the selection process and rationale for curve selection in the company base case to demonstrate why we believe the ERG's selection to not only be too conservative but also clinically implausible. Comparison of the PFS curves for the responder / non-responder analysis showed [redacted] to be most conservative option for PFS. Gompertz and generalised gamma were deemed clinically implausible for PFS. See below the projections for PFS as per Table 118 from Appendix L of the company submission. [redacted]

	For the OS curve selection process, a different trend was observed. See below for the projections for OS as per Table 121 from Appendix L of the company submission. Exponential, log normal, log logistic, and generalised gamma would be unreasonable selections because responders appear to live for a biologically implausible amount of time past the 80 year time horizon. The Gompertz curve has the opposite trend where patients appear to die relatively quickly after for both responders and non-responders. Additionally, the OS percentages at any given point in time are very similar to the PFS percentages, indicating death occurs very soon after progression, and therefore is likely too conservative of an OS curve. [redacted] appeared to be the most biologically plausible curve and it aligned with the base case settings. [redacted] We explored the uncertainty of curve selection in the responder/non-responder analysis, implementing the [redacted] based on the company base case to the ERG preferred responder/non-responder dual-partitioned survival model. This scenario resulted in an ICER of £[redacted], in line with the company submitted responder/non-responder scenario result. The table below provides an overview of the various ICERs resulting when altering the distribution for PFS and OS in the ERG base case model. Applying Gompertz to OS produces a very high ICER inconsistent with the ICERs when assuming any other distribution for OS. Similarly, when applying the Gompertz or generalised gamma for PFS, results in very high ICERs inconsistent with the ICERs when assuming any other distribution for PFS.

Review of ICERs when altering distribution setting for PFS and OS in the ERG base model

	Larotrectinib			Pooled comparator			
Description	**Costs**	**QALYs**	**Life years**	**Costs**	**QALYs**	**Life years**	**ICER**
ERG base case (PFS: Weibull, OS: Gompertz)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OS changed to Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OS changed to Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OS changed to Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OS changed to Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OS changed to Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS changed to Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS changed to Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS changed to Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS changed to Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS changed to Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Issue 13: Drug wastage and adherence	
Have potential drug wastage costs and adherence costs been appropriately included in the model?	The company presented scenario analyses around wastage in response to the ERG clarification questions. These analyses made little difference to the ICER. Regarding adherence, whilst the ERG have made an adjustment to the paediatric dose in their analysis, and in doing so, including increased costs, the clinical benefit of receiving the increased dose has not been modelled.
Issue 14: Administration costs and resource use	
Have administration costs and resource use been adequately captured in the company's model?	The technical report outlines that the uncertainty around administration costs was unclear from the company submission, specifically a lack of inclusion of oral chemotherapy administration costs for larotrectinib. Administration costs for oral chemotherapies is not consistently applied in NICE TAs. Administration costs were not applied in either the larotrectinib arm or for oral therapies in the comparator arm within the economic model. However, we have run a scenario analysis including the cost of one administration of exclusive oral chemotherapy (NHS reference costs code SB11Z - £140.82 per administration) per 30-day treatment cycle, reflecting clinical practice for oral chemotherapies. The inclusion of oral chemotherapy administration costs results in increased costs for larotrectinib and therefore an increase in the ICER from £[redacted] to £[redacted]. It should be noted that the full resource impact of avoiding intravenous chemotherapy is unlikely to have been accounted for in the modelling.

Inclusion of oral chemotherapy administration costs

Scenario category	Description	Larotrectinib			Pooled comparator			
		Costs	QALYs	Life years	Costs	QALYs	Life years	ICER
Base case results		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Inclusion of oral chemotherapy administration for larotrectinib at the start of each 30-day treatment cycle		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Additional queries were based on the justification of administration costs for STS paediatrics, with a perceived inclusion of a weekly cost. STS paediatric comparator treatment of irinotecan in combination with vincristine was included within the economic model with specific treatment regimes. As previously outlined in Table 43 of the original company submission and based on pivotal clinical trials, irinotecan was implemented as 50 mg/m2 per day for 5 days at weeks 1, 4, 13, 25, 34, 46, 49 and vincristine implemented as 1.5mg/m2 on day 1 of weeks 1, 2, 4, 5, 13, 14, 25, 26, 34, 35, 46, 47, 49, 50. As a result, administration costs were only incorporated on the specific days of treatment, in line with the outlined regimes and explains the variation in administration costs each week and the potential to receive treatment up to 50 weeks. Scenario analyses identified that comparator costs were not a key driver of the results, with uncertainty around administration costs for a single tumour location within the mixed comparator not resulting in any significant impact on the ICER.

Issue 15: Implementation and training costs	
What additional infrastructure and training requirements could be considered for this appraisal?	Training on identification and place in treatment of new therapies is not routinely included in appraisals. This is part of CPD for clinicians, often supported by pharmaceutical companies. Training on handling and collection of biopsies does not apply solely for identification of patients potentially suitable for larotrectinib and therefore the costs of this should not be assigned to a single technology.
Issue 16: Utility values	
How closely do the utility values modelled match the utility values of patients in clinical practice?	Utility values for larotrectinib were informed by EQ-5D-5L and PedsQL estimates taken directly from the patients enrolled in the larotrectinib clinical trial programme. In the comparator arm, health state utilities were applied independently per health state in each comparator engine and the data were sourced where possible from NICE TAs or if no relevant TA could be identified, the literature. Given the process and scrutiny undertaken in each technology appraisal to select the Committee's preferred inputs and assumptions these sources were determined to be most suitable for decision making, and allowed the data and assumptions used in the model to reflect the Committee's preferred assumptions. This minimises uncertainty, and allows incorporation of input from the wide range of stakeholders who contributed to previous appraisals.
Is there justification for considering post-progression utility values to be different between larotrectinib and best supportive care?	The company consider there is justification for maintaining a difference in quality of life post-progression after larotrectinib treatment compared to standard of care and that this is clinically plausible. It would be overly conservative and unreasonable to equalise post-progression utility. The overarching aim during the development of larotrectinib was to create a precision medicine specifically targeting patients with an NTRK gene fusion while minimising off-target toxicity. Whilst data is immature, the company believe the side effect profile we see reinforces this aspiration. Whilst the company acknowledge the small number of post progression data points, the results presented in Appendix N of the company submission show that indeed the mean post-progression utility is sometimes higher compared to pre-progression utility at the same cycle. This difference is not statistically

	significant however since the number of assessments at progression is very low at a given cycle. Further, these descriptive results do not show consistent differences in mean utility values when comparing pre- and post-progression scores. When applying the OLS and MMRM models, on the other hand, the results suggest that there is a significant negative consequence for patient utility associated with disease progression as the post-progression utility estimates were consistently lower than the progression-free state across all models and scenarios. This trend remains when removing the paediatric patients contributing to the post-progression assessments while remaining on treatment. Regardless of the statistical methodology applied to the clinical trial-based utilities, the utilities for pre-progression range between [redacted], while post-progression utilities varies from [redacted]. These ranges of utilities observed for patients treated with larotrectinib are consistently higher in general compared to the literature-based utilities for the pooled comparator of [redacted] and [redacted] for pre and post-progression and therefore suggest a higher utility for patients treated with larotrectinib than has been observed for the comparator treatments. Assuming an equal post-progression utility across all treatment options is unjustified as it is inconsistent with the trend for overall higher utility observed in the larotrectinib trial. Furthermore, it is not unusual to have differences in post progression utilities especially where comparators are older chemotherapy products which generally have poor side effect profiles. Indeed in the last three years NICE has assessed 48 oncology products in the last line setting and there were a number of appraisals that reported significant QALY gains between treatment and comparator in the post-progression state. Whilst we agree the post progression utility data collected in the trial are limited, the difference in utilities identified is completely plausible. There are many examples we could put forward to support this approach but below are a few for committee consideration: **Avoiding amputation or disfiguring surgery** In the SCOUT trial, [redacted] patients were listed as having no other curative options besides amputation or disfiguring surgery. Larotrectinib treatment enabled an increased rate of limb sparing surgery. In all [redacted] patients, amputation was avoided. Disfiguring surgery, such as amputation, can have devastating, lifelong consequences. A study in patients with lower extremity bone sarcoma, explored the difference in quality of life between those patients who had an amputation and those patients who went on to have limb-

	preserving surgery. The overall quality of life of patients with limb preservation was significantly higher than patients with amputation (*p*-value<0.01) (1). **Persistent adverse effects of comparator treatment** Larotrectinib was well tolerated in the clinical studies. Most drug-related AEs were grade 1 or 2, whereas chemotherapy for example, can be associated with significant adverse effects. Lasting effects from the comparator side effects could plausibly lead to the difference in the post progression utilities. Examples include: **a. Persistent symptoms of neuropathy** Cisplatin for example can cause neuropathies which may be irreversible. Neurotoxicity with ifosfamide is reported to persist and occasionally, recovery has been incomplete. A systematic review has reported on the negative impact on quality of life of peripheral neuropathies induced by chemotherapy (2). **b. Nephrotoxicity** Cisplatin for example can cause severe cumulative nephrotoxicity. Renal failure associated with gemcitabine may not be reversible with discontinuation of therapy and dialysis may be required. Disorders of renal function following ifosfamide administration are very common. Glomerular or tubular dysfunction may resolve with time, remain stable, or progress over a period of months or years, even after completion of ifosfamide treatment. A number of studies have reported that people with ESRD experience significantly reduced quality of life relative to those with normal kidney function (3). **c. Cardiotoxicity** Many cancer therapies are potentially cardiotoxic and cardiotoxicity adversely affects prognosis in cancer patients (4). For example, cardiotoxicity is a risk of anthracycline treatment that may be manifested by early (i.e. acute) or late (i.e. delayed) events. Delayed cardiomyopathy is manifested by reduced left ventricular ejection fraction (LVEF) and/or signs and symptoms of congestive heart failure (CHF). In CHF, quality of life is reported to decrease as New York Heart Association (NYHA) functional class worsens (5). As another example of persistence of effect, radiation-induced pericardial effusion has been reported as late as 15 years following radiotherapy (4).

	d. Infertility Fertility can be impaired in men and women after chemotherapy and radiotherapy. A recent systematic review found that fertility-related psychological distress persists from diagnosis through to survivorship, with cancer patients reporting a range of negative emotional experiences brought about by threatened infertility. In survivorship, reproductive concerns, unfulfilled desire for a child, nulliparous status, and early menopause were linked to higher rates of mental health disorders and psychological distress (6). Larotrectinib was well tolerated in patients with NTRK fusion-positive cancer from a pooled safety analysis across the three clinical studies. Most drug-related AEs were grade 1 or 2. Long-term follow-up of patients with > 2 years exposure as at 30 July 2018 has not indicated new or cumulative toxicities. The overarching aim during the development of larotrectinib was to create a precision medicine specifically targeting patients with an NTRK gene fusion while minimizing off-target toxicity. Patients who progress on larotrectinib would have higher quality of life because the treatment is very well tolerated, with extremely rare cases of grade 3 and 4 side effects. The incidence of grade 1 and 2 side effects is low as well, compared to treatments such as chemotherapy. Additionally, the type of side effects potentially occurring while taking larotrectinib—fatigue and weight gain—have a less negative/serious impact on patient's quality of life than with chemotherapy (e.g. peripheral neuropathy or cardiotoxicity) which have the potential of long term or irreversible damage. Although larotrectinib's long term data is not yet mature, the extremely rare occurrence of grade 3 and 4 side effects suggest that assuming different utility weight between larotrectinib and comparator treatments for post-progression is a highly plausible approach. 1. Mason et al. Quality of life following amputation or limb preservation in patients with lower extremity bone sarcoma.Frontiers in Oncology. 2013. August 2013; Volume 3; Article 210. 2. Mols et al. Chemotherapy-induced peripheral neuropathy and its association with quality of life: a systematic review. 2014. Support Care Cancer; 22:2261–2269 3. NHS England. Chronic Kidney Disease in England: The Human and Financial Cost.2012] 4. Koutsoukis et al. Cardio-oncology: A Focus on Cardiotoxicity. 2018. European Cardiology Review 2018;13(1):64–9. 5. Juenger et al. Health related quality of life in patients with congestive heart failure: comparison with other chronic diseases and relation to functional variables. 2002. Heart 2002;87:235–241. 6. Logan et al. Systematic review of fertility-related psychological distress in cancer patients: Informing on an improved model of care. 2019. Psycho-Oncology. 2019;28:22–30.

Issue 17: End of Life	
What is the life expectancy of the patient group receiving established management?	The indication concerns a disease setting of locally advanced or metastatic malignant solid tumours after standard therapy or when there is no appropriate available therapy. According to the draft EPAR, in this setting symptoms of disease will be present or imminent and the disease is incurable, likely leading to death. In preparing for the submission, the company conducted an extensive comparator therapy SLR on a multitude of tumours known to harbour NTRK gene fusions, and this indicated a limited life expectancy with 'standard of care' treatments in patients who have received ≥ 1 prior therapy. With available treatments median PFS and OS varies across tumour types in patients with progressive, recurrent or metastatic disease. Median PFS was generally less than 12 months across included tumour types, considerably lower than that of larotrectinib (median PFS [redacted] months). On the basis that patients will be eligible for larotrectinib only if there are no other available satisfactory treatment options, and hence, as a subsequent line of therapy to those identified in the SLR, results would suggest a likely life expectancy for larotrectinib-eligible patients to be within the 24 months NICE criterion. Further detail is presented in Table 30 of the submission.
What is the extension to life of the patient group receiving larotrectinib?	In terms of extension to life, the survival data, although immature and analysis ongoing, supports durability of larotrectinib effect and extension of life of greater than the 3 months specified by NICE. Larotrectinib represents a step-change in the management of patients with refractory locally advanced or metastatic NTRK fusion-positive solid tumours in that it is a treatment option for patients who have exhausted all other satisfactory treatment options. Larotrectinib should therefore be considered as an end-of-life therapy. Further to the histology independent nature of larotrectinib, it is not appropriate to consider applying the end of life criteria to specific cancer sites based on their location. As highlighted in the draft EPAR, the patient population are those with symptoms of disease present or imminent and the disease is incurable, likely leading to death.
Issue 18: Innovation	

Is larotrectinib an innovative treatment?	Larotrectinib is considered innovative and a 'step change' in the management of NTRK fusion-positive cancer. 1. **Larotrectinib provides a specific treatment for NTRK fusion-positive solid tumours where previously no treatment was available** - Larotrectinib represents a paradigm shift in the way cancer is treated, enabling cancer treatment to be delivered according to causation (in this case, the presence of NTRK gene fusion) as opposed to tumour location as has been done traditionally. **Larotrectinib is the first histology independent therapy approved in Europe.** 2. **Innovative design to selectively target NTRK fusion cancer: a precision medicine** – Larotrectinib is a first-in-class, orally bioavailable, potent and highly selective inhibitor of TRKA, TRKB, and TRKC, rationally designed to avoid activity with off-target kinases. 3. **Treatment of adults and children within one indication -** Larotrectinib has been shown to be generally safe and effective across a broad range of tumours including rare tumours and rare subsets of more common tumours, and in paediatric and adult patients ranging in age from [redacted] years. 4. **A step towards delivering 'Personalised medicine' in cancer patients** - Personalised medicine is based on comprehensive genomic and diagnostic characterisation, meaning different subtypes of patients within a given condition can be identified, and treatment can be tailored to the underlying cause. The availability of larotrectinib enables delivery of personalised medicine to cancer patients harbouring NTRK gene fusions. **Unmeasured benefit** - In terms of unmeasured benefit, the value that an oral oncology medication brings for treating paediatric patients with advanced cancer, in terms of impact on schooling and the further impact on parents should not be underestimated. This compares favourably with treatment regimens requiring daily visits to the hospital as well as admissions to manage adverse events and was highlighted in the clinical validation interviews. These benefits were not captured in the economic model.

	Larotrectinib treatment enabled an increased rate of limb sparing surgery. Disfiguring surgery, such as amputation, can have significant, devastating, lifelong consequences. The lifelong impact of avoiding amputation or disfiguring surgery has not been captured in the QALY calculation.
Issue 19: Cancer Drugs Fund	
Does larotrectinib meet the criteria for inclusion in the Cancer Drugs Fund?	With the approved patient access scheme, the company has presented a base case ICER in line with meeting the end of life criteria. Whilst we acknowledge the data is immature and there are uncertainties, the Committee is asked to give balanced consideration to downward as well as upward uncertainty that is associated with evaluating this histology independent innovation. Further, that a recommendation to enter the CDF will go towards addressing much of the uncertainty without denying patients an effective treatment in a timely manner. In order for larotrectinib to be made available by the CDF, it is the company's understanding that the ICER needs to be plausible rather than having a high degree of certainty attached. ***Given the current level of uncertainty, the company proposes that whilst data mature, larotrectinib is made available in a timely manner through the Cancer Drugs Fund.***

What data would be most useful to collect to address the outstanding uncertainties? For example, unrepresented tumour types.	The company consider that collecting longer term data on the following could be informative:
	• Larger sample size
	• Which populations would use larotrectinib in clinical practice
	  ○ Tumour site
	  ○ Place in therapy
	  ○ Optimal timing of testing
	• OS and extension to life
	• Post progression treatments
	• Impact of avoiding amputation/ disfiguring surgery

Larotrectinib for treating NTRK fusion-positive advanced solid tumours [ID1299]

Additional evidence further to technical engagement

25th October 2019

In line with our email to Linda Landells on 18th October, Bayer are submitting further evidence to NICE as a result of the recent discussions in the technical engagement process.

The areas of additional evidence are as follows:

1. To further explore the potential impact of experimental post-progression treatments on the efficacy data
2. Further clarification on the point raised on the teleconference regarding the potential for dose escalation post progression.
3. Analysis of the updated data cut from the trial programme (ESMO 2019). Please note that the original ePAS2 results we used in our submission were based on independent central review (IRC), whereas the new dataset used investigator assessment (INV), limiting comparability.
4. Updated epidemiological data – we have now conducted a systematic literature review and meta-analysis of real world evidence on the epidemiology of NTRK gene fusion in solid tumours and present updated estimates of potentially eligible patients.

1. Potential impact of experimental post-progression treatments on the efficacy data

In section 6.3.4 of the ERG report, the technical team commented that the survival gain with larotrectinib post-progression exceeded that during the progression-free state. The ERG suggested that this may be driven by the high uncertainty in extrapolation for larotrectinib, but is also likely related to the high proportion of patients that received LOXO-195 post-progression ([redacted] out of [redacted] progressed patients) as well as larotrectinib treatment post-progression ([redacted] out of [redacted] progressed patients).

Additional individual patient data was reviewed to identify patients receiving LOXO-195 and larotrectinib post-progression. With this additional investigation, Bayer identified that the previously reported numbers of patients receiving post progression treatments was incorrect; apologies for introducing this confusion.

- Among the [redacted] patients previously reported with post-progression larotrectinib treatment, [redacted] should be recategorized:
  - [redacted] received larotrectinib only on the day of progression (and not after); and
  - [redacted] received larotrectinib after surgery and not after progression.
  - [redacted] had not progressed based on IRC assessment at the time of the data cutoff;
- Among the [redacted] patients with reported LOXO-195 use, [redacted] patients had progression at the time of the data-cutoff. Other reported use of LOXO-195 occurred after the data-cutoff for the NICE submission (i.e. July 2018).

Therefore, when evaluating post-progression larotrectinib and LOXO-195 use in the [redacted] patients with IRC progression at the time of the NICE submission (i.e. July 2018), only [redacted] patients received larotrectinib and [redacted] patients received LOXO-195 in the post-progression period. Amongst the [redacted] patients who used larotrectinib post-progression, [redacted] also used LOXO-195, resulting in a total of [redacted] patients ([redacted]% of the [redacted] progressors) who received either LOXO-195 or larotrectinib post-progression.

Survival exploration excluding patients receiving investigational treatments

To investigate the impact of LOXO-195 and larotrectinib use post-progression, overall survival was reanalyzed after excluding patients receiving these treatments. Several concepts were explored through comparison of Kaplan Meier (KM) plots, including: omitting patients who used larotrectinib post-progression, patients who used LOXO-195 post-progression and patients who used both larotrectinib and LOXO-195 post-progression.

OS for the total population and those patients receiving LOXO-195 and larotrectinib post-progression

First, OS trends for those patients receiving LOXO-195 and larotrectinib post progression were compared to the full patient sample for:

1) patients who used larotrectinib post-progression (green);

2) patients use used LOXO-195 post-progression (teal); and

3) patients who used either larotrectinib or LOXO-195 post-progression (purple).

Per the KM plots below, we observe that plots for patients using larotrectinib and LOXO-195 are [redacted] than all patients (red) initially but that their survival [redacted]. However, please note that trends in the tails of the KM curves are based on a very limited number of patients given censoring, and should be interpreted with caution.

OS when excluding LOXO-195 and larotrectinib use post-progression

To test the influence of these patients on observed model trends and resulting ICERs, KM plots were generated to show trends when patients who used larotrectinib and/or LOXO-195 were excluded from the sample. The purple plot in the figure below shows OS trends for patients who progressed but did not receive either LOXO-195 or larotrectinib post-progression. After removing patients with LOXO-195 or larotrectinib treatment post-progression (purple line), OS trends were [redacted] than the KM plot for all patients (red).

As noted in the technical report, the ERG believed that the post-progression OS gain was likely due to use of LOXO-195 and post-progression larotrectinib. When we compare OS for those with versus without LOXO-195, the lOXO-195 patients [redacted]. When we compare OS for those

with versus without post-progression larotrectinib use, the post progression larotrectinib patients [redacted]. Accordingly, there is no evidence to back up the ERG's belief on the impact of these post-progression treatments on OS trends. Specifically, based on this exploration, it does not appear that time in the progressed disease health state [redacted]. Though data immaturity remains given length of follow-up and rolling enrollment in the trial, we suggest that this uncertainty be addressed through extended data collection in the CDF.

In conclusion, given that the, albeit limited data, does not support OS benefit being driven by use of larotrectinib post-progression or LOXO-195, and that no other appraisal has taken into consideration potential benefit from experimental treatments, (see Bayer response to the technical engagement report 'Issue 10 Subsequent therapies': *Should experimental treatments be adjusted for in this analysis?,* pages 18-19), we would suggest equalizing post-progression OS benefit to match that of the comparators is overly conservative and should be reconsidered.

Other post-treatment considerations – appropriateness of the cure model and impact of Infantile fibrosarcoma (IFS) surgery

Children with advanced IFS may require limb amputation to achieve a cure.

In assessing cost-effectiveness of oncology products, the survival data from the clinical trials with limited duration of follow-up requires extrapolation to project life-time outcomes. However, the accuracy of extrapolation is highly-dependent on selection of the survival models. With the existence of heterogeneity between patients and long-term survivors, standard models do not capture the patients who might be 'cured' and will not experience cancer-related events. A cure model has been applied in multiple previous NICE TAs and published studies.

While our base case was conservative and did not consider curative effect, we have explored the cure model using larotrectinib OS. The results from the cure model show that long-term survival was improved when incorporating a statistical cure. Applying a cure model approach to this data may improve the ICER; however, cure cannot be determined based on the individual patient data (IPD) due to short follow-up, limited sample size and high censoring of data (not clustered in the plateau).

The impact of larotrectinib on the costs and benefits of increased surgery eligibility for paediatric patients were not considered in the company base case model since the tumour type (IFS) was not directly modeled and was grouped under STSp instead. Had IFS been modeled separately to reflect the health-related quality of life benefits following limb sparing surgery, the cost effectiveness of larotrectinib may be more favourable.

2. Further clarification on the point raised on the teleconference regarding the potential for dose escalation post progression

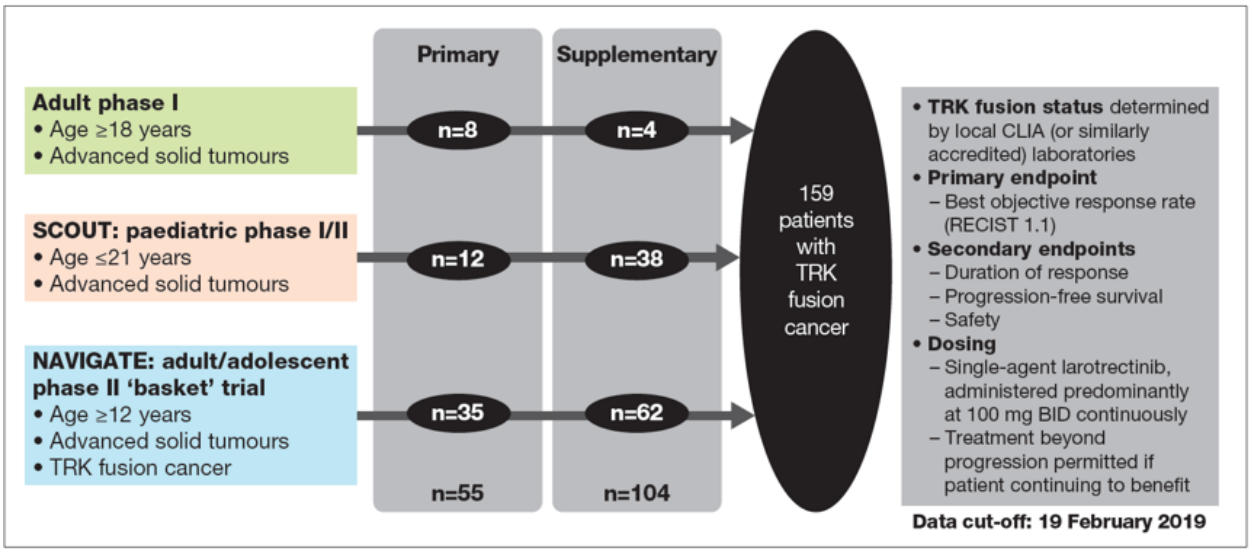
Inadvertently, Bayer created confusion on the technical engagement teleconference on 8th October regarding larotrectinib received by patients post progression. We can confirm that the average dose used in the original economic model submitted to NICE included post progression exposure and any dose escalation. However, we expand below on the post-progression exposure to larotrectinib.

Additional individual patient data were reviewed to further investigate those patients receiving larotrectinib post-progression. With this additional investigation, Bayer can confirm that there were [redacted] patients who received post-progression larotrectinib (see topic 1, page 2). Of these there were [redacted] patients who experienced a dose escalation of larotrectinib. The table below presents a summary of this information. [redacted] of the [redacted] were paediatric patients, and the dose increased due to growth/change of body surface area (BSA). The other was an adult who received [redacted]. All [redacted] cases of dose escalation post-progression were included in the original economic model submitted to NICE.

subjid	Dose increase?	Dose escalated to	Reason for dose escalation
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

3. Analysis of the updated data cut from the trial programme

Updated information on the larotrectinib clinical trial programme was presented at ESMO in October 2019 (1). The data cut-off for this integrated expanded dataset was 19 February 2019. The integrated dataset included 55 patients from the primary dataset and 104 patients in the supplementary dataset. Of the 159 patients included in the integrated dataset, 153 were evaluable for efficacy. Six patients were not evaluable due to post-baseline assessments being incomplete.



BID, twice daily; CLIA, Clinical Laboratory Improvement Amendments; RECIST, Response Evaluation Criteria In Solid Tumors; TRK, tropomyosin receptor kinase.

The expanded data cut includes a total of 153 patients (PAS+SAS1), where the primary endpoint was best objective response rate based on investigator assessment. Secondary endpoints included duration of response, progression-free survival, overall survival and safety.

It is of note that in this analysis set, efficacy was judged based on investigator assessment as opposed to independent review committee assessment, as per the dataset on which the original submission was based. As such, this leads to limitations in the comparability of the data, but Bayer have provided these data as this is the most up-to-date data available at this time. It should also be noted that the ESMO dataset does not include patients with primary CNS tumours.

A summary of the efficacy assessments are presented in the table below.

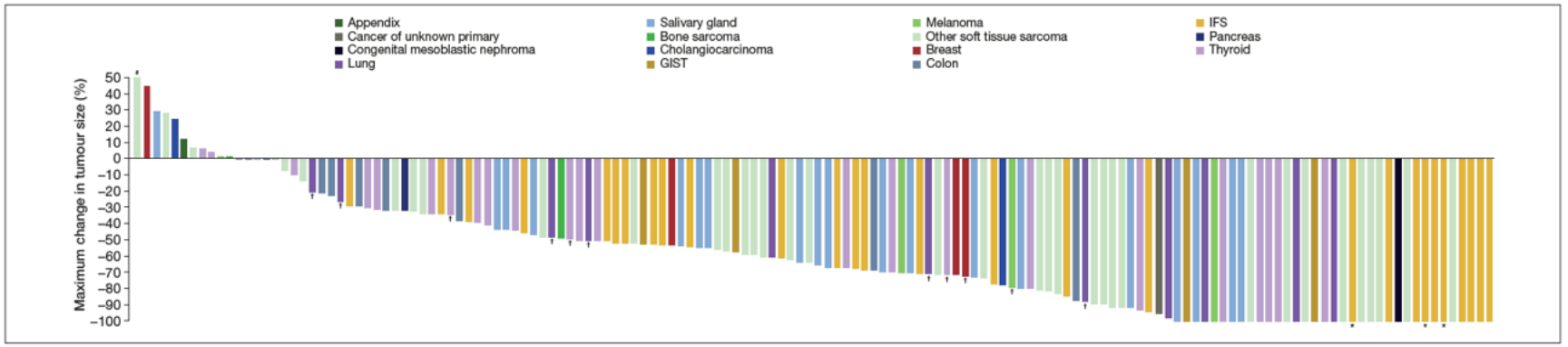
	Integrated dataset (N=159)
Response	
Evaluable patients, n	153[a]
ORR (95% CI)	79% (72–85)
Best overall response, n (%)	
Complete response	24 (16)[b]
Partial response	97 (63)[c]
Stable disease	19 (12)
Progressive disease	9 (6)
Not determined	4 (3)
Duration of response	
Median, months (95% CI)[d]	35.2 (22.8–NE)
Range, months	1.6+ to 44.2+
Rate of ongoing response at 12 months, % (95% CI)[e]	80%
Median follow-up, months	12.9
Progression-free survival	
Median, months (95% CI)	28.3 (22.1–NE)
PFS rate at 12 months, % (95% CI)[e]	67 (58–76)
Median follow-up, months	11.1
Overall survival	
Median, months (95% CI)	44.4 (36.5–NE)
OS rate at 12 months, % (95% CI)[e]	88 (83–94)
Median follow-up, months	13.9

[a]Six patients not evaluable because post-baseline assessments were not yet done at data cut-off. Best response percentages are calculated from the evaluable patient population. [b]Including three patients with pathological complete response; two patients had complete responses pending confirmation. [c]13 partial responses pending confirmation. [d]In patients with confirmed responses (n=108). [e]Kaplan–Meier estimates. CI, confidence interval; DOR, duration of response; NE, not estimable; ORR, objective response rate; OS, overall survival; PFS, progression-free survival.

In the integrated dataset, larotrectinib was efficacious regardless of tumour type (note, two further tumour locations were included (hepatic and prostate cancer) in the dataset). Please see figure on the following page which presents the maximum change in tumour size.

Further, in the expanded safety population of 260 patients, and with longer follow-up than in the initial report, no new safety signals of larotrectinib were identified.

Maximum change in tumour size



Excludes four patients who had clinical deterioration prior to an initial response assessment and six patients who were not evaluable due to insufficient time on therapy. *Patients with a pathological complete response. #Maximum change in tumour size of +93.2%. †Patients with brain metastases. GIST, gastrointestinal stromal tumour; IFS, infantile fibrosarcoma.

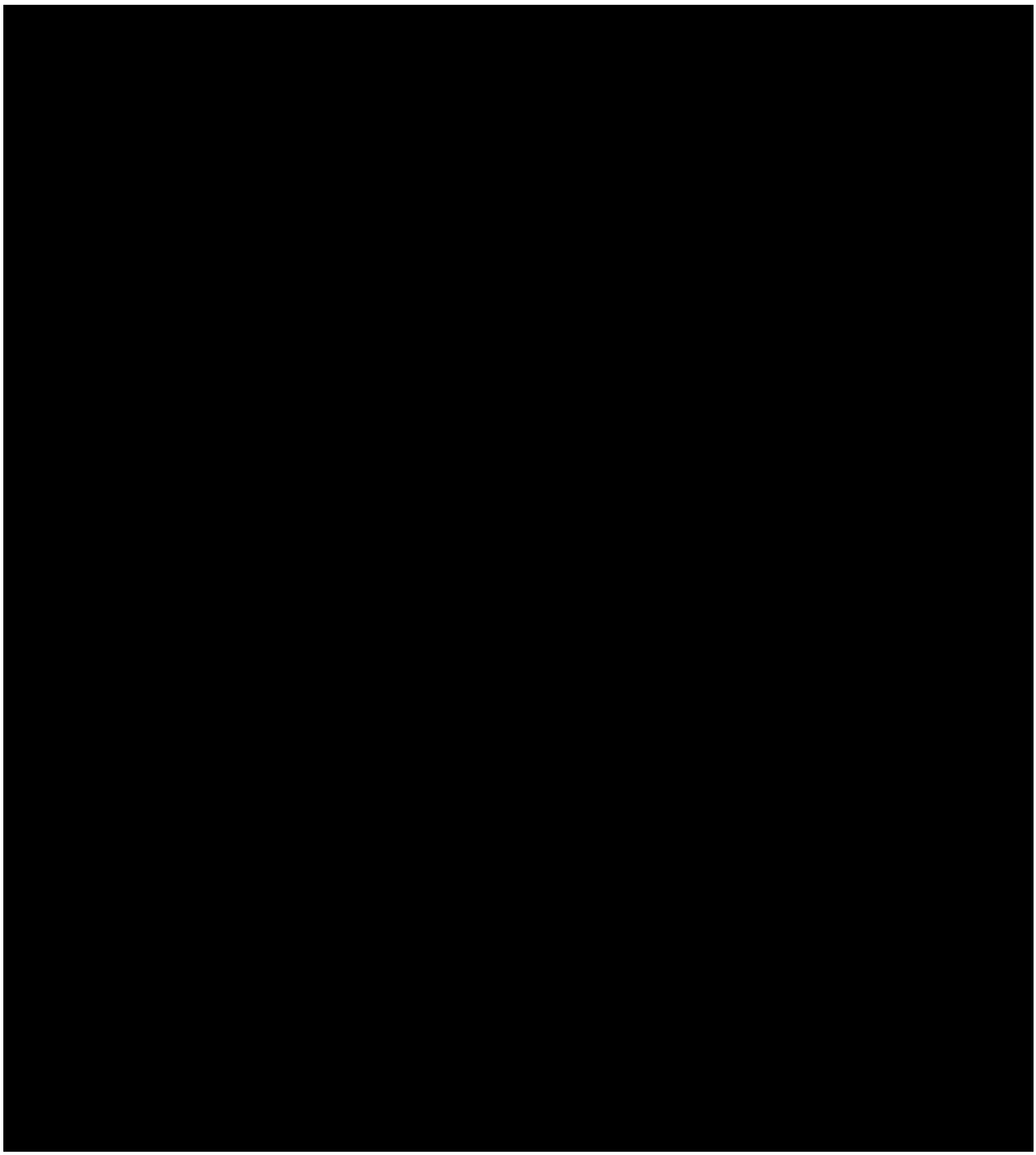
With the additional patients and longer follow-up in the Feb 2019 data cut, survival analyses were conducted for different patient sets to understand the overall impact. Overall survival (OS) and progression-free survival (PFS) Kaplan Meier curves were generated by response status (responders, non-responders) and for the pooled population for the full sample of 153 patients (PAS+SAS1).

When reviewing the following graphs, please note the limitations of Kaplan Meier (KM) curve comparisons. Specifically, KM curves estimate survival probability using product limit, where this involves computing the probability of occurrence of event among patients at risk. However, when the majority of patients are censored before the last event time, it is not uncommon to see an abrupt drop in the tail. This trend may not be indicative of a sudden shift in performance but can be caused by the nature of data, as the estimated survival probability could have been less biased if there were more patients at risk when the very last survival probability is calculated. Therefore, the small patient numbers at the tails of the following KM curves should be interpreted with caution given the small sample sizes from censoring.

PFS for total population

Kaplan Meier plots for progression-free survival for the PAS+SAS1 patient set (n=153) and the original dataset (ePAS2+SAS3, n=102) are generally similar. Cumulative number of events increased from [redacted] events in the original dataset to [redacted] events in the updated dataset. When using the expanded data, [redacted] when compared to the ePAS2+SAS3 patient set (n=102). However, please note that trends in the tails of the KM curves are based on a very limited number of patients given censoring, and should be interpreted with caution. After displaying the KM curves, [redacted] projections are shown at 40 months and 400 months, showing [redacted] for PFS with the expanded dataset.

PFS for total population comparing different data cuts (green=153, blue=102)

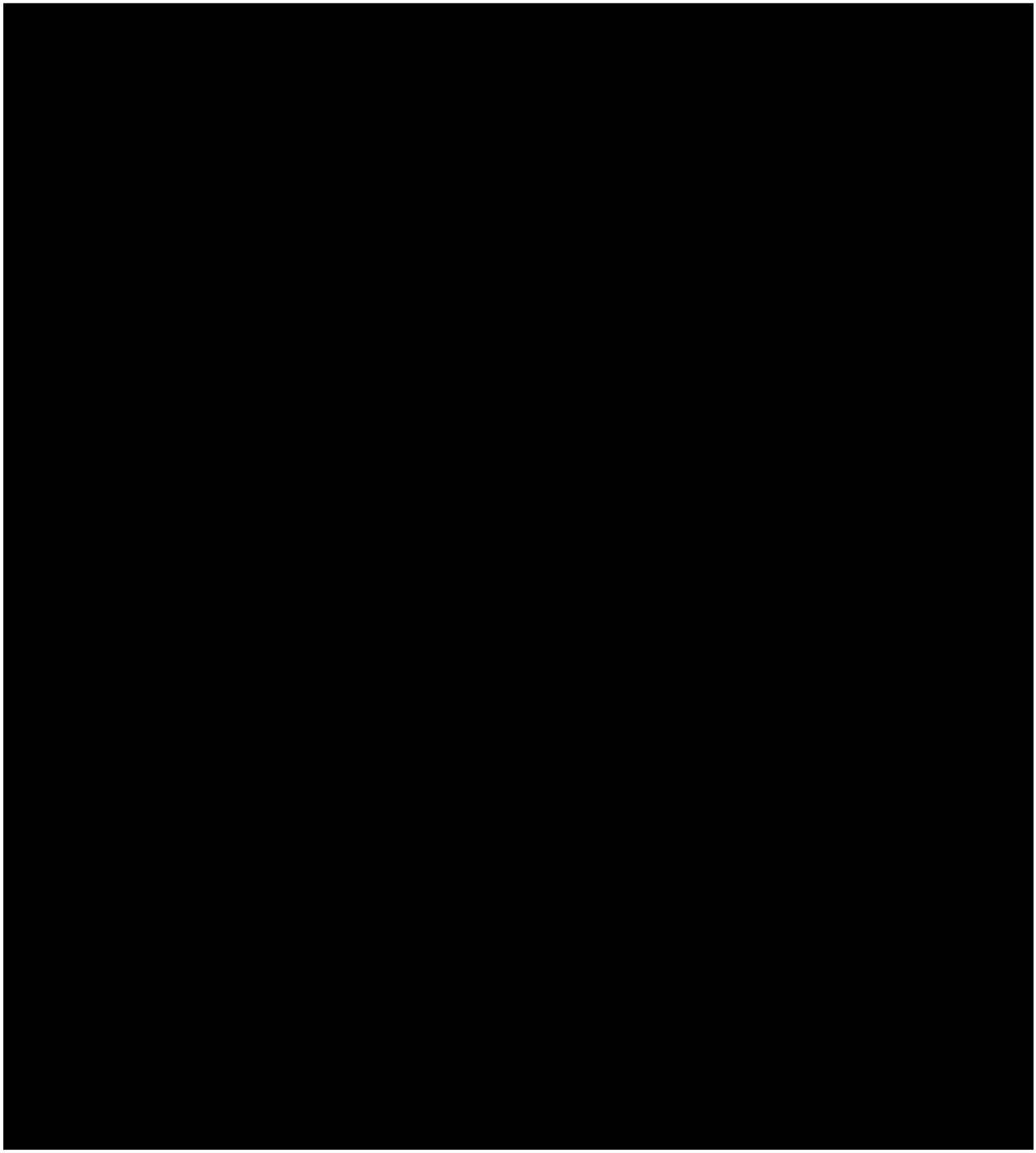


[redacted] **projections for PFS across datasets (green=153, blue=102) at 40 and 400 months**

OS for total population

Similar to PFS, we see [redacted] of the KM curve for the expanded data [PAS+SAS1 patient set (n=153)] as compared to the initial ePAS2+SAS3 patient set (n=102) using the July 2018 data cut. Cumulative number of events increased from [redacted] events in the original dataset to [redacted] events in the updated dataset. However, please note that trends in the tails of the KM curves are based on a very limited number of patients given censoring, and should be interpreted with caution. After displaying the KM curves, [redacted]_projections are shown at 40 months and 400 months, showing [redacted] for OS with the expanded dataset.

OS for total population comparing different data cuts (green=153, blue=102)



[redacted] **projections for PFS across datasets (green=153, blue=102) at 40 and 400 months**

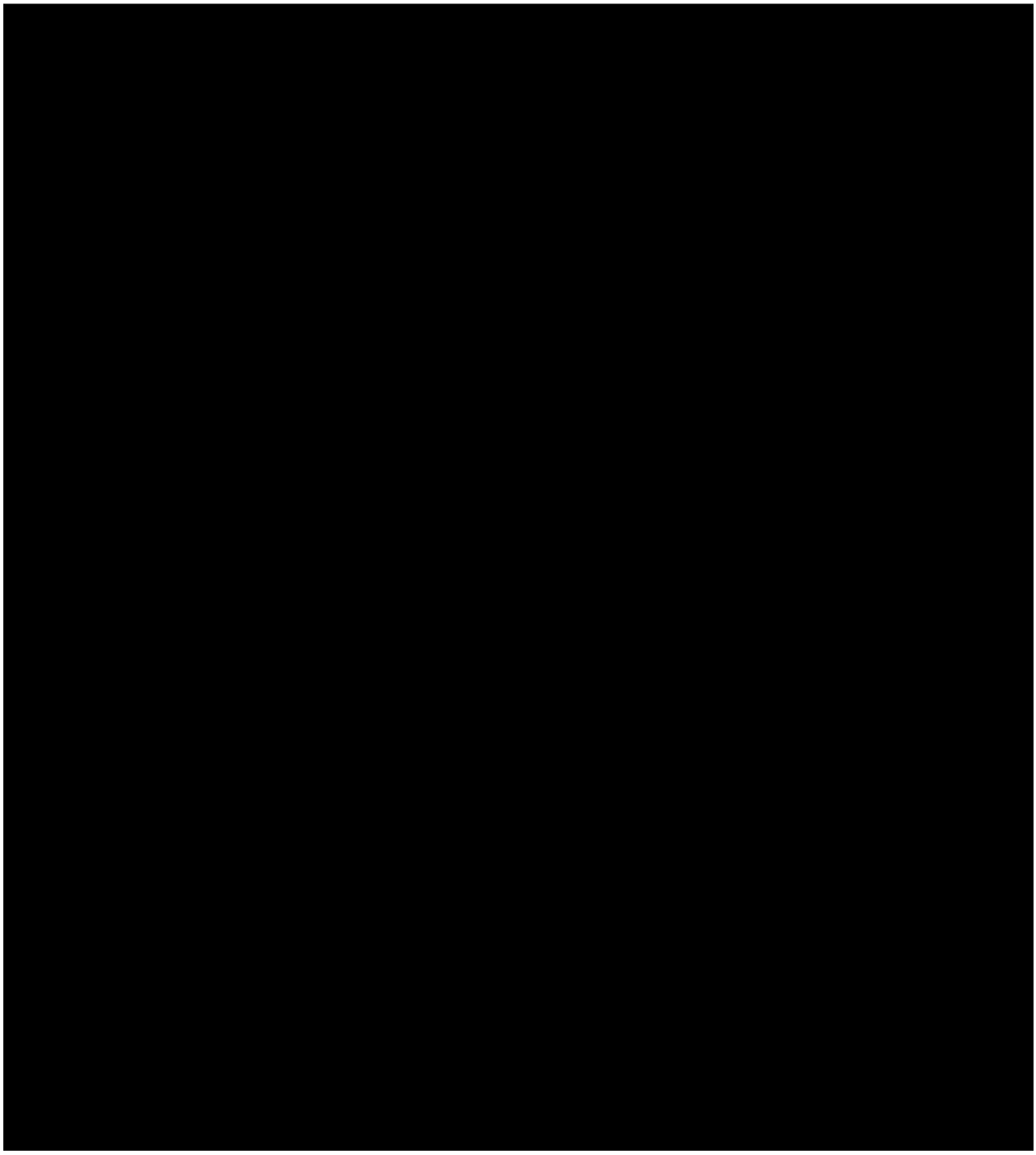
Progression Free Survival by response status

Progression-free survival was estimated by response status for the expanded PAS+SAS1 patient set (n=153). Relative to the initial dataset, Kaplan-Meier curves for responders [redacted] as compared to non-responders in the expanded data. While PFS KM curves for responders are very similar across datasets, a [redacted] is seen for non-responders with the expanded dataset (n=153). A log-rank test was applied to both samples (n=102, n=153) to test whether differences between responders and non-responders was statistically significant. For both datasets, p values indicate [redacted] in PFS for responders and non-responders. Please note that trends in the tails of the KM curves are based on a very limited number of patients given censoring, and should be interpreted with caution.

PFS by response status using July 2018 cut (n=102)

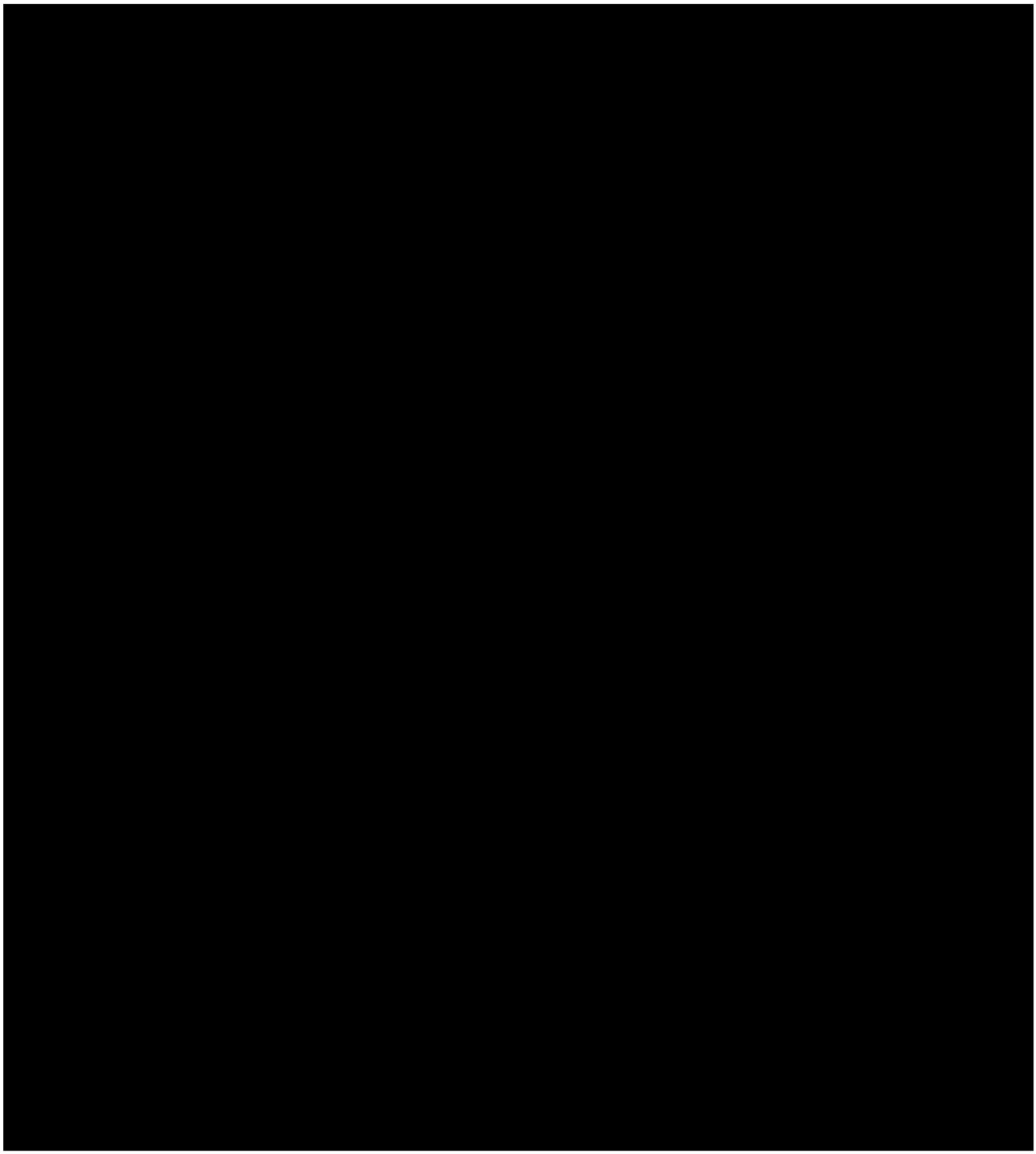
PFS by response status using Feb 2019 cut (n=153)

PFS for <u>responders</u> comparing different data cuts (green=153, blue=102)



PFS for <u>non- responders</u> comparing different data cuts (green=153, blue=102)

When comparing across datasets, a reduction in the number of non-responders is observed. However, this may be due to the difference in patient datasets.



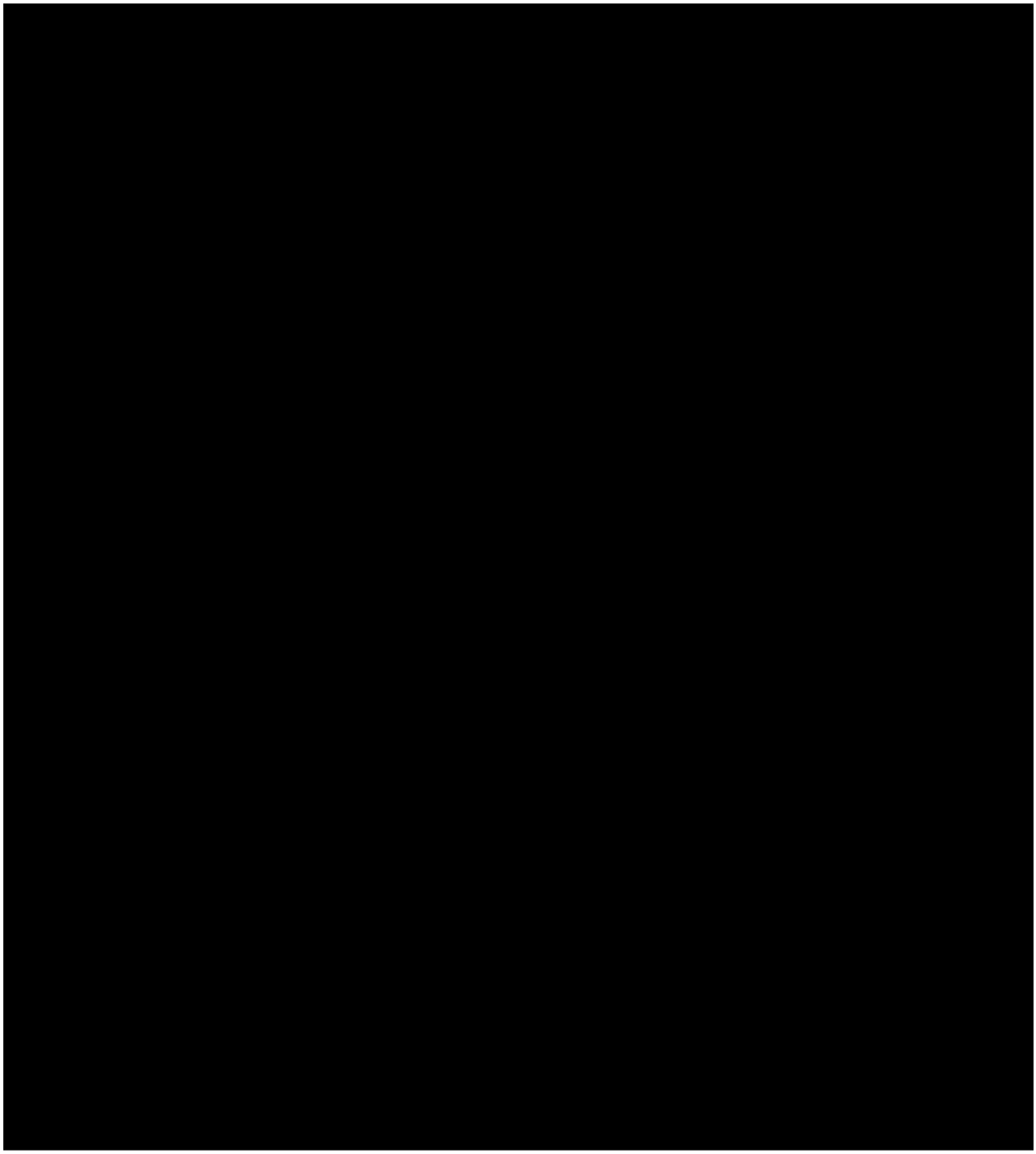
Overall Survival by response status

Overall survival was estimated by response status for the expanded PAS+SAS1 patient set (n=153). Kaplan-Meier curves for responders [redacted]. Similar to the trend seen in PFS, the KM curve for non-responders for the expanded dataset shows [redacted] when comparing to non-responder KM trends from the original data (n=102). A log-rank test was applied to both samples (n=102, n=153) to test whether differences between responders and non-responders was statistically significant. For both datasets, p values indicate [redacted] in OS for responders and non-responders. Please note that trends in the tails of the KM curves are based on a very limited number of patients given censoring, and should be interpreted with caution.

OS by response status using July 2018 cut (n=102)

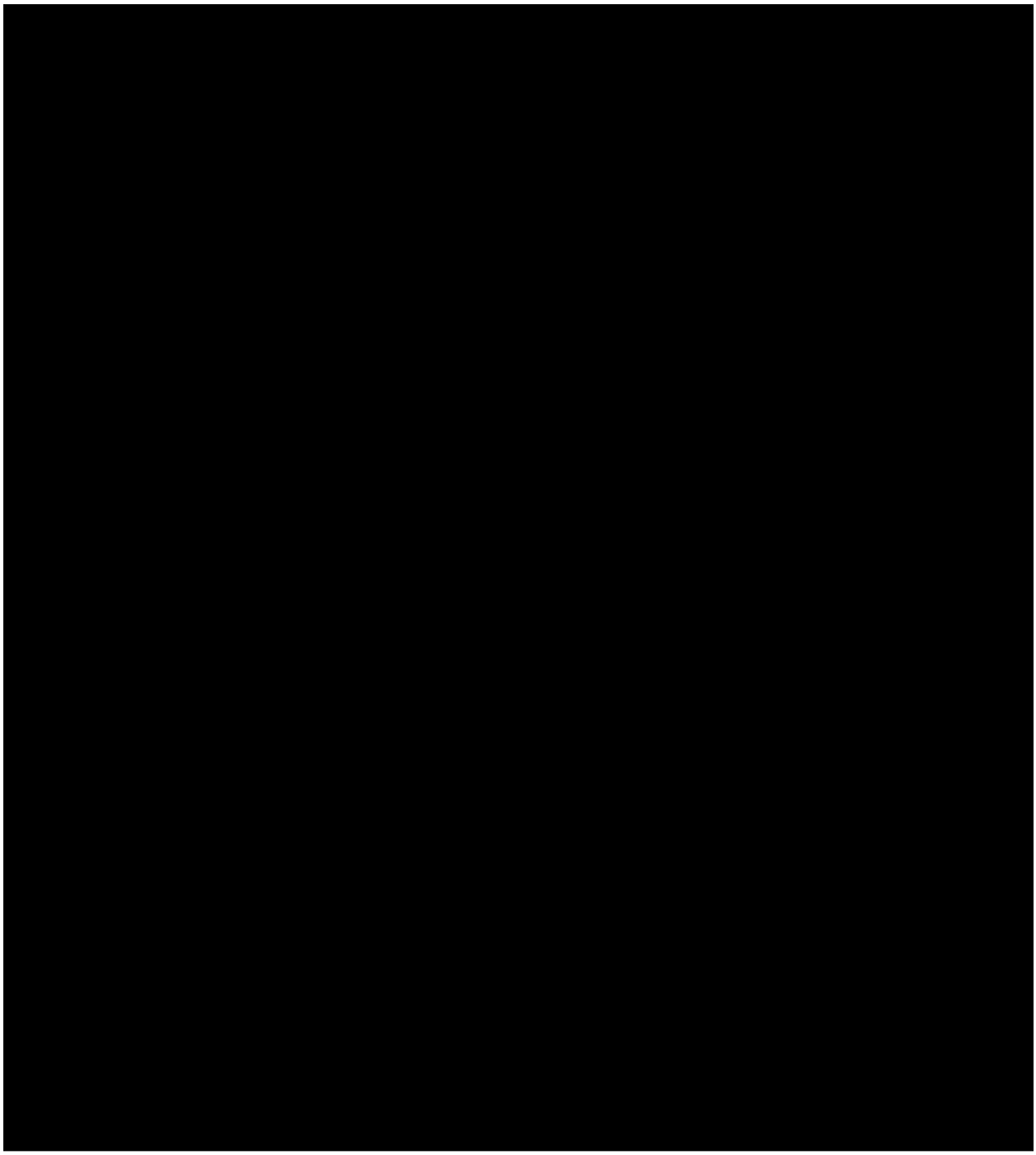
OS by response status using Feb 2019 cut (n=153)

OS for <u>responders</u> comparing different data cuts (green=153, blue=102)



OS for <u>non-responders</u> comparing different data cuts (green=153, blue=102)

When comparing across datasets, [redacted] in the number of non-responders is observed. However, this may be due to the difference in patient datasets.



The availability of an expanded dataset allows us to make additional observations about the effectiveness of larotrectinib in NTRK fusion cancers. The very impressive response rate of 72% in the initial cohort has been exceeded with the response rate in the expanded dataset of 79% (CI: 72-85%). Accordingly, there is no evidence that adding more patients and new tumour types has reduced the response rate. We draw attention again to the fact that for this analysis set, efficacy was judged based on investigator assessment as opposed to independent review committee assessment, as per the dataset on which the submission was based. As such, this leads to limitations in the comparability of the data.

Parametric fits to PFS now predict [redacted] PFS at future timepoints than was the case with the original dataset. PFS for responders is generally [redacted] than in the original dataset although median has now been reached. PFS for non-responders is generally [redacted] with [redacted] median PFS in the expanded dataset. The expanded PFS data confirm and [redacted] our estimates of PFS, with additional robustness as these estimates are based on larger sample size and longer follow up. The expanded dataset confirms (investigator assessed) response as a strong predictor of benefit.

In the extended dataset OS to around [redacted] months is very close to the original data. OS beyond [redacted] months however looks quite different in the expanded dataset, with a [redacted] in survival after month 30. Inspection of the OS curves by response status and the numbers at risk suggests these deaths occurred in patients who had experienced response in the original dataset.

The expanded dataset has confirmed low risk of early mortality, as observed in the original data. However it appears that [redacted]. This observation is based on small numbers (<20 at risk at month 30), and other patients continue to experience extended survival. Parametric estimates of OS based on the expanded cohort will be [redacted], but additional follow up is needed to better understand the pattern of long term survival in this population.

Model results with expanded dataset

The ERG amended model was updated to include the latest ESMO PAS+SAS1 (n=153) data cut and to replicate all of the scenarios included within Table 8 of the ERG technical report. It should be noted that two of the scenarios included within the original table were not in line with the description in the ERG report (namely scenarios 3 and 4 in the table below). Specifically, scenarios 3 and 4 use the 72% response rate from the trial data rather than ERG report preferred [redacted]%. Applying the 72% response rate seems a more reasonable approach especially considering that the ORR in the updated dataset is 79%. The description in the ERG report claims to have used [redacted]% response rate in its base case model results. However, only when applying a rate of 72% was it possible to replicate the ERG model results. Therefore, 72% was similarly applied in scenarios 3 and 4 using the company model to allow comparison of results with the ERG model results. Similarly, for scenario 3, we could only replicate the ERG results if the same utility was applied for larotrectinib and the comparators in the post-progression health state, in addition to equalizing post-progression survival.This brief summary of changing trends across data sets corrects these discrepancies to ensure transparency in the scenarios explored, as outlined in the footnotes of the scenario table below.

When reviewing these exploratory analyses, please note that these results are based on replacing efficacy data (PFS, OS) with the expanded data cut. However, the model was not updated to include edits to reflect expanded tumour locations seen in the 153 patient data set. Further, a full update to nuanced assumptions for all inputs was not feasible in the short timeframe for the response. Therefore, this should be viewed as a close proxy to how final results would look if the model was fully updated to reflect the expanded ESMO data. Further, when comparing results, please note that the original data was based on outcomes measured by independent review committee (IRC) but the most recent available data (presented at ESMO) reports outcomes as defined by the trial investigator, which may also contribute to small differences in trends.

The inclusion of additional patients, bringing the total cohort to N=153, suggests [redacted] projections for progressed-disease life years ([redacted]), [redacted] projections of progression-free life years ([redacted]) and [redacted]_treatment cost from [redacted] [redacted]), resulting in an ICER of £[redacted] for the company base-case submission (see Table below). However, results incorporating the additional data should be noted with some caution. Whilst there is added follow-up for the original n=102 patients, the additional patients (n=51) only have a short follow-up with outcomes based on investigators assessment which contributes only for the initial period within the KM.

Comparison of n=102 and n=153 company submitted base case detailed results

		Results with Initial Data (n=102)			Results with Expanded Data (n=153)		Change in incremental differences (n=102 - n=153)
		Comparators (n=102 and n=153)	Larotrectinib (n=102)	Incremental (n=102)	Larotrectinib (n=153)	Incremental (n=153)	
Life years	**Progression-free**	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Progressed disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Total LYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
QALYs	**Progression-free**	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Progressed disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Treatment specific	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Adverse events	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Costs	**Progression free survival**	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Progressed disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Death	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Adverse event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Societal cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Treatment cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Total costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ICER (per QALY)		[redacted]		[redacted]	[redacted]	[redacted]	[redacted]

Using the updated PFS and OS data to replicate the ERG base-case, we found [redacted] ICER with the updated data, which is in line with the company submitted base case change. Most scenarios resulted in consistent relative changes in the ICER from the ERG base case; inclusion of SAS3 dataset (for the BHM ORR only), inclusion of testing costs and inclusion of matched post-progression utility. The ERG response-based model scenarios matching larotrectinib post-progression survival outcomes to the non-responder-based comparator also continued to result in [redacted] ICER than the ERG base case. However, the relative difference between the ERG base case and scenarios are smaller and the overall ICER when combining all ERG scenarios is lower than with the original data, suggesting reduced uncertainty with the inclusion of updated data.

ICER comparison under the ERG response-based model structure of submitted n=102 data cut and updated n=153 data cut

	Current ICERs from tech report	N=153 (ESMO data cut)
Company base case	[redacted]	[redacted]
Scenario 1: ERG base case	[redacted]	[redacted]
Scenario 2: ERG base case + SAS3 dataset	[redacted]	[redacted]
Scenario 3: ERG base case + post-progression survival equal for larotrectinib and comparator	[redacted]	[redacted]
Scenario 4: ERG base case + same post-progression utility for larotrectinib	[redacted]	[redacted]
Drug wastage and adherence, administration costs and resource use	[redacted]	[redacted]
Scenario 5: ERG base case + diagnostic testing costs	[redacted]	[redacted]
Scenario 6: Cumulative impact of the technical team's preferred assumptions on the cost-effectiveness estimate – no diagnostic testing costs	[redacted]	[redacted]
Scenario 7: Cumulative impact of the technical team's preferred assumptions on the cost-effectiveness estimate – plus diagnostic testing costs	[redacted]	[redacted]

Notes on changes to the scenario:
~ Uses the 72% response rate from the trial data rather than ERG preferred [redacted]%. The reason for this is that when replicating the ERG model results for this scenario, it was only possible to replicate the results using the 72% response rate, although the description in the ERG report claims to have used the [redacted]%
** Also includes same post-progression utility as per the ERG model, although the ERG report does not mention that this was done for scenario 3*

- The **base-case** from the company submission was updated to refresh only the PFS and OS data for larotrectinib. No other changes were applied to the model and [redacted] curves were used for both PFS and OS. With the use of the expanded dataset, the ICER [redacted] due to overall survival [redacted] and progression-free survival [redacted], and therefore treatment costs [redacted].
- **Scenario 1**: The ERG base case was updated to only refresh PFS and OS, stratified by response, for larotrectinib. No other changes were applied to the model, with [redacted] and [redacted] used for PFS and OS, respectively. In line with the base case, the ICER [redacted] with the same key drivers; [redacted] OS and [redacted] PFS.

- **Scenario 2:** Adjusting only the objective response rate (per the BHM) to include the SAS3 dataset resulted in consistent changes to the ICER as with the original n=102 dataset.

- **Scenario 3**: Use of comparator post-progression survival for larotrectinib results in an [redacted] ICER vs the ERG base case with the updated data. However, the relative difference [redacted] than with the n=102 dataset due to a [redacted] in responder and non-responder survival curves, [redacted] the impact of switching from responder to non-responder post-progression survival. Specifically, the tail of the KM curve for responders with the expanded data results in [redacted] projections for survival. Therefore, this scenario where non-responder survival is assumed for both arms (larotrectinib and comparator) does not have [redacted] shift (as compared to the n=102 data).

- **Scenario 4**: Including the same post-progression utility for larotrectinib as the comparator on top of the ERG base case resulted in a similar [redacted] in the ICER for both the n=102 dataset and updated n=153 dataset.

- **Scenario 5**: Including testing costs for larotrectinib in line with the ERG calculations on top of the ERG base case resulted in a similar [redacted] in the ICER for both the n=102 dataset and the updated n=153 dataset.

- **Scenario 6**: Including all the changes within the scenarios, excluding testing costs, on top of the ERG base case had a logical combined impact on the ICER. Whilst the ICER again [redacted], the magnitude of [redacted] in line with the previously described scenario 3

- **Scenario 7**: Including testing costs for larotrectinib in line with the ERG calculations on top of the previously described scenario 6 resulted in a similar relative [redacted] in the ICER for both the n=102 dataset and the updated n=153 dataset.

Compared to the original dataset, the expanded dataset predicts [redacted] QALY gain from larotrectinib ([redacted] QALYs vs [redacted] QALYs). This reflects [redacted] QALY benefit from an [redacted] progression-free interval outweighed by [redacted] benefit in the post progression state. The ICER is [redacted] in analyses that use the post progression data. In scenarios 3, 6 and 7 the ERG disregarded data collected after progression and replaced it with assumptions; the ICERs in these scenarios [redacted]. The range of ICERs generated by the ERG scenarios remains wide [redacted]) but the uncertainty is reduced compared to the original dataset (range [redacted]).

4. Updated epidemiological data

Further to a recently conducted systematic literature review (SLR) and meta-analysis (MA) of real world evidence on the epidemiology of NTRK gene fusion in solid tumours, Bayer present in this response updated estimates of potentially eligible patients. Whilst this is an evolving field and the number of patients eligible is uncertain, using a SLR and MA is the most robust evidence based approach at this time.

The SLR and MA were referred to in our response to the clarification questions, but further work has now been completed.

The output of the MA was an NTRK frequency per histology. These histologies were grouped within tumour type. The histologies are not mutually exclusive within a tumour type, since an NTRK publication could focus on a tumour, a subgroup of this tumour (histology) or multiple subgroups.

In order to calculate the NTRK incidence per tumour, we used the NTRK tumour frequency from the MA. When this value was not available or null, we identified the histology most representative of the tumour type. If the most representative histology had a null NTRK frequency, we selected the 2nd most representative, in order to avoid an underestimate of the entire NTRK gene fusion incidence in the population.

When a histology frequency was used to calculate the NTRK incidence per tumour, the global epidemiology department at Bayer identified in the literature the histology proportion within tumour. Each of the tumour incidences were then extracted from the European Cancer Information System when available, or the literature.

Finally, to calculate the "NTRK incidence per tumour (per 100,000)", we multiplied 3 terms: ‘Histology Proportion Within Tumour’, ‘Tumor Incidence (per 100,000)’ and ‘NTRK MA Frequency % (95% CI)’.

The number of expected incident patients in England was calculated by multiplying the NTRK incidence per tumour with the English population size. ***The detailed calculations are provided in an appendix.***

The estimated eligible population in England is set out by tumour location in the table below.

Tumour location	Estimated number in England with NTRK gene fusion
NSCLC	[redacted]
Salivary gland	[redacted]
Melanoma	[redacted]
Colorectal	[redacted]
Appendix	[redacted]
STS adults (GIST)	[redacted]
STS adults (non-GIST)	[redacted]
Bone Sarcoma	[redacted]
STS paediatrics (CMN)	[redacted]
STS infantile sarcoma	[redacted]
Secretory breast	[redacted]
Cholangiocarcinoma	[redacted]
CNS	[redacted]
Pancreas	[redacted]
Thyroid	[redacted]
TOTAL	[redacted]

When considering all tumour types with NTRK fusions identified in the SLR i.e. including and over and above those included in the trial programme in the dataset used in the submission, the total estimated number for England is [redacted].This is very close to the estimate Bayer presented in the original submission, lending confidence to these figures. Further, this piece of work identified that the tumour types covered in the trial represent [redacted]% of all those identified in the literature as being associated with NTRK gene fusion.

References

(1) *Hyman DM, van Tilburg CM, Albert CM, et al. Durability of response with larotrectinib in adult and paediatric patients with TRK fusion cancer. Presented at ESMO Congress 2019; September 27-October 1, 2019; Barcelona, Spain. Poster 445PD*

NICE National Institute for Health and Care Excellence

Technical engagement response form

Larotrectinib for treating NTRK fusion-positive advanced solid tumours [ID1299]

As a stakeholder you have been invited to comment on the technical report for this appraisal. The technical report and stakeholders' responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

We need your comments and feedback on the questions below. You do not have to answer every question. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be summarised and used by the technical team to amend or update the scientific judgement and rationale in the technical report.

Deadline for comments: **5pm Friday 18 October 2019**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the technical report which summarises the background and submitted evidence. This will provide context and describe the questions below in greater detail.
- Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.
- Do not include medical information about yourself or another person that could identify you or the other person.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow. If confidential information is submitted, please also send a second, fully

redacted, version of your comments (AIC/CIC shown as [redacted]). See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	**Jayne Bressington**
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**GIST Support UK - Respondent**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**

Questions for engagement

Issue 1: Prevalence of NTRK gene fusions	
Is the distribution of patients in the pooled analysis generalisable to NHS clinical practice?	
What is the total number of patients who would receive larotrectinib?	We estimate that the total number of GIST cancer patients who might qualify to receive larotrectinib is less than 8 per annum. e.g. There are c.800 new GIST patients diagnosed each year in England (numbers confirmed by PHE). It is estimated that c.125 are wild-type GIST and that of these 14% (17) are quadruple wild-type GIST. Quadruple wildtype GIST's lack mutations in the KIT, PDGFRA, RAS pathways and their succinate dehydrogenase (SDH) complex is intact. It is within the quadruple WT GIST patient group that NTRK fusions have been found. Of these 50% are likely to be clear of disease further to surgery and will not require therapy. The remaining 50% (c. 8 patients) when screened for NTRK fusions may find rare ones with the fusion.
Issue 2: Treatment pathway and positioning	
How will the term 'satisfactory' be defined in clinical practice?	
For each tumour type, at what point(s) in the respective treatment pathways will larotrectinib be used in clinical practice?	Currently the standard lines of treatment for GIST patients involve surgery then three lines of TKI therapy so we assume that larotrectinib will be used in fourth line. In the future, when sequencing becomes more readily available it may be that testing happens differently and identifies NTRK fusions earlier which would logically mean that patients with NTRK fusions should receive the drug earlier in the treatment pathway.
Is the clinical evidence for larotrectinib generalisable to the positioning in the treatment pathways in clinical practice in England?	

Issue 3: NTRK gene fusion testing	
What is the likely screening pathway to identify NTRK fusion positive solid tumours?	The standard screening pathway for GIST is as follows: 1. Molecular testing to review KIT (exons 8, 9, 11, 13, 17) and PDGFRA (exons 12, 14 and 18) mutation analysis. KIT exon 8 mutation has been reported in less than 0.5% of sporadic and familial GIST and its analysis should be performed in KIT (exons 9, 11, 13, 17) and PDGFRA (exons 12, 14, 18) wild type and cSDH GIST where KIT exon 8 was not analysed previously (Hartmann2005)(Huss2013)(Ito2014). 2. Following molecular testing, SDHB immunohistochemistry should be performed on all wtGIST and can be considered after multidisciplinary discussion for TKI-mutant GIST with a high risk clinical phenotype (e.g gastric GIST in patients under the age of 50, multifocality or a personal or family history of PPGL) (Mietenen 2011). 3. Immunohistochemistry for SDHB has proven to reliably identify GIST with SDH complex deficiency (Gill2010)(Gaal2011). 4. SDHA immunohistochemistry 5. BRAF is often performed as part of a NGS multi gene panel. IHC (Huss2017) 6. NF1, skeinoid fibers, clinical input. 7. **NTRK - IHC/fusion**
At what point in the treatment pathway for each tumour type will NTRK gene fusion testing be carried out?	The point at which NTRK fusion gene testing is carried out will vary for each patient. The testing happens where surgery or a biopsy has provided a sample of the tissue that can be tested. From a GIST perspective and using current protocols, NTRK gene fusion testing will be carried out when it has been identified that the patient has tested negative for all of the other known GIST mutations (as detailed above). This group of patients are currently classified as **"quadruple negative GIST".** In practice this should ideally happen at the start of treatment but the current reality is that patients are classified by Immunohistochemistry as being either mutated GIST or quadruple WT and NTRK fusion testing is not happening as standard because it is so new. Maybe when sequencing becomes standard it will happen at the start**.**

What proportion of the screening costs should be included in the economic model for this appraisal? Or is it appropriate to exclude all testing costs from the analyses?	We understand that screening for NTRK fusions will be naturally part of the whole genome sequencing panel in GIST and Sarcoma so we think that this indicates that testing costs should be excluded.
Issue 4: Identification of NTRK gene fusions – diagnostic accuracy	
What is the expected diagnostic accuracy of NGS testing?	Very accurate.
Is there a testing sequence that could avoid a substantial number false positive results in low NTRK fusion-positive tumour types?	
Is it appropriate to limit testing to avoid false positive results and the associated costs?	
Issue 5: Primary CNS tumours	
Should patients with primary CNS tumours be included in the analysis?	
Issue 6: Trial study design	
Is it appropriate to consider the 'basket' trial design for statistical evidence of heterogeneity?	
Issue 7: Heterogeneity of response across different solid tumour types	
Is a homogeneous response to larotrectinib across different tumour types a reasonable assumption?	
Is the Bayesian Hierarchical Modelling (BHM) framework an appropriate method to capture the heterogeneity of response across different solid tumour types for this appraisal?	

Would it be appropriate to apply the BHM framework to explore the heterogeneity in the time to event outcomes?	
Issue 8: Constructing a comparator arm	
Is the company's comparator arm suitable for decision making?	
Is it appropriate to use non-responders as a proxy for patients not having an active treatment or previous line of therapy to generate a comparator arm for this appraisal?	
Issue 9: Comparator treatments	
Are the comparators identified representative of where larotrectinib would be used in the treatment pathways?	
Issue 10: Subsequent therapies	
What subsequent treatments would be expected in clinical practice after larotrectinib?	We understand that a treatment currently called Loxo 195 has been developed as a companion drug to tackle the development of resistance point mutations while on Larotrectinib. So we hope that patients will be offered this companion drug which has been designed to specifically address this issue. American clinicians advise that a percentage of NTRK fusion patients do not respond to some NTRK inhibitors but that they then respond to another. We understand this is because each of the NTRK fusion inhibiting drugs have a slightly different target panel. Thus, logically another NTRK

	inhibiting drug could be considered after Larotrectinib and LOXO195 if available to patients in England.
Should experimental treatments be adjusted for in this analysis?	
Issue 11: Model structure	
What is the most appropriate model structure for this appraisal?	
Issue 12: Extrapolation of overall and progression-free survival	
Which parametric curve is most appropriate for extrapolating overall and progression-free survival in the company base case?	
Which parametric curve is most appropriate for extrapolating overall and progression-free survival in the response-based analysis?	
Issue 13: Drug wastage and adherence	
Have potential drug wastage costs and adherence costs been appropriately included in the model?	
Issue 14: Administration costs and resource use	
Have administration costs and resource use been adequately captured in the company's model?	
Issue 15: Implementation and training costs	

What additional infrastructure and training requirements could be considered for this appraisal?	
Issue 16: Utility values	
How closely do the utility values modelled match the utility values of patients in clinical practice?	
Is there justification for considering post-progression utility values to be different between larotrectinib and best supportive care?	
Issue 17: End of Life	
What is the life expectancy of the patient group receiving established management?	
What is the extension to life of the patient group receiving larotrectinib?	
Issue 18: Innovation	
Is larotrectinib an innovative treatment?	Yes
Issue 19: Cancer Drugs Fund	
Does larotrectinib meet the criteria for inclusion in the Cancer Drugs Fund?	
What data would be most useful to collect to address the outstanding uncertainties? For example, unrepresented tumour types.	

NICE National Institute for Health and Care Excellence

Technical engagement response form

Larotrectinib for treating NTRK fusion-positive advanced solid tumours [ID1299]

Questions for engagement

Issue 1: Prevalence of NTRK gene fusions	
Is the distribution of patients in the pooled analysis generalisable to NHS clinical practice?	The substantial differences in distribution across tumour sites seen in Table 1 raise substantial concerns as to whether the NAVIGATE/SCOUT trial is representative of real patients. In particular, we note the over-representation of salivary cancers and soft tissue sarcomas, and the under representation of breast cancer. These differences may mean that estimated response rates and survival distributions may not represent what will be seen in actual practice. The ERG also note that it is unclear how the economic model was updated to reflect the alternative tumour site distribution, as no new version of the model was submitted. It appears that the company has only updated the distribution of tumour sites from which the weights for the comparator are drawn, but the ERG was to replicate the exact values reported in Table 3 and 4 of the company's response to the technical engagement. The implications of a different distribution of NTRK fusion rates across tumour sites does not appear to have been reflected in terms of overall response rates and survival outcomes for patients treated with larotrectinib, so these analyses do not allow exploring the full impact of the alternative tumour site distribution on the cost-effectiveness estimates.
What is the total number of patients who would receive larotrectinib?	No comment

Issue 2: Treatment pathway and positioning	
How will the term 'satisfactory' be defined in clinical practice?	No comment
For each tumour type, at what point(s) in the respective treatment pathways will larotrectinib be used in clinical practice?	No comment
Is the clinical evidence for larotrectinib generalisable to the positioning in the treatment pathways in clinical practice in England?	No comment
Issue 3: NTRK gene fusion testing	
What is the likely screening pathway to identify NTRK fusion positive solid tumours?	If RNA-based NGS testing is implemented for all cancer patients, diagnostic accuracy to detect NTRK fusion may be greater than presented in the ERG report. This will be addressed in our updated analyses.
At what point in the treatment pathway for each tumour type will NTRK gene fusion testing be carried out?	No comment
What proportion of the screening costs should be included in the economic model for this appraisal? Or is it appropriate to exclude all testing costs from the analyses?	The approach taken to add testing costs is different to the approach taken by the ERG, which calculated a weighted NTRK fusion testing cost based on the distribution of the tumour sites and assuming different testing strategies by tumour site (dependent on NTRK fusion rate and current testing practice in the NHS for tumour sites that already undergo routine testing with NGS). The ERG considered the unit costs per IHC and NGS test to be £75 and £350, respectively (see Table 56, ERG report). It is unclear to the ERG how the company derived the cost of [redacted] per test, which is applied in the company's analyses. .

	Instead the ERG use a weighted overall cost of testing for larotrectinib of £18,670 (Pg. 182 of the ERG report).
Issue 4: Identification of NTRK gene fusions – diagnostic accuracy	
What is the expected diagnostic accuracy of NGS testing?	The purpose of the National Genomic Medicine Service is not relevant. It is essential that the accuracy of testing for NTRK fusions is known to ascertain the likely rate of false-positives (in whom larotrectinib cannot be effective), particularly for tumour types where NTRK fusion is rare.
Is there a testing sequence that could avoid a substantial number false positive results in low NTRK fusion-positive tumour types?	No comment
Is it appropriate to limit testing to avoid false positive results and the associated costs?	No comment
Issue 5: Primary CNS tumours	
Should patients with primary CNS tumours be included in the analysis?	No comment
Issue 6: Trial study design	
Is it appropriate to consider the 'basket' trial design for statistical evidence of heterogeneity?	Given the wide diversity in tumour types across adults and children, with variation in NTRK fusion prevalence, the ERG considers that the *potential* for across-site heterogeneity must be considered. This is an *a priori* modelling decision, and is not dependent on the quantity or quality of the data An appropriate analogy here is with standard meta-analyses, where random effects methods would nearly always be used to account for possible heterogeneity, regardless of the size of the included studies.

Issue 7: Heterogeneity of response across different solid tumour types	
Is a homogeneous response to larotrectinib across different tumour types a reasonable assumption?	See response above. Also, the ERG disagrees with the claim that: "larotrectinib has demonstrated a large magnitude of effect irrespective of tumour site"; our BHM analysis found that this was not the case.
Is the Bayesian Hierarchical Modelling (BHM) framework an appropriate method to capture the heterogeneity of response across different solid tumour types for this appraisal?	The ERG notes the following with regard to Bayesian Hierarchical models (BHM). 1. The BHM considered by the ERG was developed specifically for the analysis of basket trials. 2. Bayesian modelling of this kind is particularly useful where data are limited, because it permits "borrowing strength" from tumour sites where data is more plentiful, and the use of suitable priors to characterise plausible realities. Therefore it is not the case that Bayesian models become more suitable as data accumulates; indeed the opposite is true. 3. The BHM models all tumour sites and all data in one model, assuming "exchangeability" across sites. The tumour site are not analysed separately. 4. Given the novelty of tumour-agnostic interventions and basket trials, statistical analyses of such interventions must be innovative, particularly when considering survival analysis. 5. The BHM is not intended to replicate the ORR provided by the company, but to assess whether that ORR was robust to possible heterogeneity. The ERG found it was not robust. The ERG notes that a prior for mu in the BHM of Normal(19.3472979, precision=1/10) is equivalent to assuming that all patients respond to larotrectinib. This is clearly an unreasonable prior, which lends weight to the ERG's conclusion that the original ORR overestimates the effectiveness of larotrectinib. The ERG's BHM is reasonably robust to sensible choices of prior for mu [Evidence can be provided if required].
Would it be appropriate to apply the BHM framework to explore the heterogeneity in the time to event outcomes?	The assumption of surrogacy between response and survival was required because the company did not supply any data on survival outcomes by tumour site, despite repeated requests by the ERG. With proper

	data on survival by tumour site (such as full IPD) an appropriate random-effects Bayesian survival model could be fitted.
Issue 8: Constructing a comparator arm	
Is the company's comparator arm suitable for decision making?	No comment
Is it appropriate to use non-responders as a proxy for patients not having an active treatment or previous line of therapy to generate a comparator arm for this appraisal?	No comment
Issue 9: Comparator treatments	
Are the comparators identified representative of where larotrectinib would be used in the treatment pathways?	No comment
Issue 10: Subsequent therapies	
What subsequent treatments would be expected in clinical practice after larotrectinib?	No comment
Should experimental treatments be adjusted for in this analysis?	The ERG notes that NAVIGATE and SCOUT were single-arm trials, so experimental treatments were unlikely to have been received by comparator patients outside these trials. LOXO-195 can only be given to patients who have previously received larotrectinib, so it is not the case that it "could equally be offered to patients who had received standard of care" Given these issues we consider that adjustment for reception of LOXO-195 is essential.

Issue 11: Model structure	
What is the most appropriate model structure for this appraisal?	The ERG agreed that the company appropriately explored alternative approaches to model the comparator survival outcomes and that all approaches have limitations and may result in biased estimates of treatment effectiveness (see 5.2.4.2 of the ERG report). However, the non-responder control provides a more transparent and flexible alternative to modelling comparator effectiveness than the pooled historical comparator. In addition, only by using the responder-based model, can heterogeneity according to tumour type be explored, given the lack of further data provided by the company. The ERG agree that the relationship between response status and survival outcomes is highly uncertain (as discussed in Section 6.3.1 of the ERG report). The ERG request for PFS/OS data according to tumour type was not granted by the company, and therefore the ERG assume a common distribution of PFS and OS across tumour sites, conditional on response status. The adjustments made to the ratio between the progression-free and post-progression made by the ERG in the Section 6.3.4, were used to explore the validity of the survival projections made in the company base case, in particular the gains in post-progression survival for larotrectinib. Crude adjustments to post-progression survival were made in the absence of being able to separately track where the gains in post progression gains in survival for larotrectinib and the comparator occur (as discussed in Section 6.3.4).
Issue 12: Extrapolation of overall and progression-free survival	
Which parametric curve is most appropriate for extrapolating overall and progression-free survival in the company base case?	No comment

Which parametric curve is most appropriate for extrapolating overall and progression-free survival in the response-based analysis?	The ERG explore the use of alternative survival model to illustrate the uncertainty in the survival gains for larotrectinib. As discussed in Section 6.3.3 the ERG question the face validity of the post-progression survival gains for larotrectinib ([redacted]years compared to [redacted] years in PFS). The ERG considers that the shorter post-progression survival gains for larotrectinib using the Gompertz distribution for OS, makes it more clinically plausible than the use of the Weibull distribution (the company base-case). While the use of the Gompertz distribution to extrapolate OS may generate more conservative survival predictions, given the immaturity of the OS data, this may still result in optimistic survival projections.
Issue 13: Drug wastage and adherence	
Have potential drug wastage costs and adherence costs been appropriately included in the model?	The company did not provide data that allowed adjustments to the effectiveness of larotrectinib based on the dose received by paediatric patients.
Issue 14: Administration costs and resource use	
Have administration costs and resource use been adequately captured in the company's model?	No comment
Issue 15: Implementation and training costs	
What additional infrastructure and training requirements could be considered for this appraisal?	No comment
Issue 16: Utility values	
How closely do the utility values modelled match the utility values of patients in clinical practice?	No comment
Is there justification for considering post-progression utility values to be different	No comment

between larotrectinib and best supportive care?	
Issue 17: End of Life	
What is the life expectancy of the patient group receiving established management?	No comment
What is the extension to life of the patient group receiving larotrectinib?	No comment
Issue 18: Innovation	
Is larotrectinib an innovative treatment?	No comment
Issue 19: Cancer Drugs Fund	
Does larotrectinib meet the criteria for inclusion in the Cancer Drugs Fund?	No comment
What data would be most useful to collect to address the outstanding uncertainties? For example, unrepresented tumour types.	No comment

Evidence Review Group's Report

Larotrectinib for treating NTRK fusion-positive advanced solid tumours (addendum)

ERG commentary on evidence submitted by the company further to the Technical Engagement

Produced by CRD and CHE Technology Assessment Group, University of York, Heslington, York YO10 5DD

All commercial-in-confidence (CIC) data and academic-in-confidence (AIC) have been redacted.

1 Overview

The company's additional evidence submitted in response to the technical engagement included:

1 Survival outcomes for patients who underwent post-progression experimental treatments on the efficacy data (July 2018 data cut)
2 Further clarification on the point raised on the teleconference regarding the potential for dose escalation post progression
3 Analysis of the updated data cut from the trial programme
4 Cost-effectiveness results based on a version of the electronic model updated with the latest efficacy data for larotrectinib (February 2019 data cut, investigator assessment)
5 Updated epidemiological data.

The Evidence Review Group (ERG) was requested by NICE to provide commentary on the additional evidence submitted by the company in response to the technical engagement. Due to the limited time available, the additional work undertaken by the ERG does not constitute a formal critique of the company's resubmission and hence does not accord with the procedures and templates applied to the original submission. Furthermore, the company has not submitted a version of the electronic model updated with the additional evidence, and therefore, the ERG could not check the implementation of the proposed changes or replicate the results of the additional economic analyses presented by the company.

Overall, the additional clinical evidence supplied by the company based on the February 2019 data cut adds data for a further 60 patients. It appears that most of these new patients had soft tissue sarcomas, infantile fibrosarcoma or thyroid cancer. There was a slight increase in overall response rate (ORR) compared to that presented in the original submission, rising to 79% (72 – 85). However this may be a consequence of using investigator-led data in the more recent analysis.

[redacted]The new evidence provided data by tumour site for response rate but no other outcomes. The ERG was therefore able to update the Bayesian Hierarchical model (BHM), but could not update any other analyses, or perform any new analyses.

The ERG presents, in the next sections, a brief critique of the document submitted by the company, results of the updated BHM, and a discussion of issues on the diagnostic accuracy of NGS screening which were raised at technical engagement teleconference.

2 Potential impact of experimental post-progression treatments on the efficacy data

[redacted] Since the reasons for giving experimental treatment after progression are unclear, the ERG still considers that the use of LOXO-195 may lead to overestimation of survival in post-progression patients.

3 Further clarification on the point raised on the teleconference regarding the potential for dose escalation post progression

The company clarified that, contrary to what they had suggested in technical engagement teleconference, the average dose used in the original economic model, included post-progression exposure and any dose escalation. The company also provided individual patient data for the [redacted] patients treated in post-progression with larotrectinib alongside details on dose escalation for the [redacted] patients who experienced it.

4 Latest clinical efficacy data cut

4.1 Company's updated analysis

The updated analysis of the three included trials in February 2019 included 159 patients, 153 of whom had evaluable results; an increase of 60 patients compared to the July 2018 analysis considered in the company submission. Table 1 summarises the key results of both analyses, using the investigator led results in both cases, for comparability.

These results are very similar, suggesting that updated data are consistent with the original data, and conclusions are unlikely to change regarding the effectiveness of larotrectinib. As the February 2019 results are based on investigator-led interpretations, rather than independent assessment they may overestimate actual response (particularly partial response, as was the case in the company submission).

Table 1 Comparison of key results between July 2018 and February 2019 cut-offs

	February 2019	July 2018 (Investigator-led)	July 2018 (Independent)
Response			
Evaluable patients	153	[redacted]	93
ORR	79% (72 to 85)	[redacted]	67% (62 to 81)
Complete response	16%	[redacted]	17%
Partial response	63%	[redacted]	55%
Stable disease	12%	[redacted]	15%
Progressive disease	9%	[redacted]	10%
Not determined	4%	[redacted]	3%
		[redacted]	
Median response rates (months)		[redacted]	
Duration of response	35.2 (22.8 – NE)	[redacted]	Not estimable
Progression-free survival	28.3 (22.1 – NE)	[redacted]	27.4 (13.8 – NE)
Overall survival	44.4 (36.5 – NE)	[redacted]	Not estimable

4.2 Survival curves

[redacted]

4.3 Additional ERG analyses

Table 2 presents the updated data by response and tumour site. As the [redacted] non-evaluable patients were not identified, the ERG made assumptions as to which patients these were, based on reported percentages. The ERG notes that nearly all the new patients and new responses were in the four tumour sites that were already most numerous in the July 2018 cut (STS, salivary, IFS, thyroid). There are very few new patients in tumour sites where NTRK fusion is rare, where there were few patients in July 2018, and only three new tumour sites.

Table 2: Patient numbers by data cut

	July 2018 cut		February 2019 cut	
Tumour	**Patients**	**Responders**	**Patients**	**Responders**
Soft tissue sarcoma	20	16	[redacted]	[redacted]
Salivary gland	17	15	[redacted]	[redacted]
Infantile fibrosarcoma	13	12	[redacted]	[redacted]
Thyroid	10	7	[redacted]	[redacted]
Lung	7	5	[redacted]	[redacted]
Melanoma	7	3	[redacted]	[redacted]
Colon	6	2	[redacted]	[redacted]
GIST	5	5	[redacted]	[redacted]
Bone sarcoma	2	1	[redacted]	[redacted]
Cholangiocarcinoma	2	0	[redacted]	[redacted]
Appendix	1	0	[redacted]	[redacted]
Breast	1	0	[redacted]	[redacted]
Congenital mesoblastic nephroma	1	1	[redacted]	[redacted]
Pancreas	1	0	[redacted]	[redacted]
Hepatocellular	0	0	[redacted]	[redacted]
Prostate	0	0	[redacted]	[redacted]
Unknown primary	0	0	[redacted]	[redacted]

* [redacted] *in new data, ERG have assumed this was an error*

The ERG used the same Bayesian Hierarchical Model as in its report (see ERG report section 4.6.1; tables 7 and 8; figures 8 and 9), and applied it to the new data. The predicted mean response by tumour site are shown in Figure 1.

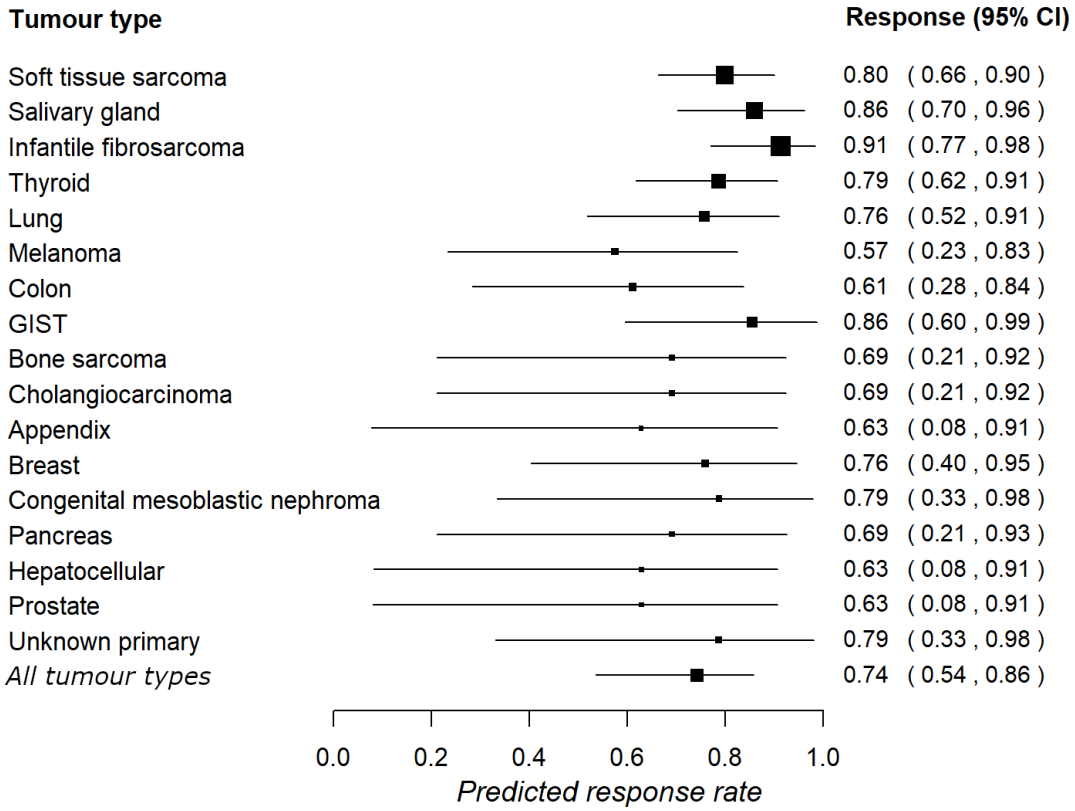
The model gave an estimated ORR across all tumour sites of 74% (95% CrI 54 to 86). This is higher than in the analysis in the ERG report [64% (95% CrI 29 to 83)]. This reflects the increase in response rate (from 72% to 79%) between the two data cut-offs, and may be a consequence of differences between investigator and independent assessment of response.

As in the ERG report, the BHM gives a lower estimate of overall response rate than that supplied by the company, because the BHM accounts for across-site heterogeneity. The ERG still considers that the BHM gives a more realistic estimate of the response to larotrectinib across all patients than the company's analysis. The BHM model shows clear evidence of heterogeneity in response across tumour sites (heterogeneity estimate: 1.01, 95% CrI 0.21 to 2.32), but this appears to have reduced when compared to the estimate in the ERG report (1.58, 95% CrI 0.38 to 3.64).

Because the ORR has increased in the updated BHM, so the by-tumour site predicted responses have also increased accordingly. The probability of the response rate exceeding 30% is now over 85% in all tumour sites. The probability of the response rate exceeding 10% is now over 95% in all tumour sites.

The ERG notes that the validity of this updated BHM is unclear, because is based on investigator-led assessments of response, rather than the independent assessment used in the ERG report.

Figure 1 Predicted median response rates from BHM based on February 2019 data



5 Cost-effectiveness results with expanded dataset

The company states that the ERG revised model has been updated to include the latest ESMO PAS+SAS1 (n=153) data cut (February 2019) for OS and PFS. The company also states that no parameters other than the PFS and OS curves were updated to reflect the latest PAS+SAS1 data (e.g. update of tumour site distribution and of the historical comparator data to reflect the additional tumour sites).

Results of the company's base-case analysis conducted in the updated model are compared to the company's base-case analysis submitted in July 2019 with the proposed patient access scheme (PAS) discount of approximately [redacted]% over expected list price of larotrectinib. The updated base-case results suggest a [redacted] on mean incremental QALYs and an [redacted] on mean incremental costs compared to the original base-case ([redacted] QALYs [redacted] and [redacted]), resulting in an [redacted] ICER of [redacted] per additional QALY. The results are driven by [redacted] in the time spent in Progression-free for patients treated with larotrectinib ([redacted]),

which [redacted] the cost of treatment ([redacted]), and a [redacted] of the OS with consequent [redacted] of time spent in progressed disease ([redacted]). The QALY gains from [redacted] time in the progression-free health state ([redacted] compared to the base-case for the July 2018 data cut-off) are offset by the [redacted] costs of treatment and by the [redacted] of QALY gains in progressed disease ([redacted] and [redacted] QALYs [redacted]). The company notes that results should be interpreted with caution due to the short follow-up of the additional patients. The comparability of the two analyses is also hindered by the fact that the PFS outcomes are based on the investigators' assessment for the later data cut, whereas PFS assessment was based on independent central review for the earlier data cut.

The ERG notes that, despite the [redacted] in the ratio post-progression and progression free QALY gains, the company's updated base-case analysis still suggests that the majority of QALY gains for larotrectinib are accrued in the progressed disease health state. As stated in Section 2 of this document, the company does not provide persuasive evidence that the experimental treatments provided to larotrectinib patients who have progressed are not contributing to the overestimation of survival benefits in post-progression, and this remains an area of uncertainty. In the absence of an updated model version, the ERG could not validate the company's updated base-case analysis or explore the impact of alternative survival assumptions in the cost-effectiveness results when the later survival data is considered.

The company also presents ICERs for a selection of the ERG additional analyses in the ERG report and with the NICE technical team preferred assumptions, estimated in the updated version of the model and using the response-based survival models. These scenarios are incorrectly referred to as being included within Table 8 of the ERG technical report. Table 8 in the NICE technical report is reproduced below in Table 3.

Table 3 Reproduction of Table 8 in the NICE technical report

Alteration	Technical team rationale	ICER	Change from base case
Company base case	–	[redacted]	-
ERG base case	Issues 7-11	[redacted]	[redacted]
ERG base case + SAS3 dataset	Issue 5	[redacted]	[redacted]
ERG base case + post-progression survival equal for larotrectinib and comparator	Issue 10	[redacted]	[redacted]
ERG base case + same post-progression utility for larotrectinib	Issue 16	[redacted]	[redacted]
Drug wastage and adherence, administration costs and resource use	Issue 13-14	[redacted]	-
ERG base case + diagnostic testing costs	Issue 3	[redacted]	[redacted]
Cumulative impact of the technical team's preferred assumptions on the cost-effectiveness estimate – no diagnostic testing costs	-	[redacted]	[redacted]
Cumulative impact of the technical team's preferred assumptions on the cost-effectiveness estimate – plus diagnostic testing costs	-	[redacted]	[redacted]

The company states that they could not replicate the ERG base-case analyses with the response rates estimated by the BHM for the ePAS2 population and the full integrated efficacy analysis for the July 2018 data cut (64% and 57%, respectively) with: i) post-progression survival equal for larotrectinib and comparator, and ii) same post-progression utility for larotrectinib (scenario 3 and 4 in the company's document, respectively). The ERG did not conduct these two cumulative scenarios in the report. The ICERs reported in the NICE technical report (Issue 10 and 16 in Table 3) correspond to the results of scenarios 9 and 11.1 in the ERG report which both assumed an ORR of 72% and not the ORR estimates from the BHM.

The majority of company additional analyses, with the updated model, result in [redacted] ICERs compared to the analyses in the model using the July 2018 data cut, which the company attributes to [redacted] of OS and an [redacted] of PFS with consequent [redacted] in treatment costs. The exceptions to this are the scenarios where the same post-progression survival is assumed for larotrectinib and the comparator (scenarios 3, 6 and 7). The company states that the ICER estimated with the updated model for these scenarios are [redacted] than for the corresponding analysis in the model with earlier data-cut (July 2018) due to [redacted] in responder and non-

responder survival curves, [redacted] the impact of switching from responder to non-responder post-progression survival. The company considers that the *"range of ICERs generated by the ERG scenarios remains wide ([redacted]) but the uncertainty is reduced compared to the original dataset (range [redacted])"*.

The ERG could not validate the results of these additional analyses given that an updated version of the model was not submitted. The ERG could not check the company's interpretation of the results, given that these are not reported in full and only ICERs are presented for these analyses. Furthermore, it is not clear which ORR was assumed for the company's additional analyses with the exception of the estimate used for scenarios 3 and 4 (ORR=72%). Importantly, and as stated in Section 4 of this document, the efficacy data in the February 2019 data cut appears to be broadly consistent with the earlier data cut (July 2019), and the data is insufficient to demonstrate that the [redacted] are due to changes in the efficacy data. The use of investigator assessment for PFS instead of independent central review may have resulted on the overestimation of PFS estimates. Finally, the company's statement on reduction of the uncertainty when using the updated dataset based on the narrowing of the ICER range across scenarios may not be strictly valid. This cannot be judged given that i) the company only updated the survival curves in the model, ii) the additional tumour sites in the February data-cut have not been reflected in the new model, iii) the use of alternative progression criteria, and iv) the additional patients in the February 2019 data-cut have a short follow-up. Moreover, the range of ICERs does not fully capture the uncertainty in the model given the constraints of the model structure and the additional assumptions that it requires (e.g. surrogate relationship between ORR and survival outcomes).

6 New epidemiological data

The company reported the results of a systematic review and meta-analysis to estimate NTRK frequency per histology, and derived estimates of number of expected incident patients in England eligible for treatment with larotrectinib. The results of the systematic review and meta-analysis were obtained by updating a previous study submitted by the company in response to clarification questions. The company does not describe the nature of the updates.

The ERG has further comments on these data.

7 Diagnostic accuracy of NGS screening

This issue is not related to new material provided by the company, but to issues raised at the technical engagement teleconference, and is included here for convenience. It was noted on that call that the

diagnostic accuracy of DNA-based or RNA-based NGS screening for NTRK fusion might be very high (99.9% specificity was suggested). There appears to be some recent evidence for this claim (81.1% sensitivity, 99.9% specificity) [1].

The ERG have reassessed the possible rates of false positive findings of NTRK fusion in the light of these data, and given the new estimates of numbers of patients with NRTK fusions by tumour site provided by the company. The results are presented in Table 4.

While the number of false positive cases is much lower than previously predicted (see ERG report, Table 4) if specificity is 99.9% rather than the previously assumed 99%, there remain several tumour sites where false positive cases will make up a sizeable proportion of all patients, which will meaningfully reduce the practical effectiveness of larotrectinib in those tumour sites.

The ERG also notes that it is perhaps unlikely that 99.9% specificity could be achieved in actual practice: it is unclear whether all patients across all tumour sites will receive NGS testing, and human error, sample deterioration or contamination could all reduce the effective specificity.

References

[1] Solomon, J.P., Linkov, I., Rosado, A. *et al. NTRK* fusion detection across multiple assays and 33,997 cases: diagnostic implications and pitfalls. *Mod Pathol* (2019) doi:10.1038/s41379-019-0324-7

Table 4 Impact of imperfect diagnostic testing on larotrectinib efficacy

Cancer type	Eligible for testing	Eligible and has NTRK fusion	True positives (has NTRK fusion detected)	False positives (no fusion but test is positive)	% of treated who are false positive	% of treated who respond (based on 80% response)
NSCLC cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Salivary cancer patients (non-MASC)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
MASC cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Melanoma cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Colorectal cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Appendix cancer patients (assumed same as pancreatic)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
STS adults (GIST) cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
STS adults (non-GIST) cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Bone Sarcoma cancer patients (assumed same as STS adults)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
STS paediatrics cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
STS infantile sarcoma cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Secretory breast cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cholangiocarcinoma cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CNS cancer patients (assumed same as brain)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CNS paediatric cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pancreas cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Thyroid cancer patients	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Technical report

Larotrectinib for treating *NTRK* fusion-positive solid tumours

This document is the technical report for this appraisal. It has been prepared by the technical team with input from the lead team and chair of the appraisal committee.

The technical report and stakeholder's responses to it are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the appraisal committee meeting.

The technical report includes:

- topic background based on the company's submission
- a commentary on the evidence received and written statements
- technical judgements on the evidence by the technical team
- reflections on NICE's structured decision-making framework.

This report is based on:

- the evidence and views submitted by the company, consultees and their nominated clinical experts and patient experts and
- the evidence review group (ERG) report.

The technical report should be read with the full supporting documents for this appraisal.

1. Topic background

1.1 Disease background: Neurotrophic tyrosine kinase (*NTRK*) fusion-positive solid tumours

- There are 3 *NTRK* gene fusions, *NTRK1/2/3*
- *NTRK* gene fusions are oncogenic drivers and are found in a wide variety of cancers including non-small cell lung cancer, breast cancer, pancreatic cancer and rare tumour types such as sarcoma and papillary thyroid cancer
- The company reports an overall prevalence of less than 1% but prevalence of *NTRK* gene fusion varies across different tumour types, ranging from less than 1% prevalence (for example in non-small cell lung cancer [NSCLC]) to 91% to 100% prevalence (for example in secretory carcinoma of the salivary gland and infantile fibrosarcoma)
- Treatment for rare, advanced cancers is often limited to standard chemotherapy with associated toxicity

1.2 Appraisal background: tumour site agnostic treatments

- This is one of the first technologies to be appraised for a histology-independent indication, with treatment determined by the presence of a specific type of genomic alteration, rather than the location of the tumour
- The marketing authorisation for larotrectinib is for the treatment of adult and paediatric patients with solid tumours who have:
  - a tumour with a *NTRK* gene fusion, and
  - a disease that is locally advanced, metastatic or where surgical resection is likely to result in severe morbidity, and
  - no satisfactory treatment options
- This appraisal considers any tumour type exhibiting the *NTRK* 1, 2 or 3 fusions

- Genomic testing is required to identify solid tumours with *NTRK* 1, 2 or 3 fusions. Testing procedures are not standardised across all tumour types at present

1.3 Treatment pathway and positioning of larotrectinib

- There is no established treatment pathway specifically for patients with *NTRK* fusion-positive tumours. Treatment is guided by tumour-specific care guidelines

- The position where larotrectinib would be offered is likely to vary by the availability of effective treatments in each tumour:

    - More common tumour sites such as NSCLC, colorectal cancer (CRC), melanoma and pancreatic have guideline recommendations for multiple lines of therapy (such as chemotherapy, targeted therapy, and/or immunotherapy)

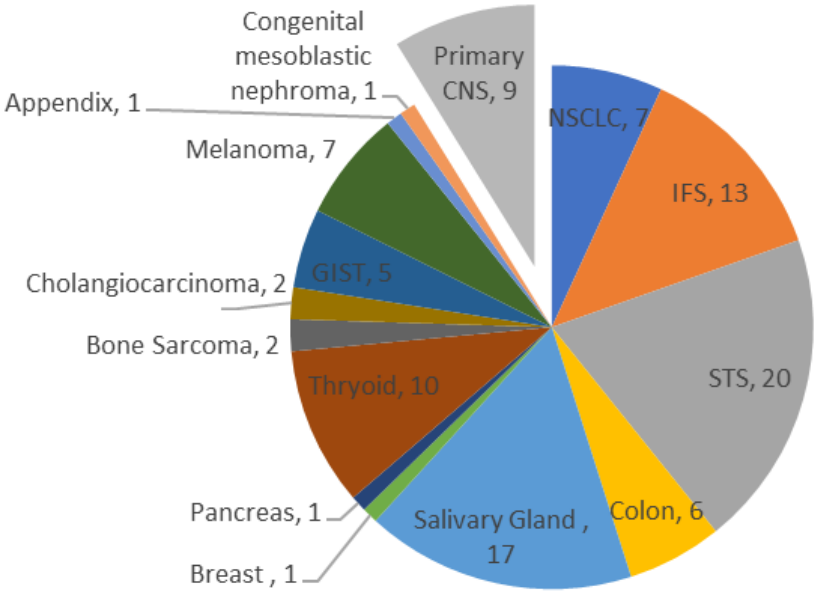
    - Less frequent tumour sites such as appendix, salivary gland, and secretory breast carcinoma have limited or no treatment guidelines or recommendations due to scarcity of evidence supporting systemic therapy. These rarer tumours are mainly treated with chemotherapy and/or surgery, or patients are enrolled in clinical trials.

1.4 Clinical evidence

- The company provided clinical effectiveness data for 68 adult patients and 34 paediatric patients with solid tumours enrolled in the NAVIGATE, SCOUT and LOXO-TRK-14001 clinical trials, combined into the integrated efficacy evaluable dataset. Results for patients with primary CNS tumours were presented separately in the SAS3 dataset (supplementary analysis set; n=9) but were included in the economic analysis. The remainder of the dataset was labelled the ePAS2 (extended primary analysis set; n=93)

dataset. The proportions of people included in the integrated efficacy analysis dataset with each solid tumour are given in Figure 1.

Figure 1 Tumour types in the company's integrated efficacy evaluable analysis (n=102), adapted from Table 8, CS



Abbreviations: NSCLC, non-small cell lung cancer; IFS, infantile fibrosarcoma; STS, soft tissue sarcoma; GIST, gastro-intestinal stromal tumours; CNS, central nervous system

- The primary CNS tumours were excluded from the company's primary analysis dataset. These patients were assessed using Response Assessment in Neuro-Oncology (RANO) criteria rather than Response Evaluation Criteria In Solid Tumours (RECIST) v1.1 criteria.

1.5 Key trial results

Median survival follow-up among all patients in the efficacy evaluable analysis set is [redacted] in ePAS2 and [redacted] in SAS3 (July 2018 data-cut). The company

provided some evidence from further data-cut at technical engagement but this is only presented in the summary of comments received at technical engagement.

Table 1: Clinical effectiveness results for larotrectinib for the company's efficacy evaluable population, July 2018 data-cut

NB: The ePAS2 dataset were assessed by independent review committee whereas the SAS3 dataset were assessed by the investigator.

1.6 Company's model structure

- The company model structure is a partitioned survival model with three separate health states.

Figure 2: Company's economic model structure

1.7 Data in economic model

- Larotrectinib arm: efficacy and safety data were based on results from the integrated efficacy evaluable dataset
- Comparator arm: the company constructed a composite comparator based on assumed established clinical management for each tumour type within the efficacy evaluable dataset. They extracted information on progression free survival (PFS) and overall survival (OS) Kaplan Meier (KM) curves, utility values, adverse event data, time on treatment data and resource use from previous NICE technology appraisals where available and literature values/ assumptions where none were available. This allowed creation of a comparator 'engine' for 12 different tumour types, each of which were weighted by the number of people in the trial to create the composite comparator arm

1.8 Key model assumptions

Area	Assumption	Company justification
Time horizon	40 years for adults 80 years for paediatric	This period is expected to represent a lifetime horizon.
Clinical effectiveness: PFS & OS (larotrectinib)	[redacted]	Best statistical fit to the larotrectinib data, judged to be clinically plausible and more appropriate than the [redacted] for modelling hazard trends with age.
Clinical effectiveness: PFS & OS (comparator)	Composite comparator, as in previous NICE TAs	Considers previous NICE committee decisions about survival extrapolation.
Clinical effectiveness: prognostic factors	No adjustment made for *NTRK* fusion positive status	Limited data available, therefore no adjustment is made for the prognostic factors of *NTRK* fusion-positive status.

Treatment duration	Larotrectinib treatment duration is equivalent to PFS	Larotrectinib is assumed to be administered until disease progression or unacceptable toxicity. Comparator treatment is assumed treatment to progression unless specified in the NICE TA.
Health-related quality of life (HRQoL)	Quality of life data for larotrectinib based on data collected by the company in NAVIGATE and SCOUT mapped to EQ-5D where necessary	To reflect the model structure with progression-free and progressed disease states. Consistent with previous appraisals.
	Quality of life data for comparator based on weighted average of data from previous NICE appraisals	Selected data were identified and accepted within previous NICE technology appraisals.
Adverse events	Treatment emergent grade 3-4 adverse events (AEs) that occurred in ≥5% of patients in the relevant treatment arm were included within the economic assessment	Threshold of 5% based on common assumption in NICE TAs. Anaemia and Neutropenia are the only adverse events included as disutilities in the model. Comparator adverse events are extracted from NICE TAs where available, weighted by the distribution of population mix in the integrated efficacy analysis.
Subsequent treatment	Subsequent treatments are not adjusted for in the model	Assumed that subsequent therapies would be equivalent between treatment arms.
Diagnostic testing	Costs of diagnostic testing are not included in the model	The company note that whole genome sequencing (WGS) is an area of innovation in the NHS Long Term Plan and since WGS can provide information about multiple targets, the company believe it is inappropriate to assign the cost of testing to this treatment.

2. Summary of the draft technical report

2.1 In summary, the technical team considered the following:

- Prevalence of *NTRK* gene fusions in each tumour site is uncertain. This will affect the number and mix of tumour types presenting in clinical practice and add uncertainty to the cost-effectiveness estimates (see issue 1)
- Issues with the generalisability of the trial to clinical practice increases the uncertainty in the cost-effectiveness estimates (see issue 2)
- Screening pathway, testing costs and testing accuracy when identifying *NTRK* gene fusion positive solid tumours remain uncertain and depend on the provisions set up by NHS England in a timeframe that aligns with this appraisal (see issue 3 and 4)
- Patients with primary CNS tumours should be included in the analysis (see issue 5)
- The 'basket' trial is designed to confirm response by tumour site but has not completed (see issue 6)
- Heterogeneity of response across tumour types is a source of uncertainty in this appraisal (see Issue 7)
- Issues with the robustness of the control arm increases the uncertainty with the cost-effectiveness estimates (see issues 8 and 9)
- The proportion of people receiving subsequent therapies, types of therapies and associated costs and health effects should reflect those given in clinical practice in England (see issue 10)
- A response-based model structure may be a plausible alternative to modelling the decision problem and allows for some exploration of heterogeneity (see issue 11)
- The immaturity of the data does not allow for meaningful interpretation of extrapolated data (see issue 12)
- Drug wastage costs and adherence to the specified dose should be considered in the costs of larotrectinib (see issue 13)

- Administration costs and resource use are uncertain and difficult to measure and model (see issue 14)
- Implementation and training costs associated with larotrectinib should be considered (see issue 15)
- The post-progression health utility state should be equivalent between larotrectinib and the comparator arm (see issue 16)
- A proportion of the tumour types included in the analysis do not meet the end-of-life criteria. Committee will consider this in its judgement. This is compounded by uncertainty in the positioning of larotrectinib and judgements around tumours types that are unrepresented in the evidence base (see issue 17)
- Larotrectinib is innovative but its adoption is dependent on innovation and development of testing infrastructure (see issue 18)
- Larotrectinib does not meet the criteria for inclusion in the Cancer Drugs Fund, it does not have plausible potential to be cost-effective at the current price that includes a patient access scheme (see Issue 19).

2.2 The technical team recognised that the following uncertainties would remain in the analyses and could not be resolved within the timeframe of this technology appraisal:

- Follow-up of the larotrectinib trials is short and overall and progression-free survival data are immature.

- The clinical evidence is based on a small number of patients with multiple tumour types, meaning comparison of baseline characteristics to a comparator is not possible.

2.3 The cost-effectiveness results include a patient access scheme for larotrectinib. Taking these aspects into account, the technical team's preferred assumptions result in an incremental cost-effectiveness ratio

(ICER) of £97,923 per QALY gained without inclusion of any diagnostic testing costs (see table 8).

2.4 The company considers larotrectinib to meet the end-of-life criteria because the median progression-free survival of the comparator treatments are generally less than 12 months and median overall survival would likely be less than 24 months. Also, the company considers larotrectinib to extend life for longer than 3 months based on the ongoing trial results.

2.5 The company considers the technology to be innovative.

2.6 No equality issues were identified by the company or ERG, although the company considered there to be an equity issue for accessing treatments for rare cancer types (see table 10).

2.7 The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund noted that the 7 Genomic Laboratory Hubs are at different stages of being able to implement Next Generation Sequencing multigene panel testing and this variation will be resolved in the next 1 to 2 years. They also note that whole genome sequencing will take time to embed within clinical treatment pathways, particularly in respect of the need for the collection and processing of fresh tissue.

3. Key issues for consideration

Issue 1 – Prevalence of NTRK gene fusions

<table>
<tr><td>Questions for engagement</td><td>1. Is the distribution of patients in the pooled analysis generalisable to NHS clinical practice?
2. What is the total number of patients who would receive larotrectinib?</td></tr>
<tr><td>Background/description of issue</td><td>Prevalence of NTRK gene fusions is uncertain and varies depending on tumour histology. Some common histologies, such as non-small cell lung cancer, have a low prevalence (<0.1%-3%) while prevalence is higher (>90%) in rarer tumour types such as infantile fibrosarcoma.
The company consider the total population of solid tumours with NTRK gene fusions to be less than 1%. The distribution of patients in the pooled analysis of the trials is considered to be equivalent to the population that would use larotrectinib in clinical practice. The total population of eligible patients was estimated to be [redacted] based on the number of patients receiving last line of cancer therapy for various tumour sites harbouring NTRK fusions.
The ERG calculated estimates of eligible patients with an NTRK fusion within these cancer types. These are shown in Table 2.

Table 2: Prevalence estimates for NTRK fusions, company ‘last line’ eligibility estimates and number needed to screen (adapted from ERG report, Table 2)
<table>
<tr><th>Cancer type</th><th>Cases per year</th><th>Company research % eligible for larotrectinib</th><th>Estimated % of tumours with NTRK fusion</th><th>Cases of NTRK per year</th><th>Eligible and has NTRK fusion</th></tr>
<tr><td>NSCLC cancer patients</td><td>[redacted]</td><td>[redacted]</td><td>0.09%</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Salivary cancer patients (non-MASC)</td><td>[redacted]</td><td>[redacted]</td><td>1.72%</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>MASC cancer patients</td><td>[redacted]</td><td>[redacted]</td><td>100.00%</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Melanoma cancer patients</td><td>[redacted]</td><td>[redacted]</td><td>0.21%</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>CRC cancer patients</td><td>[redacted]</td><td>[redacted]</td><td>0.12%</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Appendix cancer patients (assumed same as pancreatic)</td><td>[redacted]</td><td>[redacted]</td><td>4.00%</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>STS adults (GIST) cancer patients</td><td>[redacted]</td><td>[redacted]</td><td>1.28%</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>STS adults (non-GIST) cancer patients</td><td>[redacted]</td><td>[redacted]</td><td>0.56%</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Bone Sarcoma cancer patients (assumed same as STS adults)</td><td>[redacted]</td><td>[redacted]</td><td>1.00%</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>STS paediatrics cancer patients</td><td>[redacted]</td><td>[redacted]</td><td>0.56%</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>STS infantile sarcoma cancer patients</td><td>[redacted]</td><td>[redacted]</td><td>90.90%</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Secretory breast cancer patients</td><td>[redacted]</td><td>[redacted]</td><td>91.70%</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Cholangiocarcinoma cancer patients</td><td>[redacted]</td><td>[redacted]</td><td>0.10%</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>CNS cancer patients (assumed same as brain)</td><td>[redacted]</td><td>[redacted]</td><td>0.05%</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>CNS paediatric cancer patients</td><td>[redacted]</td><td>[redacted]</td><td>5.30%</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Pancreas cancer patients</td><td>[redacted]</td><td>[redacted]</td><td>0.26%</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Thyroid cancer patients</td><td>[redacted]</td><td>[redacted]</td><td>3.96%</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td></td><td>[redacted]</td><td></td><td></td><td>[redacted]</td><td>[redacted]</td></tr>
</table>
The ERG also note that there are at least 30 tumour sites with identified NTRK fusions, therefore there are tumour sites that are unrepresented within evidence from the trials.</td></tr>
<tr><td>Why this issue is important</td><td>Prevalence estimates of NTRK gene fusions are needed to estimate the eligible populations that could benefit from larotrectinib and for other inputs in the economic model. Using the total prevalence of NTRK gene fusions would affect the distribution of available tumour sites as was shown in the company model. It is also included in the calculation of the number needed to screen and so impacts on the screening costs (see Issue 3).</td></tr>
<tr><td>Technical team preliminary judgement and rationale</td><td>The technical team considered that the distribution of patients in the trial was unlikely to be representative of the histologies that would be treated in clinical practice because the trial population was recruited as a convenience sample and not systematically. It is uncertain what populations would use larotrectinib in clinical practice, but it is more appropriate to assume the distribution of all NTRK fusions than to assume the trial distribution represents NHS clinical practice because there is over-representation of rare tumour types.
The technical team consider that more patients than calculated in Table 2 would receive larotrectinib in clinical practice because of uncertainty in the prevalence estimates, treatment of people before the last line of therapy (see Issue 2) and inclusion of tumour sites not represented in the trials.</td></tr>
<tr><td>Summary of comments</td><td>Comments received from the company:
The company acknowledge that the distribution of tumour types seen in general practice may vary.
Additional tumour types may be identified but the company does not expect that other tumour sites would behave differently, due to the mode of action of larotrectinib. The company will make additional data for tumour types identified in NAVIGATE and SCOUT available if larotrectinib is accepted for us via the CDF.
The company conducted sensitivity analysis by varying the weighting by tumour site for the comparator arm to an alternative distribution. The proportions were weighted using updated NTRK prevalence data from a systematic literature review.</td></tr>
</table>

	• The company acknowledge that the number of patients who may receive larotrectinib is uncertain. • The company provided updated epidemiological data (not shown in Table 1) as part of their response which estimated [redacted] patients in England with NTRK gene fusions. Of these, [redacted] are represented in the clinical evidence or [redacted] of all tumour types. **Comments received from GIST Support UK:** • The total number of GIST cancer patients who might qualify to receive larotrectinib is less than 8 per year. Of 800 diagnoses per year, 125 are wild-type GIST and of these 17 are quadruple wild-type GIST which lack other treatment options. It is within this group that NTRK fusions have been found, 50% of these people will be clear of disease further to surgery and will not require further therapy. **ERG considerations on new evidence received during technical engagement** • The differences between the proportions of patients identified in the systematic review compared with the clinical trial proportions raise substantial concerns as to whether the pooled efficacy is representative of real patients – in particular, over-representation of salivary gland cancers and soft tissue sarcomas and under-representation of breast cancer. • It is unclear how the economic model was updated to reflect the alternative distribution, although applying this distribution to the comparator arm only does not reflect differences that an alternative distribution would have on overall response rates and survival outcomes for the larotrectinib arm.
Technical team judgement after engagement	The technical team do not consider the company sensitivity analysis to be appropriate because the alternative distribution would have further effects on the model that are not explored in this analysis. The current model does not allow for adjustment to alternative populations than the clinical evidence base with potential for substantial selection bias. The technical team maintain that prevalence estimates and number of eligible patients are highly uncertain and consider that this information could be collected in the CDF.

Issue 2 – Treatment pathway and positioning

Questions for engagement	3. How will the term 'satisfactory' be defined in clinical practice in relation to the marketing authorisation? 4. For each tumour type, at what point(s) in the respective treatment pathways will larotrectinib be used in clinical practice? 5. Is the clinical evidence for larotrectinib generalisable to the positioning in the treatment pathways in clinical practice in England?
Background/description of issue	There is no defined clinical pathway specifically for people with solid tumours expressing the *NTRK* gene fusion. Treatment is currently guided by tumour-site specific care guidelines. The anticipated marketing authorisation is "for the treatment of adult and paediatric patients with solid tumours that display NTRK gene fusion, who have a disease that is locally advance, metastatic or where surgical resection is likely to result in severe morbidity, and who have no satisfactory treatment options." This does not specify the point in the treatment pathway that larotrectinib can be used. In the larotrectinib full efficacy evaluable dataset, [redacted] of patients received prior surgery, [redacted] received prior systemic therapy, [redacted] as first or second line and [redacted] as 3rd or further line. **The company** propose that larotrectinib will be used 'last-line' after patients have exhausted all satisfactory treatment options for locally advanced or metastatic disease or where surgical resection is likely to result in severe morbidity as in the marketing authorisation and consider the clinical evidence to align with the marketing authorisation. **The ERG** consider that the positioning of larotrectinib would likely vary considerably by tumour site. The threshold for what treatments are deemed unsatisfactory would involve assessment of response rates, adverse events and discussion with patients. This would therefore mean the threshold would vary by clinicians and patients. There is not enough information provided to understand the definition of 'no further satisfactory treatments' in the trials, including prognostic indicators and other patient characteristics, and whether this is generalisable to UK clinical practice. **The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund** also considers the wording of the company's anticipated marking authorisation to be open to potentially variable interpretation. They consider that larotrectinib was given in the larotrectinib clinical trials before other alternative therapies had been exhausted and note that this could have led to potential considerable bias in the larotrectinib clinical trials.
Why this issue is important	It is unclear which populations will benefit from larotrectinib without a robust definition of satisfactory treatment. The position of larotrectinib within the pathway affects the number of people eligible for larotrectinib (see Issue 1) and the potential comparators (see Issue 8).
Technical team preliminary judgement and rationale	The technical team consider the wording of the proposed marketing authorisation to be ambiguous and this limits the technical team's understanding of where it should be positioned in the treatment pathway. To make a recommendation in line with larotrectinib's marketing authorisation, committee will need to know where larotrectinib will be positioned in each of the treatment pathways for each of the tumour types. It is likely that the decision to use *NTRK* inhibitors would depend on clinician judgement. However, clinical judgement may be difficult to elicit because of the rarity and diversity of these tumour types. It is unclear whether the trial population have exhausted all available treatment options without this information. This could be collected through the Cancer Drugs Fund.

Summary of comments	**Comments received from the company:** • The company considered that each of the 3 trials from which the population of the pooled efficacy evaluable set was drawn reflected the license because of the wording in their respective inclusion criteria. • For patients where response to prior systemic therapy was reported, the ORR to that line of therapy was [redacted] • The company consider that clinical judgement would be used to identify patients who have no satisfactory treatment options. • The real-world interpretation of 'satisfactory' and positioning in the treatment pathway could be collected in the CDF. **Comments received from GIST Support UK:** • Current treatment for GIST involve surgery followed by three lines of TKI therapy. When sequencing becomes more readily available and identification of NTRK fusions happen earlier, patients with NTRK gene fusions should receive the drug earlier in the treatment pathway.
Technical team judgement after engagement	The use of larotrectinib within the treatment pathways has not been established because the decision to use *NTRK* inhibitors would depend on clinician judgement. Section 4.4 of the summary of product characteristics states that larotrectinib should only be used if there are no treatment options for which clinical benefit has been established, or where such treatment options have been exhausted (i.e., no satisfactory treatment options).The technical team consider that there is potential for the definition of 'satisfactory' to change if larotrectinib is used in clinical practice; the magnitude and direction of bias of these changes are unknown.

Issue 3 – NTRK gene fusion testing

Questions for engagement	6. What is the likely screening pathway to identify *NTRK* fusion positive solid tumours? 7. At what point in the treatment pathway for each tumour type will *NTRK* gene fusion testing be carried out? 8. What proportion of the screening costs should be included in the economic model for this appraisal? Or is it appropriate to exclude all testing costs from the analyses?
Background/description of issue	All solid tumours types can potentially harbor an *NTRK* gene fusions. This means that the number of people who require testing for *NTRK* gene fusions is very high. A national service for cancer genomic testing has been created by NHS England and is regionally organised by 7 Genomic Laboratory Hubs. The hubs process tissue samples for whole genome sequencing (WGS) and pass these to Genomics England for analysis, perform next generation sequencing (NGS) and interpret all NGS and WGS results before returning them to the requesting clinician. NGS provides the technology for multigene panels (which provide testing for anything between 5 and 500 genes). **The company** consider that diagnostic testing for *NTRK* gene fusions is part of the NHS Long Term Plan and that whole genome sequencing (WGS) will be offered as part of routine care. Therefore, they have not provided any testing strategies or costs associated with testing. **The ERG** consider that because only patients with *NTRK* fusion can benefit from larotrectinib, considering the clinical impact of genetic testing is important. The ERG identify different testing pathways for different prevalence of *NTRK* fusions. They provide a scenario including a mean cost per patient for testing that includes testing for the total number needed to screen per patient, based on the most appropriate screening pathway for each tumour site. **The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund** highlight that for paediatric cancer and sarcoma, funding for WGS is in place. However, it is noted that NGS may be necessary for *NTRK* fusion testing in the short term until WGS is fully operational. Funding is also currently in place for MASC and the secretory variant of breast cancer through the National Genomic Test Directory for 2019. For all other adult solid cancers, *NTRK* gene fusion testing is not currently required by the National Genomic Test Directory and is not systematically performed. However, by the end of the 2019/20 financial year, the Genomic Laboratory Hubs plan to introduce NGS gene panels for solid tumour testing, which will include the capability to identify *NTRK* gene fusions. This could be for example with a 50 to 60 gene panel or a 500 gene panel. Uptake of molecular testing across the 7 genomic hubs will increase during 2020/21 as genomic pathways are embedded and links are made with the clinical teams. Given the complexity of implementation, it may take a further 12 months for molecular testing to become fully embedded in practice. The clinical lead notes that for patients and clinicians to be able to best use the information of NGS panel testing, it has to be done prior to starting any systemic therapy for the locally advanced/metastatic disease. The clinical lead considers that it is appropriate that at least part of the costs for multi-gene panel testing be covered by each company that benefits from the new service provision. They highlight that the weighted average cost of testing will be sensitive to the prevalence of *NTRK* gene fusions in each tumour type included in the calculation (see Issue 1). NHS England and NHS Improvement would like to see scenario analyses in which various percentages of the costs of NGS multi-gene panel testing are borne by the company: 100%, 50%, 33% and 0%.
Why this issue is important	The number of people who will be tested for *NTRK* gene fusions will be high and also costly. The screening pathway is currently uncertain. There is a potential equality issue as service provisional has not yet been rolled out nationally.
Technical team preliminary judgement and rationale	The technical team understands the difficulties in describing the potential screening pathway and the associated costs. For *NTRK* fusions currently tested in clinical practice, the costs should not be included in the economic analysis. However, because there will likely be a large number of NGS

	tests that are not be available in clinical practice, the company should consider how to incorporate the costs of this testing in a scenario. The technical team considers the screening pathway to depend entirely on the provisions set up by NHS England in a timeframe that aligns with the *NTRK* appraisals. The proportion of overall testing costs that should be included in the analysis for the larotrectinib appraisal is not a judgement that the technical team can make at this stage without more input from the company and NHS England. However, the technical team would like to see the scenario analyses in which various percentages of the costs of NGS multi-gene panel testing are borne by the company: 100%, 50%, 33% and 0%, as suggested by The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund.
Summary of comments	**Comments received from the company:** • The company consider that WGS would be used for all paediatric and sarcoma cancers and NGS will be used for all other solid tumours. • The company do not consider screening costs should be included because other actionable targets may be identified through these screening tests. • The company provided scenario analyses including diagnostic tests for a small number of last-line patients only, based on the number eligible for last line treatment. **Comments received from GIST Support UK:** • The standard screening pathway is: 1. Molecular testing to review KIT and PDGFRA mutations 2. SDHB IHC testing for all wild-type GIST 3. SDHB IHC reliably identifies GIST with SDH complex deficiency 4. SDHA IHC testing 5. BRAF is often performed as part of NGS multi-gene panel 6. NF1, skeinoid fibers and clinical input 7. Expected place in the pathway for NTRK through IHC • The point of testing will vary by patient – this will be part of whole genome sequencing in GIST and sarcoma and therefore testing costs should be excluded. **ERG considerations on new evidence received during technical engagement** • The approach taken to adding testing costs is different to the approach taken by the ERG. The ERG cannot verify how the cost of [redacted] per test was calculated in the company analyses.
Technical team judgement after engagement	The technical team is anticipating receiving further guidance on the diagnostic testing pathway from NHS England and Genomics England ahead of the committee meeting. The rationale for the company testing strategy and how it is incorporated in the model is not clear. The committee will make the final judgement about what diagnostic costs should be attributed to larotrectinib at the committee meeting.

Issue 4 – Identification of NTRK gene fusions – diagnostic accuracy

Questions for engagement	9. What is the expected diagnostic accuracy of NGS testing? 10. Is there a testing sequence that could avoid a substantial number of false positive results in low *NTRK* fusion-positive tumour types? 11. Is it appropriate to limit testing to avoid false positive results and the associated costs?
Background/description of issue	In addition to issues identifying the prevalence (see Issue 1) and uncertainty about the testing pathway (Issue 3), there is uncertainty about the ability to correctly identify *NTRK* fusions through diagnostic accuracy. **The company** does not explore issues of diagnostic accuracy. **The ERG** notes that for tumour sites with low *NTRK* fusion prevalence, the number of false positives (people who test positive for *NTRK* but do not have it, measured by specificity of the diagnostic test) may outnumber the true *NTRK* fusion cases. This will substantially reduce the observed effectiveness of larotrectinib because these people will not respond to it. The ERG conducted analysis to explore the impact of false positives using a test with 99% sensitivity and 99% specificity shown in Table 3 below. **Table 3: Accuracy of diagnostic testing exploratory analysis (adapted from ERG report, table 4)**

Cancer type	True positives (has *NTRK* fusion detected)	False positives (no fusion but test is positive)	% of treated who are false positive	% of treated who respond (based on 80% response)
NSCLC cancer patients	[redacted]	[redacted]	[redacted]	[redacted]
Salivary cancer patients (non-MASC)	[redacted]	[redacted]	[redacted]	[redacted]
MASC cancer patients	[redacted]	[redacted]	[redacted]	[redacted]
Melanoma cancer patients	[redacted]	[redacted]	[redacted]	[redacted]
CRC cancer patients	[redacted]	[redacted]	[redacted]	[redacted]
Appendix cancer patients (assumed same as pancreatic)	[redacted]	[redacted]	[redacted]	[redacted]

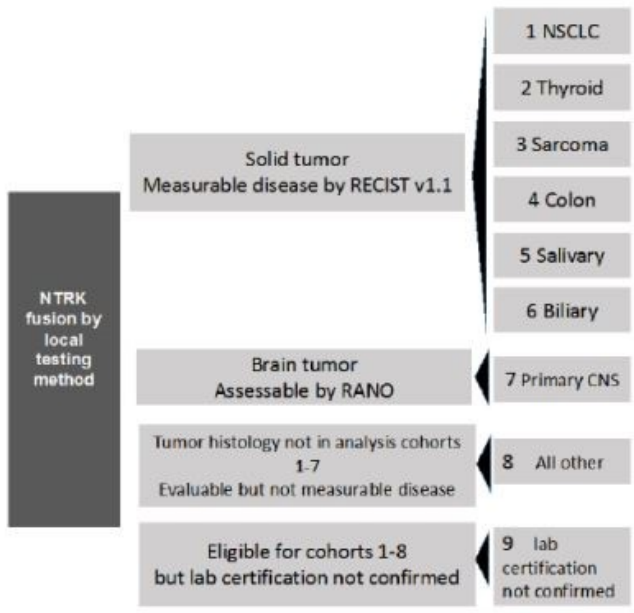
	STS adults (GIST) cancer patients [redacted] STS adults (non-GIST) cancer patients [redacted] Bone Sarcoma cancer patients (assumed same as STS adults) [redacted] STS paediatrics cancer patients [redacted] STS infantile sarcoma cancer patients [redacted] Secretory breast cancer patients [redacted] Cholangiocarcinoma cancer patients [redacted] CNS cancer patients (assumed same as brain) [redacted] CNS paediatric cancer patients [redacted] Pancreas cancer patients [redacted] Thyroid cancer patients [redacted] [redacted] **Total – all tumour sites** [redacted] **Total – sites with high *NTRK* fusion rate or paediatric cases** [redacted] This analysis shows that with highly specific tests of 99%, the low prevalence of *NTRK* in some tumour types would mean the response to larotrectinib could be substantially lowered by false positive tests. The ERG consider testing of sites with high *NTRK* fusion rates or paediatric cases could lead to better overall response rates with fewer false positives.
Why this issue is important	The issue of diagnostic accuracy may explain some of the heterogeneity of response across different solid tumour types (see Issue 6), tumour types with low *NTRK* fusion prevalence are more likely to have a false positive *NTRK* fusion test, in which case response would not be expected. Diagnostic accuracy will also affect the practicalities of implementing tests to identify eligible patients. Additionally, there are ethical considerations to giving unnecessary treatments to false positive patients.
Technical team preliminary judgement and rationale	There is uncertainty about the accuracy of testing for *NTRK* fusion testing. Although the estimates in Table 3 may be inaccurate because of the treatment pathway uncertainty and prevalence estimates uncertainty (Issue 2), the technical team would like an estimate of the sensitivity and specificity of NGS testing of *NTRK* fusions as there is a likely to be a high level of uncertainty about the practicality of testing in low prevalence tumour types, even with tests of high sensitivity and specificity. The technical team recognise that false-positive rates will impact the clinical efficacy and recognise the potential of additional uncertainty if patients falsely identified as *NTRK*-fusion positive were included in the trial population.
Summary of comments	**Comments received from the company:** • The company do not consider WGS and broad NGS panels to be appropriate for screening for NTRK fusions. The intention of the National Genomic Medicine service is to generate the relevant genomic profile for the patient to inform diagnosis, staging and treatment. • The company does not consider it appropriate to limit testing to particular tumour types. **ERG considerations on new evidence received during technical engagement** • The purpose of the National Genomic Medicine Service is not relevant. It is essential that the accuracy of testing for NTRK fusion is known to ascertain likely rate of false-positives (in whom larotrectinib cannot be effective). • In response to a comment indicating that specificity of detecting *NTRK* fusion with RNA-based NGS may be as high as 99.9%, the ERG provided a table that indicated the number of false-positives. There remain several tumour sites where false-positive cases make up a sizeable proportion of all patients. • The ERG also notes that it is unlikely that 99.9% specificity can be achieved. It is unclear whether all patients across all tumour sites will receive NGS testing and human error, sample deterioration or contamination could all reduce the effective specificity.
Technical team judgement after engagement	The technical team consider that the screening method will require implausibly high specificity to achieve an acceptable diagnostic accuracy for common tumour types with low *NTRK* prevalence. The diagnostic accuracy may vary by *NTRK* gene fusion and not all fusion partners have been identified or characterised, therefore it is difficult for committee to make judgement in the absence of information.

Issue 5 – Primary CNS tumours

Questions for engagement	12. Should patients with primary CNS tumours be included in the analysis?
Background/description of issue	The anticipated marketing authorisation for larotrectinib includes people with primary CNS tumours. The company's presented separate clinical outcomes for those with primary CNS tumours and all other tumour types pooled, but the economic analysis included both datasets. **The company** present two efficacy evaluable sets, with the 9 patients with primary CNS tumours coded as the SAS3 dataset and all other indications coded as ePAS2. The entire efficacy evaluable dataset was included in the economic analysis. **The ERG** noted that the overall response rate for SAS3 dataset was notably lower than ePAS2 (11% vs 72%) and had a high proportion of *NTRK*2 patients (7 out of 9). The ERG considered that the large difference may be for a variety of reasons: because primary CNS tumours may be less likely to respond to larotrectinib; *NTRK*2 patients may be less likely to respond to larotrectinib; *NTRK*2 patients may have a higher false positive rate (see Issue 4); or that it was a chance result.

	The ERG considered this issue as evidence of heterogeneity between subgroups (see Issue 7). They presented results separately for the ePAS2 dataset and the full efficacy evaluable dataset.
Why this issue is important	Including primary CNS patients in the base case increases the generalisability of the evidence base to the population likely to be seen in clinical practice. Including primary CNS patients increases the company's base case ICER and also reduces the Bayesian Hierarchical Model ORR to 57% which affects the ERG responder analysis (see Issue 12).
Technical team preliminary judgement and rationale	The technical team consider it appropriate to include primary CNS tumours in the base case as it increases generalisability and primary CNS tumours are included in the marketing authorisation.
Summary of comments	**Comments received from the company:** • Exclusion of primary CNS patients in the overall efficacy estimates was based on the following rationale: - response from patients with primary CNS tumours were evaluated using RANO or RECIST v1.1, whereas other solid tumours were evaluated using RECIST v1.1 only - surgery and radiation treatments can lead to a varying amount of oedema/ inflammation/ scarring which can impact radiological assessment in these patients - the response was investigator assessed instead of independent review committee assessed.
Technical team judgement after engagement	The technical team consider that there are biologically plausible reasons why larotrectinib would have a true lower response for primary CNS tumours. For example, the EPAR states larotrectinib is a substrate for P-glycoprotein, a major functional constituent of the blood brain barrier. This may reduce the effective dose within the brain. Ideally, the CNS tumours would be modelled separately and avoid using response data because of the difficulty in assessing response. However, in the absence of evidence it is appropriate to include the response data in the modelling as it provides the most conservative ORR and increases generalisability because primary CNS tumours are included in the marketing authorisation.

Issue 6 – Trial study design

Questions for engagement	13. Is it appropriate to consider the 'basket' trial design for statistical evidence of heterogeneity?
Background/description of issue	The company present a pooled analysis of 3 trials – a phase 1 study that contributed [redacted] patients, SCOUT a phase 1/2 study that contributed [redacted] paediatric patients and NAVIGATE a phase 2 basket trial that contributed [redacted] adult patients to the pooled analysis. A basket trial recruits patients to assess efficacy in distinct cancer sites and if a prespecified proportion of patients in a particular basket respond, then recruitment is expanded within this disease area. If too few responses are observed within a basket, then recruitment is stopped due to low promise of efficacy. The individual baskets are described in the protocol and summarised in Figure 3. **Figure 3 – Basket design of the NAVIGATE basket trial (from the NAVIGATE protocol)**  **The company** do not consider the study design in their analysis and present each patient in their analysis independent of the trial they initiated in, assuming homogeneity of response across all tumour types. **The ERG** consider heterogeneity of response an important consideration in the approach of the basket trial (see Issue 7). The poor response rate of patients with primary CNS tumours (1/9) suggests that the assumption of homogeneity is not appropriate, and this would be shown by the results of the statistical tests in the trial.
Why this issue is important	Homogeneity of treatment effect is an assumption of the 'histology independent' nature of the appraisal and the assumptions in the economic analysis.
Technical team preliminary judgement and rationale	The technical team consider that as the basket trial design has not completed, following the statistical protocol from the basket trial would offer more robust data on heterogeneity that could better inform the cost-effectiveness analysis. The technical team would like an update on the status of each basket within the NAVIGATE trial protocol, and statistical evidence of heterogeneity of response.

Summary of comments	**Comments received from the company:** • The studies in the development programme for larotrectinib were not designed to analyse each cohort separately. • Assessing efficacy by 'basket' is statistically inappropriate, the rarity of disease prevents a standard approach to heterogeneity. **ERG considerations on new evidence received during technical engagement** • Given the wide diversity in tumour types across adults and children, with variation in NTRK fusion prevalence, the ERG considers that the *potential* for across-site heterogeneity must be considered.
Technical team judgement after engagement	The technical team consider that although heterogeneity of response in the basket trial was in the initial protocol, the analyses are now pooled which likely introduces selection bias. The updated company trial protocol assumes a response for all tumour types and this is not formally tested for multiple tumour types within the pooled analysis.

Issue 7 – Heterogeneity of response across different solid tumour types

Questions for engagement	14. Is a homogeneous response to larotrectinib across different tumour types a reasonable assumption? 15. Is the Bayesian Hierarchical Modelling (BHM) framework an appropriate method to capture the heterogeneity of response across different solid tumour types for this appraisal? 16. Would it be appropriate to apply the BHM framework to explore the heterogeneity in the time to event outcomes?
Background/description of issue	A key assumption of the appraisal is that the response to larotrectinib is homogeneous across tumour types. **The company** assume that each of the solid tumour types will have identical response rates when treated with larotrectinib. This allows them to generate a pooled response estimate across each of the tumour types included in their efficacy evaluable dataset, 72%. This approach does not take into account the potential for heterogeneity of response across different tumour types. The company have not explored alternatives to this assumption, stating that [redacted] **The ERG** considered the issue of heterogeneity to be very important in this appraisal, highlighting that the target population (see Issue 1) will respond differently if the assumption is invalid and there is any difference in population to the trial. Additionally, making a recommendation based on a modelling approach that groups heterogeneous subgroups together may result in cost-effective subgroups being shielded by the cost-ineffective groups and vice-versa. The ERG identified the most important sources of heterogeneity to be the response by tumour types studied within the trial, the testing costs by tumour type, age of patients (paediatric/adults) and the heterogeneity of trial design. In order to explore the heterogeneity of response between tumour types, the ERG performed exploratory analysis modelling each of the tumour types as a 'basket' and analysed the response data using a Bayesian Hierarchical Modelling framework. The ERG used this method to estimate posterior probabilities of response for each tumour type, as well as a pooled posterior probability of response across all tumour types, accounting for the potential lack of uniformity of effect across tumours. The predicted outcomes of the model are summarised in Figure 4, grouped by prevalence. The overall distribution for an unevaluated tumour site is shown in Figure 5. **Figure 4 - Predicted response rate distributions from Bayesian Hierarchical model**

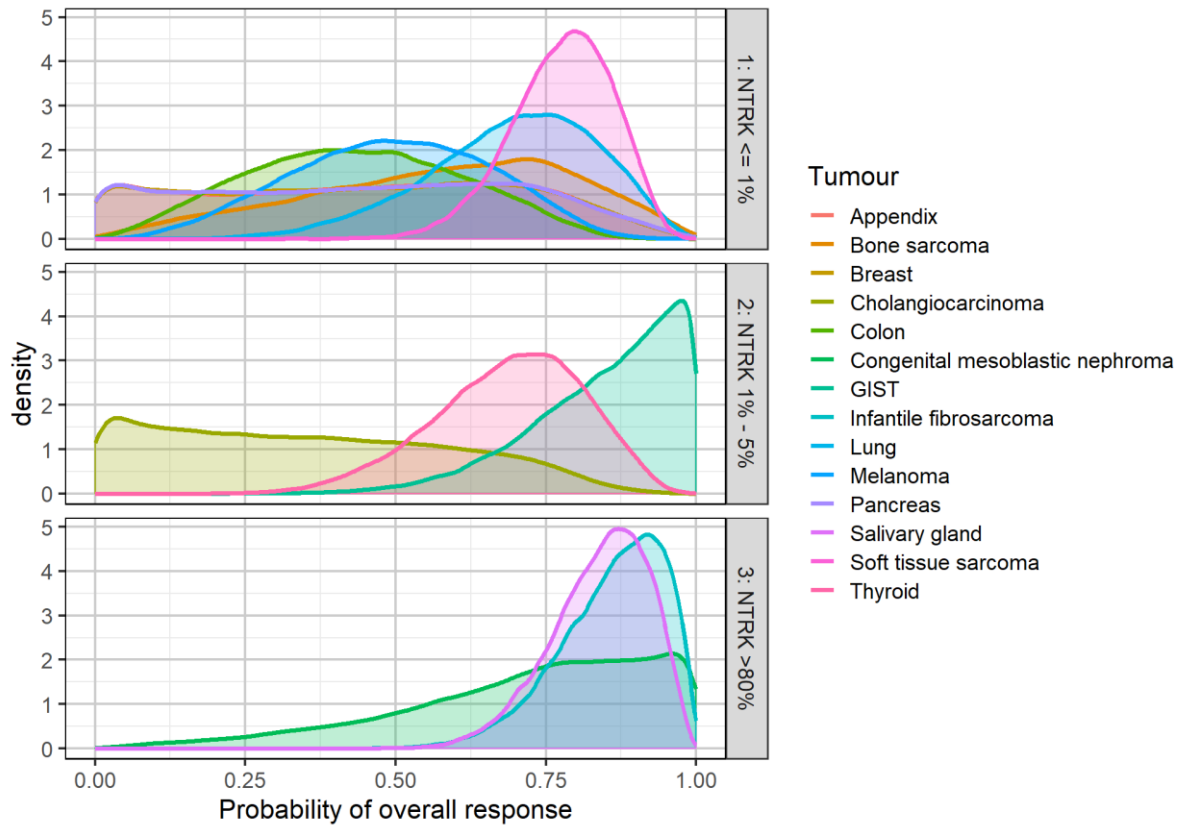
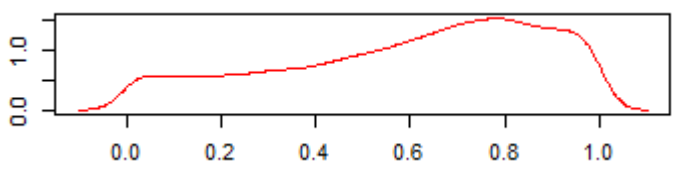


Figure 5 – Predicted distribution of response of an unevaluated tumour site



The ERG also requested time-to-event data on progression free survival and overall survival to further assess the heterogeneity of these outcomes, as they are used in the model, but the company refused to provide it. Therefore, the ERG simulated PFS using the data from response rates and the Kaplan-Meier curves for the response-based model. The ERG noted the weaknesses of this analysis and would prefer unsimulated data but considered it important to explore heterogeneity. The predicted results of the improvement in PFS of each of the histologies is shown in Figure 6 below. This analysis was repeated with the OS data, although the ERG acknowledged that the small number of events makes the analyses difficult to interpret.

Figure 6 - Potential improvement in PFS for larotrectinib versus comparator, by tumour site

[redacted]

Why this issue is important	If benefits differ by histology or other characteristics, then estimates of treatment benefit must match the target population. Also, there can be health benefits from making tailored decisions for particular groups of patients such as a potentially cost-effective treatments being withheld because of cost ineffective subgroups.
Technical team preliminary judgement and rationale	The technical team consider that patients are unlikely to have a homogeneous response across tumour types, as the trial design has not tested this (see Issue 6), there is evidence of heterogeneous response and TKIs generally have heterogeneity across tumour sites. The technical team consider that using the simulated progression-free survival data in the BHM framework is a reasonable approach, and strongly suggests important differences in progression-

	free survival between tumour type populations. Data should be made available for an appropriate assessment of heterogeneity of response and time-to-event outcomes and the risk to the NHS associated with this uncertainty of heterogeneity should be modelled. The technical team would like to see additional descriptive data, for example – median, 6 month and 12 month PFS and OS data (and unsimulated full Kaplan-Meier curves where feasible) for each tumour type and separately for paediatric patients vs adults or by clinical trial informing the integrated efficacy analysis. This will help determine where there appear to be similarities and differences in PFS and OS, determine whether hierarchical modelling is appropriate and whether pooling the clinical trials into one efficacy evaluable dataset/ by age group is appropriate.
Summary of comments	**Comments received from the company:** • The company understand the histology-independent nature of larotrectinib and the rarity of NTRK gene fusion cancers present unorthodox challenges to the traditional technology assessment process. • The company believe *"that consideration of response by tumour location only serves as a distraction and introduces the potential for decision-making to be based on chance findings."* • The ESMO Scale for Clinical Actionability of molecular Targets (ESCAT) graded NTRK gene fusions as a Tier 1c actionable driver across tumours. • The clinical and non-clinical body of evidence support a histology-independent indication since larotrectinib has a "*large magnitude of effect irrespective of tumour site*." • The company did not agree to provide time-to-event data by tumour types that the ERG intended to use with the BHM model because: - The company consider that the data are too limited in patient numbers - They consider that Bayesian methods are theoretical for this decision problem and would require extensive amounts of external data to populate the model - The company could not find examples of a Bayesian framework for time-to-event outcomes and it would therefore exploration using a Bayesian framework would be academic and not appropriate for a technology appraisal - The company are discussing post-marketing commitments with the EMA that will provide more substantial basis to assess tumour heterogeneity • The BHM methodology consistently shows a lower ORR than the trial data irrespective of the chosen prior, therefore the company consider this methodology inappropriate. The company suggest that unknown tumour types should be validated with external sources of information such as clinical opinion. • The response-based model requires surrogacy between response and survival but there is no evidence to support this when considering heterogeneity in time to event outcomes. • The company provided additional data from a new data cut from February 2019 with investigator assessed response rates for an expanded cohort of 153 efficacy-evaluable patients. **ERG considerations on new evidence received during technical engagement** • The ERG disagrees with the claim that larotrectinib has demonstrated a "*large magnitude of effects irrespective of tumour site*" as the BHM shows that this is not the case. • The BHM methodology was developed specifically for the analysis of basket trials. • Bayesian modelling is particularly useful where data are limited because it permits 'borrowing strength' from tumour sites where data is more plentiful. Therefore, it is not the case that Bayesian models become more suitable as data accumulate, the opposite is true. • The BHM models all tumour sites and all data in one model, assuming some exchangeability across sites, not separate analyses. • Given the novelty of tumour-agnostic interventions and basket trials, statistical analyses of such interventions must be innovative, particularly when considering survival analysis. • The BHM is not intended to replicate the ORR provided by the company, but to assess whether the ORR was robust to possible heterogeneity – the ERG found it was not robust. • The assumption of surrogacy was required because the company did not supply any data on survival outcomes by tumour site, an appropriate Bayesian survival model could be fitted if this data were made available. • The ERG noted that the new patients within the expanded cohort mostly comprised of patients in tumour sites that were already most numerous in the previous data cut ([redacted]). There were also [redacted] new tumour sites – [redacted] • The ERG used the same BHM model to assess heterogeneity applied to the new data which predicted a response of 72%. The ERG considered this to likely be a consequence of difference between investigator and independent review committee assessment of response.
Technical team judgement after engagement	The technical team consider consideration of heterogeneity by tumour type to be an important uncertainty within this appraisal. The EPAR states "*There are insufficient data to establish the activity of larotrectinib due to lack of comprehensive sequencing of tumour tissue prior to treatment, the small sample size in different tumour types, the significant heterogeneity observed in terms of ORR coupled with the notably very low ORR observed in different tumour types (ORR=0%-33%), especially in those common tumour types where occurrence of NTRK gene fusion is rare (lung, colon, breast)."* The technical team also consider the BHM appropriate for characterising heterogeneity of response by tumour type. The response data available do not support homogeneity of response and the BHM provides a framework to understand this uncertainty by reducing the potential effect of selection bias from high response rates in tumour types that have more patient numbers.

Issue 8 – Constructing a comparator arm

Questions for engagement	17. Is the company's comparator arm suitable for decision making? 18. Is it appropriate to use non-responders as a proxy for patients not having an active treatment or previous line of therapy to generate a comparator arm for this appraisal?
Background/description of issue	The larotrectinib trials did not include a control arm. There is difficulty in creating a basket comparator for established management without larotrectinib because of the diversity of potential comparators. **The company** generated a comparator arm by conducting a systematic literature review of tumours sites known to harbor *NTRK* gene fusions and selecting comparator data, the majority of which is from previous NICE TA guidance, and best supportive care or placebo arms were used from trials where NICE TA guidance was not available. The PFS and OS curves identified in the submissions for each appraisal were extracted, along with utility values, adverse event data and costs to create an 'engine' for each comparator in the economic model (see Issue 11). The identified comparator engines were weighted by patient enrolment as in the efficacy evaluable dataset to create a theoretical comparator arm. The company use this data in their base case cost-effectiveness analysis but also provided scenarios: • Using effectiveness data on non-responders as a proxy for patients not receiving an active treatment. • Comparing the outcomes for people on larotrectinib with their outcomes on the previous line of therapy (using a ratio of time to progression for each patient) **The ERG** considered that the validity of the comparator data is uncertain because it is drawn from arbitrarily selected past NICE appraisals, with numerous assumptions where past TAs did not exist. The main issue with this data is the inability to compare the data with the corresponding data for larotrectinib, since the company did not provide PFS or OS results by tumour type (see Issue 7). Other unresolvable issues with this approach included arbitrary grouping of tumour types, *NTRK* may be a prognostic factor (see table 10) and this was not accounted for in the analysis and the inability to compare baseline characteristics. The ERG preferred to use the responder-based analysis, this is explored in Issue 11. The ERG also considered that the patients without *NTRK* fusions would not be able to receive targeted experimental therapy (see Issue 10).
Why this issue is important	A lack of direct evidence adds uncertainty to the true comparative efficacy of larotrectinib and established management. The lack of control group in the larotrectinib trial evidence means that the relative effectiveness and safety of larotrectinib compared with relevant alternative cancer therapies are highly uncertain.
Technical team preliminary judgement and rationale	The technical team recognises the difficulty in constructing a comparator arm for this appraisal and consider multiple methods to be valuable in estimating a counterfactual. This is because methods for creating counterfactual data each have inherent biases and if the methods concur then committee can consider this in their exploration of uncertainty.
Summary of comments	**Comments received from the company:** • The company considered that process and scrutiny undertaken in each technology appraisal to select the committee's preferred inputs would suggest that their comparator arm is the most suitable for decision making. • The approach to comparator selection was validated with clinical experts and alternative comparator selection has minimal impact on the ICER
Technical team judgement after engagement	The technical team maintain that multiple methods of constructing counterfactual data would be useful in exploration of uncertainty for this appraisal.

Issue 9 – Comparator treatments

Questions for engagement	19. Are the comparators identified representative of where larotrectinib would be used in the treatment pathways?
Background/description of issue	It is unclear if the comparator dataset is drawn from a population at a line of therapy similar to the clinical trials or the population that would be used in clinical practice (see Issue 2). Without information on the line of treatment in the trial, it is difficult to determine whether the comparator choices are appropriate. **The company** selected comparators based on the most advanced patients (e.g. last line of systemic therapy), the extent to which the publication matched the treatment criteria in the larotrectinib trials, the acceptance of the comparator arm as a proxy of BSC and the date of publication. The chosen comparators are summarised in Table 4. **Table 4 – Comparator treatment by tumour site (adapted from Table 14, ERG report)**

Tumour Site	Comparator treatment	Data source
NSCLC	Best Supportive Care	TA374
Salivary gland	Cisplatin + vinorelbine	Airoldi 2001- survival outcomes and AE rates
Melanoma	Mixed chemotherapy including dacarbazine, paclitaxel, carboplatin, temozolomide and paclitaxel + carboplatin	TA357
CRC and appendix	Best Supportive Care	TA405
GIST	Best Supportive Care	TA488
Adult STS (non-GIST) and bone sarcoma	Best Supportive Care	TA185

	<table><tr><td>STS (paediatric), IFS and CMN</td><td>Irinotecan and vincristine</td><td>Mascarenhas et al 2010</td></tr><tr><td>Breast</td><td>Treatment of physician's choice including vinorelbine, gemcitabine, paclitaxel, doxorubicin and docetaxel</td><td>TA423</td></tr><tr><td>Cholangiocarcinoma</td><td>Gemcitabine + cisplatin</td><td>Valle et al 2010</td></tr><tr><td>Pancreatic</td><td>5-FU + leucovorin</td><td>TA440</td></tr><tr><td>Gliomas (CNS)</td><td>Lomustine</td><td>Batchelor et al 2010</td></tr><tr><td>Thyroid</td><td>Best Supportive care</td><td>TA535</td></tr></table> **The ERG** preferred to consider the responder-based analysis due to uncertainty in the analysis and did not consider further analysis of this comparator data, either across tumour or within site. **The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund** considers some of the comparators used in the company's blended comparator to not reflect the treatments that larotrectinib would displace in clinical practice, including some patients that would have received chemotherapy.
Why this issue is important	Line of therapy is an important determinant of prognosis. Selecting the most appropriate comparator treatments is important as it also impacts on the costs and utility values included in the economic model for the comparator arm.
Technical team preliminary judgement and rationale	The technical team acknowledge the difficulty in selecting comparator data. Although the larotrectinib trial population is mostly pre-treated, the technical team consider estimates of PFS and OS for the comparator may be confounded in favour of larotrectinib for some tumour types. However, this is difficult to confirm without further information about trial participants' treatment history and other factors that would affect treatment choice. The technical team would like to see line of therapy for larotrectinib by tumour type if available.
Summary of comments	**Comments received from the company:** • The company restated their approach to comparator selection and their response to Issue 8. All data chosen for the weighted comparator arm reported patients had received previous lines of therapy.
Technical team judgement after engagement	The technical team maintain that selection of the most appropriate comparators is important and this is difficult without knowledge of the treatment pathway and positioning (Issue 2). The company have not provided clinical data to support each line of therapy within the trial so there is considerable uncertainty as to the generalisability of the trial results.

Issue 10 – Subsequent therapies

Questions for engagement	20. What subsequent treatments would be expected in clinical practice after larotrectinib? 21. Should experimental treatments be adjusted for in this analysis?
Background/description of issue	No subsequent treatments are modelled for larotrectinib in the company base case. In the trials, [redacted] ([redacted]) progressed patients received an experimental therapy called 'LOXO-195', developed for patients who become resistant to TRK inhibitors. [redacted] ([redacted]) patients continued to receive larotrectinib post-progression. [redacted] patients with [redacted] also received surgery following a partial response to larotrectinib. **The company** did not adjust for any subsequent treatments. **The ERG** acknowledges that some people with resistant tumours could receive experimental treatments in the NHS, but [redacted] specifically receiving LOXO-195 suggests that the impact of LOXO-195 on survival after progression should not be ignored. They did not consider that larotrectinib would be used after progression. The modelled utility gain in each state is shown in Table 5. **Table 5 – Modelled utility gain in each state** <table><tr><th>Modelled Utility Gain</th><th>Larotrectinib arm</th><th>Comparator arm</th></tr><tr><td>Pre-progression</td><td>[redacted]</td><td>[redacted]</td></tr><tr><td>Post-progression</td><td>[redacted]</td><td>[redacted]</td></tr><tr><td>Adverse event disutility</td><td>[redacted]</td><td>[redacted]</td></tr><tr><td>Total QALYs</td><td>[redacted]</td><td>[redacted]</td></tr></table> The ERG considered the post-progression survival gains to be confounded by LOXO-195 and larotrectinib use. They provided two scenarios to further explore the uncertainty surrounding post-progression survival. The first scenario uses the ERG base case assumptions but also assumes that the mean discounted post-progression survival is equal to the comparator arm, and also equalises costs and QALYs accrued in the progressed disease. The second scenario assumes that the mean discounted post-progression survival for larotrectinib is equal to the overall survival in the comparator arm. Both scenarios significantly increase the ICER.
Why this issue is important	Subsequent treatments likely to be received following progression within the NHS should be accounted for and modelled appropriately. This can have a significant effect on cost-effectiveness estimates. Greater health gains and increased costs can be expected from treatment with an active subsequent therapy compared with best supportive care. Including greater health gains associated

	with active subsequent therapy, but not including the increased costs for these active therapies in the larotrectinib arm could bias the cost-effectiveness result in favour of larotrectinib.
Technical team preliminary judgement and rationale	The majority of the QALY gains occur in the post-progression state which is clinically implausible. Although crude scenarios, both scenarios produced by the ERG are more plausible than the company base case. Larotrectinib may be used post-progression and the associated costs should be included in the economic analysis, with clinical validation of the length of time larotrectinib would be continued. Costs and benefits of surgery should be included in the economic analysis. The technical team does not consider that experimental treatments would be used in clinical practice and consider that use of LOXO-195 would bias the results and needs to be adjusted for.
Summary of comments	**Comments received from the company:** • The company considers that the only potentially active treatment a patient would receive in clinical practice would be via clinical trials, compassionate use programmes or individual funding requests, none of which would be approved in clinical practice. • Post-progression management could feasibly be collected within the CDF. • The company consider it unreasonable to assume that any additional post-progression benefit comes from experimental treatment for which there is no data. • There is no evidence of any drug demonstrating effectiveness in the post-progression TRK inhibitor setting, therefore equalising benefit in the post-progression setting would be overly conservative and unreasonable. • After reviewing previous NICE technology appraisals in a 'last line' setting, the company could not find criticism for exclusion of experimental therapies in the modelling that may have been received post-progression. • Experimental therapies could equally be offered to patients who had received stand of care or larotrectinib and therefore, this would have minimal impact on the cost-effectiveness results. • The company provided clarification on patient numbers who received subsequent therapies, reporting ([redacted]) receiving larotrectinib and ([redacted]) patients receiving LOXO-195 post progression. • The company provided OS KM curves for patients that received these treatments against those that did not. These showed that patients that received these treatments died sooner than the rest of the cohort. • At the technical engagement teleconference, the company mentioned that some patients may receive an increased dose upon progression and that this may account for the post-progression survival. This was clarified with individual patient data showing [redacted] adult patient receiving an increased dose of larotrectinib and [redacted] paediatric patients receiving dose escalation through natural growth. • The company also considered that because of post-treatment surgery, it may be appropriate to model a cure which may improve the ICER but cannot be determined due to short follow-up, limited sample size and high censoring of data. **Comments received from GIST Support UK:** • LOXO-195 has been developed as a companion drug for development of resistance mutations on larotrectinib. Other NTRK gene fusion inhibiting drugs could be considered after larotrectinib and LOXO-195 in England. **ERG considerations on new evidence received during technical engagement:** • NAVIGATE and SCOUT were single-arm trials, so experimental treatments were unlikely to have been received by comparator patients outside these trials. • LOXO-195 can only be given to patients who have developed resistance to larotrectinib, so the ERG do not agree that it could be offered to patients who receive standard of care. • Given these issues, the ERG consider that adjustment for LOXO-195 is essential. • The ERG noted that the KM curves for patients who continued to receive larotrectinib or LOXO-195 after progression and those who do not appear broadly similar. The unadjusted survival curves may be valid prediction of post-progression survival or indicate that experimental treatments improve survival if patients that progress have below-average survival prognoses.
Technical team judgement after engagement	The technical team consider that post-progression survival is implausibly high, with no biologically plausible reasoning provided by the company. It is unclear if this is an artefact of the extrapolation (see Issue 12) or due to post-progression treatments. The company have suggested a potential cure model for patients that receive post-treatment surgery (for IFS and paediatric sarcomas). The technical team consider that it is inappropriate to model these patients in the same model as all other tumour types as the natural history of the disease is substantially different and that a different model structure is essential to model this. In the company base case, the current comparator engine for all paediatric sarcomas does not model any lifelong survival (5 year OS: [redacted], 10 year OS: [redacted]). However, patients may have received mutilating or disfiguring surgery which would produce a 'cure' with a lower post-surgery utility value. The technical team consider that post-treatment surgery has not been modelled appropriately. The technical team consider that the additional cost of post-progression larotrectinib should be modelled in the base case. The costs and benefits of LOXO-195 are unknown and cannot be accounted for in the model. The benefits of surgery have been inappropriately modelled which introduces significant bias in favour of larotrectinib.

Issue 11 – Model structure

Questions for engagement	22. What is the most appropriate model structure for this appraisal?
Background/description of issue	A partitioned survival model populated with extrapolated PFS and OS data requires the availability of reliable, mature PFS and OS data for both the intervention and the comparator. **The company** structured its base case economic model as a partitioned survival model with 3 mutually exclusive health states: progression free, progressed and dead. The intervention arm used PFS and OS extrapolated data from the full integrated efficacy analysis set (n=102). The comparator arm used PFS and OS data extracted from the review of populations that harbor the *NTRK* fusion (see Issue 8). The company also provided scenarios that used a dual partitioned survival model based on response and a model that used time to progression and ORR from previous line of therapy. **The ERG** preferred the response-based analysis, which assumed non-responder patients are a proxy for people that receive best supportive care. It also assumes a surrogate response between response and time-to-event outcomes, which the ERG acknowledges is a weakness of the analysis, but it allows a flexible framework to explore alternative assumptions on the predicted survival outcomes for larotrectinib. This is because fewer observations are required on response outcomes to draw meaningful conclusions about differences between tumour types. The ERG adapted the company response-based model, using ORRs across tumour types that were estimated by the ERG's Bayesian Hierarchical Model (see section 4.6.1 of ERG report). The survival of non-responder patients was used to estimate survival predictions in the established management arm. The larotrectinib arm was based on a weighted average of responder and non-responder survival predictions. The ERG noted that the non-responder analysis predicted that PFS appears to be worse for non-responders than in the comparator data. This suggests that non-responders are not representative of people who do not receive larotrectinib and may have a poorer prognosis. The opposite is true for OS, non-responders have better survival than the comparator data which may be due to post-progression treatments. This shows the limitations of the assumption that non-responders are equivalent to the comparator arm.
Why this issue is important	The comparator arm used in the company's economic model may not be appropriate for decision making because it does not allow any consideration of heterogeneity (see issues 7- 9). The response-based model with BHM response outcomes may partially overcome these issues. Choice of model structure and parametric curve fit are the main driver in change in ICER between the company and ERG base cases.
Technical team preliminary judgement and rationale	The technical team consider the issue of heterogeneity to be an important issue in decision making (see Issue 7) – therefore, the technical team consider the response-based model to be more appropriate for answering this question. However, it considers the response-based analysis to rely on a small number of patients and relies on different assumptions about the link between response with OS and PFS as seen in the comparison to the weighted comparator arm in the company base case. The technical team consider the response-based model to be a plausible alternative to the company's base case model.

Summary of comments	**Comments received from the company:**
	• There is no precedence or guidance for evaluating the cost-effectiveness of histology independent treatments. A number of steps were taken to validate the approach taken for economic evaluation.
	• Appreciating the limitations of the different approaches, the company considered that all model structures produced very similar ICERs, lending credibility to the results.
	• The company believe that a simple partitioned survival model that does not require the assumption of a surrogacy relationship between response and survival to be a more appropriate approach for the current appraisal. The response based model needs strong assumptions to incorporate the analysis into the model:
	  ○ Low numbers of events, especially important for non-responders
	  ○ Uncertainty in the projected survival curves, given the relatively short, variable follow-up in the clinical evidence
	  ○ Differences in the distribution of tumour sites/disease severity between responders and non-responders are not accounted for
	  ○ The assumption that the non-responders would represent a control arm – patients who do not respond may be inherently different to patients that do respond.
	• The company also consider the methodology implemented to adjust the ratio between progression-free and post-progression survival to be inappropriate, citing NICE DSU:
	"*The level of evidence available supporting a relationship between PFS/TTP and OS varies considerably by cancer type and is not always consistent even within one specific cancer type… Therefore, any cost-effectiveness analysis which makes a strong assumption regarding the relationship between PFS and OS should be treated with caution."*
	ERG considerations on new evidence received during technical engagement:

	• The ERG agreed that the company appropriately explored alternative approaches to model the comparator survival outcomes and that all approaches have limitations and may result in biased estimates of treatment effectiveness. • However, the non-responder control provides a more transparent and flexible alternative to modelling comparator effectiveness than the pooled historical comparator. In addition, only by using the responder-based model, can heterogeneity according to tumour type be explored, given the lack of further data provided by the company. • The ERG agree that the relationship between response status and survival outcomes is highly uncertain. The ERG request for PFS/OS data according to tumour type was not granted by the company, and therefore the ERG assume a common distribution of PFS and OS across tumour sites, conditional on response status • Crude adjustments to post-progression survival were made in the absence of being able to separately track where the gains in post-progression survival occur.
Technical team judgement after engagement	The technical team acknowledge the uncertainty of surrogate relationships between response, PFS and OS. However, the uncertainties inherent in this appraisal are likely of greater magnitude than potential issues with surrogacy outcomes. The technical team consider the response-based analysis a plausible alternative to the company base case.

Issue 12 – Extrapolation of overall and progression-free survival

Questions for engagement	23. Which parametric curve is most appropriate for extrapolating overall and progression-free survival in the company base case? 24. Which parametric curve is most appropriate for extrapolating overall and progression-free survival in the response-based analysis?
Background/description of issue	The progression-free survival and overall survival measured in the trial is immature and extrapolation is needed to estimate survival over the full time horizon. The comparator extrapolation is informed by multiple fitted parametric curves from multiple sources (see Issue 9) **The company** considered the [redacted] distribution to be most appropriate to extrapolate PFS and OS based on clinical plausibility because the visual and statistical fit were similar for all six standard parametric curves explored by the company. **The ERG** noted the considerable uncertainty associated with the extrapolation of both datasets with only ~37% of patients progressed and ~14% of patients that have died. The ICER is sensitive to the choice of distribution in PFS and it appears to be driven by the cost of treatment, as well as differences in QALY gains. The overall survival extrapolation shows substantial separation between extrapolated PFS and OS curves and these survival gains may be driven by post-progression treatments (see Issue 10). The extrapolations for the pooled comparator analysis are shown in Figure 7. **The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund** considers that the [redacted] extrapolation for both overall and progression-free survival is as clinically plausible as the [redacted] distribution in the extrapolation of the larotrectinib arm. **Figure 7 – Extrapolations in the pooled comparator analysis for progression-free survival (left) and overall survival (right)** [redacted] **The ERG** prefer to use the responder analysis, the ERG used the extrapolations that the company considered clinically plausible for this analysis, but also included the [redacted] extrapolation in their considerations. The ERG considered two scenarios with the [redacted] distribution, both of which provided more clinically plausible estimates of post-progression survival than the [redacted] and [redacted] extrapolations. However, because of [redacted] using the [redacted] distribution, the ERG considered the most appropriate scenario to be the [redacted] for OS and [redacted] for PFS in the response-based analysis. The ERG highlighted that this scenario may still overestimate overall survival in larotrectinib and did not consider the results to be robust. These extrapolations for the response-based analysis are shown in Figure 8. **Figure 8 – Extrapolations in the response-based analysis for progression-free survival (left) and overall survival (right)**

	[redacted]
Why this issue is important	Choice of function for the survival extrapolation impacts on the clinical plausibility of the estimated survival. The company's choice of the exponential function gives survival estimates that are longer in the post-progression health state than the progression-free health state. Choice of model structure and parametric curve fit are the main driver in change in ICER between the company and ERG base case.
Technical team preliminary judgement and rationale	There is substantial uncertainty surrounding the extrapolation because the data is immature. An updated data cut would resolve some of the uncertainty about the issue of maturity, but even mature data may not be robust because of structural uncertainty, particularly because of potential confounding of post-progression treatments. The technical team consider that extrapolation does not provide meaningful results for the model but in the absence of robust evidence, the most conservative assumptions should be assumed.
Summary of comments	**Comments received from the company:** • The company acknowledge uncertainty around the long-term efficacy and survival profile for larotrectinib. The uncertainty is driven by the immaturity of the data, with ongoing data collection providing more certainty of the outcomes over time. • Reasonable to consider clinical plausibility rather than seeking definitive answer or taking a highly conservative approach • The company followed the NICE DSU preferred fitting of standard parametric curves and explored uncertainty around the extrapolations extensively. • The company provided PFS and OS data from an updated data-cut provided in Figure 9. **Figure 9 – Kaplan-Meier curves from the updated data-cut for progression-free survival (top) and overall survival (bottom)** [redacted]

	ERG considerations on new evidence received during technical engagement: • The ERG consider that the shorter post-progression survival gains for larotrectinib using the Gompertz distribution for OS makes it more clinically plausible than the company base case. While the Gompertz distribution generates more conservative survival predictions, given the immaturity of OS data, this may still results in optimistic survival projections. • [redacted] The exact nature of the survival curve at longer follow-up remains uncertain. • Applying the updated data set with the company base case modelling assumptions to the results in an increased ICER. The results are driven by an increase in progression-free survival which increases cost of treatment which are offset with a reduction in overall survival ([redacted]).
Technical team judgement after engagement	The technical team maintain that extrapolation does not provide meaningful results for the model because of the immaturity of the data, particularly as the company have suggested a potential cure for some tumour types (see Issue 10). The updated survival data from the new data cut do not provide enough evidence to support any change in technical team judgement. The updated progression-free survival KM curve is based on investigator-assessed progression which limits comparability with the earlier dataset. The OS KM curve may suggest a [redacted] however the ICER is very sensitive to this extrapolation of the data.

Issue 13 – Drug wastage and adherence

Questions for engagement	25. Have potential drug wastage costs and adherence costs been appropriately included in the model?
Background/description of issue	Drug wastage can occur when people stop taking treatment that has been distributed to them. Additionally, not all patients within the trial adhered to the full recommended therapeutic dose. **The company** assumed that the acquisition costs of larotrectinib were applied at the start of the treatment cycle, with no half-cycle correction to account for treatment wastage due to patients discontinuing treatment during the 7 days cycle. Also, the average adult and paediatric doses used in the trials were considered instead of the recommended daily dose. The company provided scenarios including supplying 2 and 4 week treatment supplies for drug wastage and full adherence to the specified doses. **The ERG** consider that the company model may underestimate the larotrectinib dose in the paediatric population because [redacted] ([redacted]) paediatric patients may have received sub-therapeutic doses of larotrectinib because some patients received a lower dose in the phase I dose-finding study, SCOUT. The ERG provided a scenario which applied the adult adherence percentage to the paediatric dose to estimate the effect of including the costs of the full licensed dose in the paediatric patients.
Why this issue is important	Drug wastage and adherence costs are important to capture with high cost specialist oncology drugs and increase the ICER.
Technical team preliminary judgement and rationale	**The technical team** consider the ERG correction to the paediatric dose to be appropriate, although the clinical benefit of receiving the licensed dose would not be modelled from this correction. There is no standard clinical practice for prescribing larotrectinib, so it is difficult to make an assumption about drug wastage costs. However, the technical team consider that the effect on the ICER is likely to be minimal in comparison to the structural uncertainty inherent in this appraisal.
Summary of comments	**Comments received from the company:** • The ERG have made an adjustment to the paediatric dose in their analysis, and in doing so, including increased costs, the clinical benefit of receiving the increased dose has not been modelled. **ERG considerations on new evidence received during technical engagement:** • The company did not allow adjustments to the effectiveness of larotrectinib based on the dose received by paediatric patients.
Technical team judgement after engagement	The technical team consider that the ERG correction to the paediatric dose is appropriate and that drug wastage should be incorporated in the base case.

Issue 14 – Administration costs and resource use

Questions for engagement	26. Have administration costs and resource use been adequately captured in the company's model?
Background/description of issue	**The company** estimates tumour site specific acquisition and administration costs according to the comparator treatments that are included in the pooled comparator analysis. Health state resource use and costs also differ across tumour sites for the comparator. **The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund** highlighted that the company did not include the chemotherapy tariff cost for oral treatments in the larotrectinib arm of the model. The appropriate cost for inclusion is SB11Z oral chemotherapy tariff (£120 per visit). In addition, the clinical lead notes that the estimates of the weekly comparator chemotherapy costs for paediatric soft tissue sarcoma and infantile sarcoma as ranging from [redacted] which is a very wide range and requires detailed justification and the treatment duration for this group of patients is [redacted] days which the clinical lead considers remarkably long for a non-standard therapy which is 'not satisfactory'.

Why is this issue important	It is unclear how administration costs and resource use affect the ICER because there is considerable uncertainty for the costs and resource use of some comparator treatments that were not appraised by a NICE committee and the costs and resource use of larotrectinib in NHS clinical practice.
Technical team preliminary judgement and rationale	The technical team consider that there may be issues with some of the assumptions used for administration costs and resource use, however these issues are likely to increase the ICER and there are no scenarios exploring these costs. Additionally, all costs are linked to the distribution of patients in the trial population which will not be representative of the target population (see Issue 1).
Summary of comments	**Comments received from the company:** • The company notes that administration costs for oral chemotherapies are not consistently applied in NICE TAs • The company have provided a scenario that include the administration costs of one oral chemotherapy per 30-day treatment cycle, reflecting clinical practice for oral chemotherapies. • The full resource impact of avoiding intravenous chemotherapy is unlikely to have been accounted for in the modelling.
Technical team judgement after engagement	The technical team consider that oral chemotherapy costs should be incorporated into the base case.

Issue 15 – Implementation and training costs

Questions for engagement	27. What additional infrastructure and training requirements could be considered for this appraisal?
Background/description of issue	Site agnostic oncology treatments are a new concept in clinical practice in England. Oncologists will likely require training about what tumour types may be eligible for treatment with larotrectinib and at what point in the treatment pathway larotrectinib can be used. Further, training will likely be required around the collection and handling of the tissue biopsies for testing for *NTRK* gene fusions.
Technical team preliminary judgement and rationale	It is important to capture the impact of using larotrectinib in clinical practice. Any likely constraints on the resources required to support the implementation of the appraised technology should be highlighted, and comment should be made on the impact this may have on the implementation timescale.
Summary of comments	**Comments received from the company:** • Training on identification and place in treatment of new therapies is not routinely included in appraisals. This is part of CPD for clinicians, often supported by pharmaceutical companies. • Training on handling and collection of biopsies does not apply solely for identification of patients potentially suitable for larotrectinib and therefore the costs of this should not be assigned to a single technology.
Technical team judgement after engagement	This issue will be discussed by the committee with input from NHS England and Genomics England. Any likely impact on the implementation timescale will be noted in the appraisal documents.

Issue 16 – Utility values

Questions for engagement	28. How closely do the utility values modelled match the utility values of patients in clinical practice? 29. Is there justification for considering post-progression utility values to be different between larotrectinib and best supportive care?
Background/description of issue	Utility values for pre- and post-progression were derived from health-related quality of life data collected in two of the larotrectinib single-arm trials. Utility values for the comparator were assumed to be different and were collected as part of the construction of the comparator arm (see Issue 8). **The company** mapped EQ-5D-5L measurements in NAVIGATE to EQ-5D-3L utilities and PedsQL generic core scales from SCOUT were mapped to EQ-5D-3L utilities. These utilities were used to create two regression models, stratified by progression status. The company base case is shown in Table 6. **Table 6: Utility values considered in the economic model**

	Larotrectinib		Weighted comparator	
	Pre-progression	Progressed	Pre-progression	Progressed
Company base case	[redacted]	[redacted]	[redacted]	[redacted]
ERG scenario	[redacted]	[redacted]	[redacted]	[redacted]

The ERG considered the utility estimates to be affected by several sources of uncertainty:

- The analysis is informed by a small number of observations (progressed disease state is derived from [redacted] assessments in [redacted] patients [redacted] of which were paediatric patients)
- The unadjusted mean utility values show that mean post-progression utility values are often higher than for progression-free at the same cycle in NAVIGATE.

The ERG provided a scenario to show the uncertainty surrounding the assumption of differential utility weight for post-progression on larotrectinib compared to comparator treatments. The scenario considers that the pooled progressed disease utility is equal for larotrectinib and the comparator arm. This increased the ICER by [redacted] compared to the ERG base case.

	The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund considers it appropriate to use the similar utility values in the progression-free health state for the larotrectinib and comparator arms and that for the post-progression utility values further divergence after progression is implausible compared with what is seen in clinical practice.
Why this issue is important	Inaccurate utility values could bias estimates. Differential utility values for the post-progression health state between the two treatment arms biases the cost-effectiveness results in favour of larotrectinib in the company base case.
Technical team preliminary judgement and rationale	It is unclear what the utility values in the post-progression health state would be because there is limited health-related quality of life data from the trials. The technical team consider that the company have not provided robust justification for why the post-progression health state would be different between treatment arms so prefer the scenario that assumes equivalency. The technical team also consider that the difference in utility values for the progression-free state in the model may be larger than would be seen in clinical practice.

Summary of comments	**Comments received from the company:**
	• Larotrectinib utility values were derived from the clinical trial programme.
	• Comparator utility values were applied using NICE TAs or the literature if none could be identified. The process and scrutiny undertaken in each technology appraisal selected committee's preferred input, these were determined to be the most suitable for decision making.
	• The company consider that there is justification for maintaining a difference in quality of life post-progression and that this is clinically plausible.
	• The company assessed 48 previous TAs in a 'last line' setting and found a number of appraisals that reported significant QALY gains between treatment and comparator in the post-progression state.
	• In [redacted] patients in the SCOUT trial who had no other curative options besides amputation or disfiguring surgery, larotrectinib enabled an increased rate of limb sparing surgery. Disfiguring surgery can have devastating, lifelong consequences. Limb preservation has a higher quality of life than patients with amputation.
	• Larotrectinib was well tolerated, whereas chemotherapy can be associated with significant adverse effects that could plausibly lead to the difference in post-progression utilities. Some examples of these include:
	  ○ Persistent symptoms of neuropathy (e.g. cisplatin and neurotoxicity from ifosfamide)
	  ○ Nephrotoxicity (cisplatin, gemcitabine, ifosfamide)
	  ○ Cardiotoxicity (risk associated with anthracycline, radiation induced pericardial effusion)
	  ○ Infertility (chemotherapy and radiotherapy)

Technical team judgement after engagement	The technical team consider that the difference in post-progression utility values is implausible. The potential for mutilating surgery could account for the difference in post-progression utility values. However this reduction in utility in the weighted comparator arm does not apply over the full time horizon as patients in the comparator engine for paediatric sarcomas are not modelled to survive for their natural lifespan (see Issue 10). Also, this utility difference would only apply to a proportion of patients, not the entire cohort and would depend on the distribution of patients (see Issue 1). The persistent chemotherapy symptoms in the post-progression state could plausibly offer some utility decrement, but a large proportion of patients receive best supportive care in the company model and this would likely increase with a more appropriate distribution of patients.

Issue 17 – End of life

Questions for engagement	30. What is the life expectancy of the patient group receiving established management? 31. What is the extension to life of the patient group receiving larotrectinib?
Background/description of issue	There is no direct comparator data on survival without larotrectinib which makes assessment of whether larotrectinib meets the end of life criteria (specified in NICE's guide to the methods of technology appraisal) difficult. **The company** consider larotrectinib meets the end of life criteria. They commissioned a systematic literature review of tumours known to harbour *NTRK* fusions which showed median PFS was less than 12 months in patients who receive more than 1 prior therapy across the included tumour types. This suggests that the life expectancy would be less than 24 months. The data for overall survival is immature so the extension to life criterion is not estimable, although the company consider the increase in PFS shows evidence of extension to life. **The ERG** consider end of life criteria to be met when data are pooled across tumour sites, although there is no meaningful treatment benefit for primary CNS tumours. The ERG considered end of life by individual tumour site by comparing median PFS and OS of the comparator data to results in the ERG analysis of PFS and OS. This analysis showed that life expectancy is below 24 months for all tumour types except [redacted], and close to 24 months for [redacted]. For extension to life, the ERG considered good evidence of extension to life to be an increase in PFS of at least 3 months, based on the 5th centile of the simulated PFS data (see Issue 7). This showed 3-month extension to life for [redacted], and the medians are presented in Table 7. **Table 7: Overall survival estimates of the comparator arm and increase in PFS from ERG simulated data analysis (from ERG table 52)**

	Tumour site / **Median* overall survival (from comparator data)** / **Median* increase in PFS (from ERG analysis)** Bone sarcoma [redacted] Breast [redacted] Cholangiocarcinoma [redacted] Colon [redacted] CMN [redacted] GIST [redacted] Infantile fibrosarcoma [redacted] Lung [redacted] Melanoma [redacted] Pancreas [redacted] Salivary gland [redacted] Soft tissue sarcoma [redacted] Thyroid [redacted] *Committees prefer to see the mean life expectancy and mean extension to life in their considerations of whether a technology meets the end-of-life criteria.
Why this issue is important	The appraisal committee's judgements about the acceptability of the technology as an effective use of NHS resources will take into account whether the technology meets the criteria for special consideration as a 'life-extending treatment at the end of life'. A technology which meets NICE's end of life criteria has an increased cost-effectiveness threshold.
Technical team preliminary judgement and rationale	The technical team acknowledge that this analysis is based on the simulated data but consider it to be appropriate. [redacted] is unlikely to meet the short life expectancy criterion, and a recommendation in favour of larotrectinib would distort research incentives for this indication. Committees prefer the mean estimates which are generally longer than the median, because of this, it is likely that [redacted] also do not meet the short life expectancy criterion. It is unclear the percentage of the population receiving larotrectinib that would have these tumour types (see Issue 1). Likewise, for the extension to life criterion, there appear to be some tumour types that do not increase progression free survival and would not meet this criterion, the lack of robustness of the model means that it is unclear the percentage of the target population that would not meet the extension to life criterion. A proportion of the tumour types included in the analysis do not meet the end-of-life criteria and committee will consider this in their deliberations when making a judgement. Additionally, there are some unrepresented tumour types which are unknown if they meet the end-of-life criteria, or their prevalence. There is considerable uncertainty for the end of life decision, including the robustness of the data, the positioning of larotrectinib in the treatment pathway (see Issue 2) and the target population (see Issue 1).
Summary of comments	**Comments received from the company:** • The indication concerns a disease setting of locally advanced or metastatic solid tumours where symptoms of disease will be present or imminent and the disease is incurable, likely leading to death. • The company restated their justification for why larotrectinib meets the end-of-life criteria using median PFS from the comparator arm and likely extension to life of greater than 3 months. • It is not appropriate to consider applying the end of life criteria to specific cancer sites based on their location.
Technical team judgement after engagement	The technical team recognise the uncertainties around applying the end-of-life criteria to histology-independent treatments such as larotrectinib. The committee will consider these in their deliberations when making a judgement on whether larotrectinib meets the end-of-life criteria.

Issue 18 – Innovation

Questions for engagement	32. Is larotrectinib an innovative treatment?
Background/description of issue	**The company** claim that larotrectinib is an innovative treatment. They describe larotrectinib as a step-change in the treatment of cancer, as a paradigm shift in the way cancer is treated from using tumour location to delivering treatment based on causation (i.e. the presence of *NTRK*). Utilising novel genomic technologies such as NGS to identify *NTRK* fusion positive solid tumours may also provide benefits to patient health and cost efficiencies for health care systems as multiple different actionable targets may be identified, even where *NTRK*-fusion negative and this could lead to clinical trial availability or treatment with other targeted therapies.
Why this issue is important	Committee can take into account the potential innovative nature of the technology, in particular its potential to make a significant and substantial impact on health-related benefits that are unlikely to be included in the QALY calculation during the appraisal.
Technical team preliminary judgement and rationale	Larotrectinib is potentially innovative in that it is a treatment for a newly identified rare gene fusion that occurs in a wide range of tumour types. The technical team recognise that it is one of the first site-agnostic treatments to be appraised by NICE. However, they are aware of other targeted inhibitors that can be used to treat a range of different tumour types and in those cases (for example treatments for the BRAF V600E mutation), larger studies have been done. Larotrectinib could represent a step-change in the treatment of cancer however, there is a lack of evidence of demonstrable and distinctive benefits of a substantial nature which may not have been adequately captured in the reference case QALY measure.

	The technical team consider that a major innovation is already being led by the NHS in developing more sophisticated strategies to improve genomic testing in clinical practice. These advances may facilitate uptake of treatments such as larotrectinib if it is to be recommended. However, larotrectinib is one of the first site-agnostic treatments to be appraised by NICE and represents potential for a future service redesign based on biological marker rather than histology. The committee will consider the innovative nature of larotrectinib when making its recommendations.
Summary of comments	**Comments received from the company:** • Larotrectinib provides a treatment for a new target, NTRK fusion-positive solid tumours • Larotrectinib selectively inhibits TRKA, TRKB and TRKC, designed to avoid activity with off-target kinases • Larotrectinib treats both adults and children within one indication • Larotrectinib is a step towards delivering personalised medicine for cancer patients. • Unmeasured benefit from providing an oral oncology medication which does not impact on schooling and parents as much as treatment regimens that require daily visits to the hospital. Also unmeasured benefit from decreasing
Technical team judgement after engagement	Committee will take into account the potential innovative nature of larotrectinib as part of its decision making.

Issue 19 – Cancer Drugs Fund

Questions for engagement	33. Does larotrectinib meet the criteria for inclusion in the Cancer Drugs Fund? 34. What data would be most useful to collect to address the outstanding uncertainties? For example, unrepresented tumour types.
Background/description of issue	**The company** have proactively positioned larotrectinib for funding via the Cancer Drugs Fund (CDF) as opposed to by routine commissioning in the NHS. **The technical team** is aware of the arrangements for the Cancer Drugs Fund agreed by NICE and NHS England in 2016, noting NICE's Cancer Drugs Fund methods guide (addendum). The technical team consider that there is substantial clinical uncertainty that could be partially reduced through data collection via ongoing studies.
Why this issue is important	The CDF is a potential option if there is plausible potential for the drug to satisfy the criteria for routine commissioning, but there is significant remaining clinical uncertainty which needs more investigation, through data collection in the NHS or clinical studies. This means the CDF will fund the drug, to avoid long delays, but would require information on its effectiveness before it can be considered for routine commissioning (when the guidance is reviewed). The company have not provided evidence to demonstrate that larotrectinib has plausible potential for cost-effectiveness.
Technical team preliminary judgement and rationale	The technical team acknowledges the company's positioning of larotrectinib for consideration within the Cancer Drugs Fund, however it considers that larotrectinib does not meet the criteria for inclusion in the Cancer Drugs Fund at its current price. The company's base case ICER is above the range that NICE would normally consider cost effective when a treatment meets the end-of-life criteria. This ICER does not take into account the ERG and technical team's preferred assumptions. The technical team consider that if larotrectinib does become plausibly cost-effective, committee will be interested in the practicalities of data collection within the Cancer Drugs Fund during the course of the appraisal. In particular, they will be interested in NHS England's intentions around data collection being prioritised based on unmet need, tumour types where no data has been collected previously or tumour types with high prevalence of *NTRK* gene fusions.

Summary of comments	**Comments received from the company:**
	• The company request the committee give balanced consideration to downward as well as upward uncertainty associated with evaluating an histology independent innovation.
	• Given the current level of uncertainty, the company proposes that whilst data mature, larotrectinib is made available in a timely manner through the CDF
	• Company considers the following longer term data could be informative:
	  ○ Larger sample size
	  ○ Populations who would use larotrectinib in clinical practice (by tumour site, place in therapy and optimal time of testing)
	  ○ OS and extension to life
	  ○ Post-progression treatments
	  ○ Impact of avoiding amputation/ disfiguring surgery

Technical team judgement after engagement	The technical team maintains that larotrectinib does not meet the criteria for inclusion in the Cancer Drugs Fund because it does not have plausible potential to be cost-effective at the current price.

4. Issues for information

Tables 8 to 10 are provided to stakeholders for information only and not included in the technical report comments table provided.

Table 8: Technical team preferred assumptions and impact on the cost-effectiveness estimate

Alteration	Technical team rationale	ICER	Change from base case
Company base case	**–**	£35,309	**-**
ERG base case	Issues 7-11	£48,872	+£13,563
ERG base case + SAS3 dataset	Issue 5	£49,621	+£14,312
ERG base case + post-progression survival equal for larotrectinib and comparator	Issue 10	£94,444	+£59,135
ERG base case + same post-progression utility for larotrectinib	Issue 16	£58,047	+£22,738
Drug wastage and adherence, administration costs and resource use	Issue 13-14	Unknown	-
ERG base case + diagnostic testing costs	Issue 3	£54,154	+£18,845
Cumulative impact of the technical team's preferred assumptions on the cost-effectiveness estimate – no diagnostic testing costs	**-**	**£97,923**	**+£62,614**
Cumulative impact of the technical team's preferred assumptions on the cost-effectiveness estimate – plus ERG diagnostic testing costs	**-**	**£109,965**	**+£74,656**

Table 9: Outstanding uncertainties in the evidence base

Area of uncertainty	Why this issue is important	Likely impact on the cost-effectiveness estimate
Clinical evidence for larotrectinib	The clinical evidence base for larotrectinib is small, (n=102) including people with 15 different tumour types. The largest group of patients is [redacted] for a single tumour type ([redacted]) and only [redacted] patient for some tumour sites ([redacted]) pooled from 3 separate clinical trials. There is also considerable uncertainty regarding the extent to which the high response rates seen in the results from the larotrectinib clinical trials results translate into clinically meaningful survival benefits. It is unclear if the high complete response rate (16% had a complete response) could be considered a cure for some tumour types and stages.	The effect of the limitations of the evidence base introduces substantial uncertainty with an unknown direction of bias.
Immature evidence base	The analysis from the larotrectinib trials is of short duration. Median overall survival in the trial has not yet been reached (14% had an event). Progression-free survival data are also immature (37% had an event). Analyses are based on extrapolated values.	Substantial uncertainty around the extrapolations of estimates of the partitioned-survival model – unknown direction of bias.

Table 10: Other issues for information

Issue	Comments
Larotrectinib has been granted a conditional marketing authorisation	Larotrectinib was granted a positive CHMP opinion of a conditional marketing authorisation on 25th July 2019, the conditional marketing authorisation specified the following post-authorisation obligations: • Confirm histology-independent efficacy by submitting a pooled analysis with larger sample size including the final NAVIGATE report • Study the long-term toxicity and developmental effects of larotrectinib in paediatric patients from the SCOUT study, including 5 year follow up data • Confirm the appropriate paediatric dose from SCOUT with an update pharmacokinetic model based on additional sample of patients between 1 month to 6 years.
Oncology genomic testing is not yet fully operational in clinical practice	The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund states that WGS will be fully operational by quarter 2 in the year of 2020/21 and NGS panel testing will be available by quarter 1 in 2020/21. Uptake of molecular testing across the 7 genomic hubs will increase during 2020/21 as genomic pathways are embedded and links are made with the clinical teams. Given the complexity of implementation, it may take a further 12 months for molecular testing to become fully embedded in practice.
Using a single ICER to represent the cost-effectiveness of larotrectinib conceals the potential for significant variation in tumour specific ICERs	Larotrectinib may not be cost-effective across all tumour types that are *NTRK* fusion positive. The ERG utilised its response-based model to integrate the results of the Bayesian hierarchical analysis and generated tumour type specific ICERs. This exploratory analysis showed that the tumour type specific ICER's varied from [redacted] per QALY in [redacted] to [redacted] per QALY in [redacted]
Company's estimate of the eligible population is uncertain	The company estimate [redacted] patients eligible for last line therapy in the tumour types that are known to harbour *NTRK* fusion positive cancers. The ERG analysis analysing those that are likely to have *NTRK* fusion positive tumours reduces this to [redacted]. However, there is a great deal of uncertainty in these estimates and of the prevalence of *NTRK* fusions in total.
***NTRK*2 gene fusion positive tumours are under-represented in the larotrectinib clinical trials**	A small number of *NTRK*2 gene fusion positive tumours were identified in the trial (n=10) and the majority of these had primary CNS tumours. It is unclear how prognosis and response are related to the *NTRK* gene fusion.
Company's model in original submission did not include data from the most recent data cut from the larotrectinib clinical trials	The company presented data from the July 2018 data-cut of the ePAS2 data. The technical team were aware that a more recent data-cut (February 2019) from the larotrectinib pooled efficacy study was available but had not been fully analysed for this submission. The updated efficacy analysis contained an expanded cohort of n=51 additional efficacy-evaluable patients.
Equality considerations	The company do not consider there to be any equality issues; however, they consider that the uncertainty inherent in this appraisal may pose an equity issue. There is no precedent for appraising technologies with basket trial design and a high number of comparators across multiple tumour sites. The company consider that patients should have equity of access whilst health technology assessment methods adapt to these challenges.

Authors

Gary McVeigh

Appraisal committee chair

Adam Brooke

Technical lead

Christian Griffiths

Technical adviser

Linda Landells

Associate director

With input from the lead team:

Guy Makin

Lead team member

David Meads

Lead team member

Malcolm Oswald

Lead team member