

Single Technology Appraisal

Sotatercept for treating pulmonary arterial hypertension [ID6163] Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Contents:

The following documents are made available to stakeholders:

The [final scope and final stakeholder list](#) are available on the NICE website.

1. [Company submission summary from Merck Sharp & Dohme](#)
2. [Company summary of information for patients \(SIP\) from Merck Sharp & Dohme](#)
3. [Clarification questions and company responses](#)
4. [Expert personal perspectives from:](#)
 - a. [Dr Daniel Knight, Consultant cardiologist – clinical expert, nominated by Merck Sharp & Dohme \(UK\) Limited](#)
 - b. [Dr Iain Armstrong, PHA UK Chair & Nurse Consultant – clinical expert nominated by PHA UK](#)
5. [External Assessment Report prepared by Aberdeen HTA Group](#)
6. [External Assessment Report – factual accuracy check](#)

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

Sotatercept for treating pulmonary arterial
hypertension [ID6163]

Company evidence submission



April 2025

File name	Version	Contains confidential information	Date
ID6163 Company Evidence Submission v3.0 (01-APR-2025) (CIC).docx	3.0	Yes	01-APR-2025

Contents

Contents	2
Tables	4
Figures	7
Abbreviations	9
Executive summary	12
1 Decision problem, description of the technology and clinical care pathway	15
1.1 Decision problem	16
1.2 Description of the technology being evaluated	22
1.3 Health condition and position of the technology in the treatment pathway	22
1.4 Equality considerations	32
2 Clinical effectiveness	33
2.1 Identification and selection of relevant studies	35
2.2 List of relevant clinical effectiveness evidence	35
2.3 Summary of methodology of the relevant clinical effectiveness evidence	37
2.4 Statistical analysis and definition of study groups in the relevant clinical effectiveness evidence	46
2.5 Critical appraisal of the relevant clinical effectiveness evidence	51
2.6 Clinical effectiveness results of the relevant studies	51
2.7 Subsequent treatments used in the relevant studies	99
2.8 Subgroup analysis	99
2.9 Meta-analysis	112
2.10 Indirect and mixed treatment comparisons	113
2.11 Adverse reactions	122
2.12 Ongoing studies	128
2.13 Interpretation of clinical effectiveness and safety evidence	128
3 Cost effectiveness	131
3.1 Published cost-effectiveness studies	133
3.2 Economic analysis	148
3.3 Clinical parameters and variables	157
3.4 Measurement and valuation of health effects	183
3.5 Cost and healthcare resource use identification, measurement and valuation ..	191
3.6 Severity	203
3.7 Uncertainty	204
3.8 Managed access proposal	205

3.9 Summary of base case analysis inputs and assumptions 206

3.10 Base case results 215

3.11 Exploring uncertainty 216

3.12 Subgroup analysis..... 227

3.13 Benefits not captured in the QALY calculation 227

3.14 Validation 230

3.15 Interpretation and conclusions of economic evidence..... 231

4 References 235

5 Appendices 245

Tables

Table 1. The decision problem	17
Table 2. Technology being evaluated.....	22
Table 3 Treatments for PAH currently available for use in the UK as defined in NHS clinical commissioning policies	25
Table 4 2022 ESC/ERS simplified four-strata risk assessment tool for guidance of therapeutic decision-making.....	27
Table 5 Positioning of sotatercept in the treatment pathway.....	31
Table 6. Clinical effectiveness evidence.....	36
Table 7. Summary of trial methodology	37
Table 8 Inclusion and exclusion criteria for the STELLAR trial	40
Table 9 Demographics and Baseline Characteristics - Full Analysis Set	41
Table 10 Studies from which participants may enter the SOTERIA study.....	44
Table 11 SOTERIA study baseline characteristics of trial participants.....	45
Table 12 Statistical analyses for trial endpoints.....	47
Table 13 Summary of SOTERIA study statistical analysis plan	48
Table 14 STELLAR study - Disposition of participants – Full Analysis Set	52
Table 15 STELLAR study - Summary of follow-up time (Weeks) - Full Analysis Set	52
Table 16 STELLAR – Study drug exposure by duration – Safety set.....	53
Table 17 STELLAR – Study drug exposure – Safety set – Cumulative results (DBPC+LTDB)	54
Table 18 STELLAR study - Overall summary of results - Full study population	54
Table 19 Analysis of change from baseline in 6-Minute Walk Distance (metres) at Week 24 - Full Analysis Set - (Standard Multiple Imputation)	60
Table 20 Analysis of change from baseline in 6-Minute Walk Distance (by at least 30 metres) at Week 24 - Full Analysis Set	61
Table 21 Change from Baseline in 6-Minute Walk Distance (metres) - Full Analysis Set.....	61
Table 22 Analysis of Multicomponent Improvement endpoint at Week 24 - Full Analysis Set	64
Table 23 Change from Baseline in PVR (dynes*sec/cm ⁵) by Visit 24 - Full Analysis Set	65
Table 24 Analysis of change from baseline in PVR (dynes*sec/cm ⁵) at Week 24 - Full Analysis Set (Standard multiple imputation).....	66
Table 25 Analysis of change from baseline in PVR (dynes*sec/cm ⁵) at Week 24 - Full Analysis Set (Standard multiple imputation with ANCOVA model)	66
Table 26 PVR responder analysis - Full Analysis Set.....	66
Table 27 Change from baseline in NT-proBNP (pg/mL) by visit - Full Analysis Set	67
Table 28 Analysis of change from baseline in NT-proBNP (pg/mL) at Week 24 - Full Analysis Set (Standard multiple imputation)	68
Table 29 Analysis of change from baseline in NT-proBNP (pg/mL) at Week 24 - Full Analysis Set (Standard multiple imputation with ANCOVA model)	69
Table 30 Analysis of change from baseline in WHO FC at Week 24 - Full Analysis Set.....	70
Table 31 Shift from baseline in WHO functional class by Visit 24 – Full Analysis Set.....	70
Table 32 Time to first WHO FC improvement.....	72
Table 33 Time to first clinical worsening or death - Full Analysis Set.....	74
Table 34 Low risk score (French Risk Score calculator) by visit - Full Analysis Set.....	77

Table 35 Analysis of change from baseline in low risk score (French Risk Score calculator) at Week 24 - Full Analysis Set	77
Table 36 Four-strata risk status assessment by treatment at baseline and Week 24 – Full Analysis Set	78
Table 37 Four-strata risk status assessment transitions from baseline to Week 24 – Full Analysis Set	79
Table 38 Analysis of change in four-strata risk status from baseline at Week 24 based on complete cases - Full Analysis Set	79
Table 39 Analysis of change in four-strata risk status from baseline at Week 24 based on imputed cases - Full Analysis Set	80
Table 40 Change from baseline in Physical Impacts Domain Score of PAH-SYMPACT® by visit - Full Analysis Set	82
Table 41 Analysis of change from baseline in Physical Impacts Domain Score of PAH-SYMPACT® at Week 24 - Full Analysis Set (Standard multiple imputation)	82
Table 42 Analysis of change from baseline in Physical Impacts Domain Score of PAH-SYMPACT® at Week 24 - Full Analysis Set (Standard multiple imputation with ANCOVA model)	83
Table 43 Change from baseline in Cardiopulmonary Symptoms Domain Score of PAH-SYMPACT® by visit - Full Analysis Set	83
Table 44 Analysis of change from baseline in Cardiopulmonary Symptoms Domain Score of PAH-SYMPACT® at Week 24 - Full Analysis Set (Standard multiple imputation)	84
Table 45 Analysis of change from baseline in Cardiopulmonary Symptoms Domain Score of PAH-SYMPACT® - Full Analysis Set (Standard multiple imputation with ANCOVA model) ..	84
Table 46 Change from baseline in Cognitive/Emotional Impacts Domain Score of PAH-SYMPACT® by visit - Full Analysis Set	86
Table 47 Analysis of change from baseline in Cognitive/Emotional Impacts Domain Score of PAH-SYMPACT® at Week 24 - Full Analysis Set (Standard multiple imputation)	86
Table 48 Analysis of change from baseline in Cognitive/Emotional Impacts Domain Score of PAH-SYMPACT® at Week 24 - Full Analysis Set (Standard multiple imputation with ANCOVA model)	87
Table 49 Analysis of change from baseline in EQ-5D-3L Index Score based on the UK crosswalk (from EQ-5D-5L to EQ-5D-3L) mapping algorithm at Week 24 - Full Analysis Set	87
Table 50 Change from baseline in EQ-5D-5L VAS at Week 24 - Full Analysis Set	88
Table 51 SOTERIA: change from Baseline ^a at Week 24 and Year 1 in clinical parameters ..	91
Table 52 SOTERIA study summary of clinical worsening events	93
Table 53 Analysis of the proportion of WHO FC improvement or maintenance relative to STELLAR baseline - Participants who enrolled to SOTERIA from STELLAR sotatercept arm and continued receiving sotatercept	94
Table 54 Analysis of the proportion of WHO FC improvement or maintenance relative to STELLAR Week 24 - Participants who enrolled to SOTERIA from STELLAR sotatercept arm and continued receiving sotatercept	94
Table 55 Analysis of the proportion of risk status improvement or maintenance relative to STELLAR baseline - Participants who enrolled to SOTERIA from STELLAR sotatercept arm and continued receiving sotatercept	95

Table 56 Analysis of the proportion of risk status improvement or maintenance relative to STELLAR Week 24 - Participants who enrolled to SOTERIA from STELLAR sotatercept arm and continued receiving sotatercept.....	95
Table 57 Analysis of overall survival - Period on STELLAR and SOTERIA - Without adjusting for treatment switch to sotatercept in placebo arm (STELLAR Full Analysis Set population)	97
Table 58 Analysis of overall survival - Period on STELLAR and SOTERIA - Adjusting for treatment switch to sotatercept in placebo arm using rank-preserving structural failure time model (STELLAR Full Analysis Set population).....	97
Table 59 Summary of STELLAR study subgroup analyses interaction test results.....	101
Table 60 Summary of subgroup results for participants with baseline ESC/ERS intermediate-low risk.....	105
Table 61 STELLAR study – within-trial analysis of the comparison of PDE5i + ERA + sotatercept versus PDE5i + ERA + selexipag – participant baseline characteristics.....	108
Table 62 Analysis of change from baseline in risk status at Week 24 – STELLAR study – Within-trial analysis of the comparison of PDE5i + ERA + sotatercept versus PDE5i + ERA + selexipag – Full Analysis Set.....	111
Table 63 Sotatercept doses used in the PULSAR and STELLAR studies	113
Table 64 Overview of outcomes of interest	117
Table 65 Summary of key Bucher ITC results	119
Table 66 Overall summary of treatment-emergent adverse events - Safety Set - Cumulative results (DBPC+LTDB).....	124
Table 67 SOTERIA study overall safety and tolerability results	126
Table 68 SOTERIA study summary of adverse events of special interest	126
Table 69 Summary list of cost-effectiveness studies and model designs.....	135
Table 70 Summary list of cost-effectiveness studies for PAH therapies identified in HTA submissions	142
Table 71 Health state definitions (base case).....	150
Table 72 Features of the economic analysis	154
Table 73 Risk Status Deterioration Between Baseline and Week 24.....	156
Table 74 Baseline characteristics (ITT Population)	158
Table 75 Summary of clinical data sources	159
Table 76 Demographics and baseline characteristics – Participants who received double therapy (sotatercept: ERA+PDE5i) or triple therapy (placebo: ERA+PDE5i+prostacyclin [selexipag])	163
Table 77 Relative risks of risk stratum worsening/improving	164
Table 78 Distribution of patients at baseline and at 24 weeks (Sotatercept).....	166
Table 79 Distribution of patients at baseline and at 24 weeks (Selexipag)	166
Table 80 Analysis of the proportion of risk status improvement or maintenance relative to STELLAR Week 24.....	167
Table 81 Risk stratum transition probabilities for sotatercept observed between week 12-24 in STELLAR.....	168
Table 82 Risk stratum transition probability inputs used in model for sotatercept (long-term): no improvement allowed beyond week 24.....	169
Table 83 Risk stratum transition probability inputs for selexipag (long-term)	169
Table 84 Risk stratum transition probability inputs (off sotatercept and selexipag).....	170
Table 85 Risk stratum transition probability inputs (off sotatercept only)	170
Table 86 Probability of lung/heart transplant	171

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Table 87 A Comparison of baseline patient characteristics of STELLAR and COMPERA .	172
Table 88 Model inputs: HRs for all-cause mortality	177
Table 89 PAH hospitalisation	183
Table 90 Utility by health state UK crosswalk (from EQ-5D-5L to EQ-5D-3L) mapping algorithm utility scores: pooled records (baseline and Week 24)	186
Table 91 Scenario analysis: health state utilities based on WHO Functional Class as a proxy for risk status (123)	187
Table 92 Disutility for administration method	188
Table 93 Disutility for clinical event	188
Table 94 Disutility for caregivers	190
Table 95 Abbreviated summary of costs relevant to UK decision-making.....	191
Table 96 Abbreviated summary of HCRU relevant to UK decision-making.....	192
Table 97 Treatment options	193
Table 98 Treatment distribution by health state and treatment arm	193
Table 99 Treatment package costs	195
Table 100 Patient weight and required dosing	196
Table 101 Calculated drug acquisition costs per cycle and per year	196
Table 102: Sotatercept resource use and costs	197
Table 103: Selexipag regimen resource use and costs	198
Table 104 Intravenous prostacyclin resource use and costs	198
Table 105 Inhaled prostacyclin resource use and costs	200
Table 106 Recurring HCRU costs used in the CEM	202
Table 107 Summary of clinical event costs	203
Table 108 Summary features of QALY shortfall analysis.....	203
Table 109 Summary of QALY shortfall analysis	204
Table 110 Summary of variables applied in the economic model	206
Table 111 Summary of assumptions and corresponding rationale	213
Table 112 Deterministic base case results	216
Table 113 Deterministic net health benefit and net monetary benefit	216
Table 114 Probabilistic base case results	218
Table 115 Ten most impactful parameters from one-way deterministic sensitivity analysis on ICER (severity modifier applied).....	221
Table 116 Scenario Analyses.....	223
Table 117 Scenario analyses deterministic results	226

Figures

Figure 1 Pulmonary arterial hypertension treatment algorithm based on the European society of cardiology/European respiratory society guidelines.....	27
Figure 2 Pulmonary arterial hypertension outline of overall treatment recommendations/treatment algorithm by the World Symposium on Pulmonary Hypertension	28
Figure 3 Schematic of the STELLAR trial design.....	39
Figure 4 Schematic of SOTERIA trial design.....	44

Figure 5 Change from Baseline in 6-Minute Walk Distance (metres) by visit - Full Analysis Set	63
Figure 6 WHO functional class by visit – Full Analysis Set – Cumulative results	71
Figure 7 WHO FC improvement Kaplan-Meier plot – Full Analysis Set	73
Figure 8 Clinical worsening or death Kaplan-Meier plot.....	76
Figure 9 SOTERIA study clinical efficacy results summary	90
Figure 10 Kaplan-Meier curves of overall survival - Period on STELLAR and SOTERIA - Adjusting for treatment switch to sotatercept in placebo arm using rank-preserving structural failure time model (STELLAR Full Analysis Set population)	98
Figure 11 Subgroups analyses of change from baseline in 6-Minute Walk Distance (metres) at Week 24 - Full Analysis Set (Standard multiple imputation)	102
Figure 12 Subgroups analyses of change from baseline in NT-proBNP (pg/mL) at Week 24 - Full Analysis Set (Standard multiple imputation)	103
Figure 13 Subgroups analyses of change from baseline in PVR (dynes*sec/cm ⁵) at Week 24 - Full Analysis Set (Standard multiple imputation)	104
Figure 14 Model diagram – Possible Movements (sotatercept)	153
Figure 15 Model diagram – Possible Movements (selexipag).....	153
Figure 16 Kaplan-Meier curves for survival from Rosenkranz et al. (2023) (79) (original – left; re-created patient-level data per the Guyot et al. algorithm (85) – right).....	173
Figure 17 Fitted independent parametric regression models for low risk–specific survival from Rosenkranz et al. (2023)(79): hazard plots	175
Figure 18 Fitted independent parametric regression models for low risk–specific survival from Rosenkranz et al. (2023)(79): Survival.....	176
Figure 19 Fitted independent parametric regression models for all risk–specific survival from Rosenkranz et al. (2023)(79): hazard plots of Weibull, Gompertz and Gamma	177
Figure 20 Fitted dependent parametric regression models for risk stratum–specific survival from Rosenkranz et al. (2023)(79): hazard plots	178
Figure 21 Fitted dependent parametric regression models for risk stratum–specific survival from Rosenkranz et al. (2023)(68): Survival.....	179
Figure 22 Comparison of HR-based and Dependent Model Projections.....	181
Figure 23 Cost-effectiveness plane scatter plot (1,000 iterations) – Incremental Costs and QALYs (severity modifier applied).....	219
Figure 24 Cost-effectiveness acceptability curve (1,000 iterations) (severity modifier applied)	220
Figure 25 Tornado diagram for ICER impact for one-way deterministic sensitivity analysis (severity modifier applied)	222

Abbreviations

Abbreviation/ acronym	Definition
6MWD	6-minute walk distance
A&E	Accident and emergency department
ActRIIA	Activin receptor type IIA
ADA	Anti-drug antibody
AE	Adverse event
AESI	AE of special interest
AIC	Akaike information criterion
ANCOVA	Analysis of covariance
ARSW	Adjective Rating Scale of Withdrawal
ASE	Asymptotic standard error
AWMSG	All-Wales Medicines Strategy Group
BAT	Best available therapy
BGT	Background therapy
BIC	Bayesian information criterion
BIM	Budget impact model
BMI	Body mass index
BMPR-II	Bone morphogenetic protein receptor type 2
BNF	British National Formulary
BNP	Brain natriuretic peptide
CADTH	Canadian Agency for Drugs and Technologies in Health
CAMPHOR	Cambridge Pulmonary Hypertension Outcome Review
CEAC	Cost-effectiveness acceptability curve
CEDAC	Canadian Expert Drug Advisory Committee
CEM	Cost-effectiveness model
CHD	Congenital heart defect
CI	Confidence interval
CMH	Cochran-Mantel-Haenszel
COVID-19	Coronavirus disease 2019
CSR	Clinical study report
CTD	Connective tissue disease
CTEPH	Chronic thromboembolic pulmonary hypertension
DBPC	Double-blind placebo-controlled period
DMC	Data monitoring committee
DSA	Deterministic sensitivity analysis
EMA	European Medicines Agency
eMIT	Electronic Market Information Tool
ERA	Endothelin receptor antagonists
ERG	Evidence review group
ERS	European Respiratory Society
ESC	European Society of Cardiology

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

EU	European Union
EUR	Euro
FAS	Full analysis set
FC	Functional class
FDA	Food and Drug Administration
FDC	Fixed-dose combination
FRS	French risk score
FSH	Follicle-stimulating hormone
GBP	Great British Pound
GP	General practitioner
HCP	Healthcare professional
HCRU	Healthcare resource use
HES	Hospital Episode Statistics
Hgb	Haemoglobin
H-L	Hodges-Lehmann
hPAH	heritable PAH
HR	Hazard ratio
HRQoL	Health-related quality-of-life
HTA	Health technology assessment
IA	Interim analysis
ICER	Institute for Clinical and Economic Review
ICER	Incremental cost-effectiveness ratio
ICU	Intensive care unit
ICUR	Incremental cost-utility ratio
IgG	Immunoglobulin G
iPAH	idiopathic PAH
IPCW	Inverse Probability of Censoring Weighting
IRP	International Recognition Procedure
IRT	Interactive response technology
ITC	Indirect treatment comparison
ITT	Intention-to-treat
IV	Intravenous
KM	Kaplan-Meier
KOL	Key opinion leader
LL	Lower limit
LOCF	Last observation carried forward
LOS	Length of stay
LTDB	Long-Term Double-Blind
LTFU	Long-term follow-up
LY	Life-year
LYG	Life-years gained
MAIC	Matching-adjusted indirect comparison
MCI	Multicomponent improvement
MedDRA	Medical Dictionary for Regulatory Activities

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

MHRA	Medicines and Healthcare products Regulatory Agency
MLFHQ	Minnesota Living With Heart Failure Questionnaire
mPAP	Mean pulmonary arterial pressure
NCPE	National Centre for Pharmacoeconomics
NHB	Net health benefit
NHS	National Health Service
NHSE	NHS England
NICE	National Institute for Health and Care Excellence
NMB	Net monetary benefit
NO	Nitrous oxide
NT-proBNP	N-terminal pro-brain natriuretic peptide
NYHA	New York Heart Association
ONS	Office for National Statistics
PAH	Pulmonary arterial hypertension
PAH-SYMPACT	Pulmonary arterial hypertension - symptoms and impact questionnaire
PBAC	Pharmaceutical Benefits Advisory Committee
PBS	Pharmaceutical Benefits System
PCA	Prostacyclin analogue
PD	Pharmacodynamic
PDE5i	Phosphodiesterase 5 inhibitor
PGI2	Prostaglandin I2
PK	Pharmacokinetic
PRA	Prostacyclin receptor agonist
PSA	Probabilistic sensitivity analysis
PVR	Pulmonary vascular resistance
QALY	Quality-adjusted life-year
QoL	Quality of life
RCT	Randomised controlled trial
RPSFT	Rank-Preserving Structural Failure time
RR	Risk ratio
RWE	Real-world evidence
SC	Subcutaneous
SD	Standard deviation
SE	Standard error
SEM	Standard error of the mean
SF-36	Short-Form 36-item questionnaire
sGC	Soluble guanylate cyclase
SLR	Systematic literature review
SMC	Scottish Medicines Consortium
SoC	Standard of care
TEAE	Treatment-emergent adverse event
TGF- β	Transforming growth factor beta
TRIPHIC	Translational Research in Pulmonary Hypertension at Imperial College database

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

UK	United Kingdom
UL	Upper limit
ULN	Upper limit of normal
USA	United States of America
WHO	World Health Organization
WSPH	World Symposium on Pulmonary Hypertension
WTP	Willingness-to-pay

Executive summary

Pulmonary arterial hypertension (PAH) is a highly progressive orphan disease that affects the blood vessels in the lungs, leading to right ventricular failure and ultimately death (1, 2). PAH imposes a considerable burden on patient quality of life and daily functioning, with decreased independence and feelings of distress and debilitation due to limited social functioning (3).

The availability of National Health Service (NHS) treatments for PAH in the United Kingdom (UK) is currently defined by NHS commissioning policies. However, the treatment pathway for PAH patients in the UK is aligned with that recommended in the 2022 European Society of Cardiology (ESC)/European Respiratory Society (ERS) guidelines, which utilises an algorithmic approach based upon a patient's risk status (predictive of one-year risk of mortality) and comprises four categories ("low" risk, "intermediate-low" risk, "intermediate-high" risk, and "high" risk). The goal of this approach is to achieve and maintain a low risk status (4).

Based on the available evidence and clinical advice this submission considers the use of sotatercept in intermediate-low risk PAH patients who would otherwise receive selexipag.

Data on the clinical effectiveness of sotatercept in the relevant population comes from the STELLAR study, a multicentre, double-blind, phase 3 trial in which adults with PAH (World Health Organization [WHO] functional class II or III) who were receiving stable background therapy were randomly assigned in a 1:1 ratio to receive subcutaneous sotatercept or placebo every 3 weeks (5).

Sotatercept provided statistically significant and clinically meaningful benefits to adult PAH patients in combination with background therapy, across key endpoints in the STELLAR trial (5). Participants treated with sotatercept in the STELLAR study achieved clinically meaningful and statistically significant improvement in 6-Minute Walking Distance (6MWD; 40.4m) and showed statistically significant reduction in N-terminal pro-brain natriuretic peptide (NT-proBNP), compared to participants treated with placebo in combination with background

therapy. Patients in the sotatercept treatment group were significantly more likely to experience an ESC/ERS risk status improvement at Week 24 than patients in the placebo treatment group. Sotatercept represents a step-change in the treatment of PAH.

A within-trial analysis of the STELLAR trial that specifically compared PDE5i + endothelin receptor antagonist (ERA) patients randomised to receive sotatercept and PDE5i + ERA + selexipag patients randomised to receive placebo found that the patients treated with sotatercept are more likely to have an improvement in ESC/ERS risk status from baseline to Week 24 than patients treated with selexipag as well as less likely to have a worsening in risk status. An indirect treatment comparison (ITC) using the STELLAR trial to provide data on sotatercept and the GRIPHON (NCT01106014) and TRACE (NCT03078907) trials to provide data on selexipag statistically significant results in favour of sotatercept versus selexipag for the outcomes of WHO functional class (FC) worsening, change in 6MWD, and change in NT-proBNP.

SOTERIA is a long-term follow-up study wherein adults with PAH who completed prior sotatercept trials (including STELLAR) receive open-label sotatercept in combination with background PAH therapy for up to 4 years (6). Data collected so far in SOTERIA show that improvements in clinical efficacy measures (in 6MWD, NT-proBNP, improved/maintained WHO FC II, meeting Low French Risk score) were maintained at 1 year during the study's open-label period (6).

Adverse events data from the STELLAR and SOTERIA studies indicate that sotatercept is generally well tolerated with a manageable safety profile for the treatment of PAH.

The cost-effectiveness of sotatercept was assessed in a model based on the ESC/ERS risk classification system. Previous models in PAH typically used a WHO FC-based structure, however this was deemed to be incongruent with current clinical practice and to risk presenting clinically implausible outcomes. The model employed data from STELLAR, published studies, real-world evidence sources and UK clinical expert input where appropriate. The trial provided data on patient baseline characteristics, clinical efficacy (transition probabilities), and health-related quality of life (HRQoL) collected using NICE's preferred measure.

The model results indicate sotatercept to be a cost-effective use of NHS resources, and extensive sensitivity and scenario analyses largely corroborate this finding. In particular, sotatercept demonstrates value for the healthcare system through retaining patients in lower risk states for longer than the current standard of care, which are associated with a reduced resource use burden. The value of these lower risk states to the patient manifests as improved quality-of-life and a lower mortality risk. Assessment of severity modifiers concluded the treatment qualifies for a QALY weighting of 1.2.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Given the innovative nature of the treatment in a disease of substantial unmet need, it is expected that significant value remains uncaptured by the cost-effectiveness analysis presented. As is typical in a rare disease, uncertainty is inevitable and MSD believes a greater acceptance of this should apply for decision-making, given the promising results of the pivotal trial and the step-change sotatercept offers to the treatment pathway with high unmet need.

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

- PAH is a rare and serious disease that affects the blood vessels in the lungs, leading to right ventricular failure and ultimately death. Although rare, it is a chronic, life-limiting condition that affects people of all ages and backgrounds.
- PAH can significantly affect all aspects of daily life with symptoms including dyspnoea, fatigue, and dizziness/syncope, which become increasingly debilitating with disease progression and severity due to loss of functioning and greater disability. PAH imposes a considerable burden on patient quality of life and daily functioning, with decreased independence and feelings of distress and debilitation due to limited social functioning.
- Sotatercept is a first-in-class fusion protein consisting of the Fc domain of human immunoglobulin G (IgG) linked to the extracellular domain of human activin receptor type IIA (ActRIIA), which acts as a ligand trap for selected transforming growth factor beta (TGF- β) superfamily members. Inhibition of these ligands by sotatercept is proposed to help rebalance pulmonary vascular homeostasis toward growth-inhibiting and proapoptotic signalling.
- The availability of NHS treatments for PAH in the UK is currently defined by NHS commissioning policies. However, the treatment pathway for PAH patients in the UK is largely aligned with those recommended in the 2022 ESC/ERS guidelines, which utilises an algorithmic approach based upon a patient's risk status (predictive of one-year risk of mortality) and comprises four categories ("low" risk, "intermediate-low" risk, "intermediate-high" risk and "high" risk). The goal of this approach is to achieve and maintain a low risk status.
- This submission assesses the treatment of adults with PAH who are at intermediate-low risk according to the ESC/ERS four-strata risk rating system (low, intermediate-low, intermediate-high and high). This is a subset of the licensed population for sotatercept, and hence differs from the published final scope for this appraisal which reflects the full population of the marketing authorisation.
- The relevant comparator in this population is triple combination therapy with PDE5i + ERA + selexipag.

1.1 Decision problem

The submission focuses on a subset of the technology's marketing authorisation, specifically patients with **intermediate-low risk** status on the ESC/ERS four-strata risk rating system. Full details and rationale are provided in Table 1.

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	Adults with pulmonary arterial hypertension	Adults with PAH, (WHO Group 1), who are at intermediate-low risk according to the ESC/ERS four-strata risk rating system.	Given the available evidence and advice from clinicians, this submission considers the use of sotatercept in patients with PAH who have an intermediate-low risk status (further details provided in section 1.3.3). - Patients with low risk status as per ESC/ERS guidelines would not require the addition of sotatercept as they would already be in the lowest risk category, which is the treatment goal in PAH management. - Patients who progress to intermediate-high risk status are currently treated with PCA in combination with PDE5i + ERA. The sotatercept-based regimen in these patients is sotatercept in combination with PDE5i + ERA instead of PCA. Robust comparative evidence between these (from either direct trial evidence or indirect comparison) is not available, so this population has been excluded from the decision problem addressed in the company submission. - Evidence to support the treatment of <u>high risk status</u> patients with sotatercept is [REDACTED] as the pivotal STELLAR clinical trial [REDACTED] patients with a high risk status [REDACTED], so this population has

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
			been excluded from the decision problem addressed in the company submission.
Intervention	Sotatercept in combination with established clinical management	Sotatercept in combination with established clinical management	N/A
Comparator(s)	Established clinical management without sotatercept (given in combination as dual or triple therapy): • endothelin receptor antagonists: ambrisentan, bosentan and macitentan • phosphodiesterase 5 inhibitors: sildenafil and tadalafil • prostaglandins/prostanoids: epoprostenol, selexipag, intravenous/inhaled iloprost and treprostinil • soluble guanylate cyclase stimulators: riociguat.	Triple combination therapy of PDE5i + ERA + selexipag	The relevant comparators should reflect current clinical practice. Current reimbursed dual- and triple therapy regimens are specified in the relevant NHS England commissioning policies below (further details provided in Appendix J): • Targeted Therapies for use in Pulmonary Hypertension in Adults, July 2015, A11/P/c. • Riociguat for pulmonary arterial hypertension, February 2017, 16055/P. • Selexipag for treating pulmonary arterial hypertension (adults), December 2018, 170104P. However: • Riociguat-based regimens (usually in combination with ERA) are not a relevant comparator for sotatercept, based on clinical advice that the target patient groups for these treatments do not overlap. Riociguat is currently administered for patients who require further pharmacological targeting of the NO pathway. This position exists for a limited subset of

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
			patients prior to the initiation of a third pharmacological therapy. The most recent National PAH Audit reports fewer than 100 PAH patients (i.e., not including CTEPH patients) currently treated with riociguat (7). • Monotherapy/dual therapy with PDE5i +/- ERA are used to manage patients with ESC/ERS low risk status. Patients with low risk status would not require the addition of sotatercept (or any other intervention) as they have already achieved the lowest risk category, which is the treatment goal of PAH management and so these would not be relevant comparators to sotatercept. Both treatment guidelines and clinical opinion indicate this. • Triple combination therapy including PCA are used in the treatment of patients with ESC/ERS intermediate-high or high risk status i.e., patients who may have failed to respond to, or had subsequent progression with selexipag. Therefore, triple combination therapy of PDE5i + ERA + selexipag, which is used to treat patients with ESC/ERS intermediate-low risk status, is the relevant comparator for the proposed positioning.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Outcomes	The outcome measures to be considered include: - time to clinical worsening - hospitalisations - overall survival - transplant-free survival - exercise capacity (for example 6-minute walking distance) - haemodynamic assessment (e.g., cardiac index, cardiac output, right atrial pressure, pulmonary arterial pressure and pulmonary vascular resistance) - adverse effects of treatment - health-related quality-of-life.	6MWD - PVR - NT-proBNP - MCI - WHO FC - Simplified French Risk Score - Time to death or the first occurrence of a clinical worsening event - HRQoL	As provided by pivotal clinical trial, STELLAR
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. 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. Costs will be considered from an NHS and Personal Social Services perspective.	Cost-effectiveness analysis	N/A
Subgroups to be considered	If the evidence allows, the following subgroups will be considered: WHO functional class 2 or 3	- ESC/ERS four-strata risk status - WHO functional class 2 or 3	Subgroups relevant to the current treatment pathway in PAH in UK clinical practice.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
	• ESC/ERS four-strata risk status • background PAH therapy (dual or triple therapy) • clinical subgroups within group 1 PAH ○ idiopathic ○ heritable ○ associated with drugs/toxins ○ associated with connective tissue disease ○ associated with congenital heart disease.	• background PAH therapy (dual or triple therapy)	
Special considerations including issues related to equity or equality	N/A	N/A	N/A

6MWD: 6-minute walk distance; CTEPH: chronic thromboembolic pulmonary hypertension; ERA: endothelin receptor antagonist; ERS: European Respiratory Society; ESC: European Society of Cardiology; FC: functional class; HRQoL: health-related quality-of-life; MCI: multicomponent improvement; N/A: not applicable; NHS: National Health Service; NO: nitrous oxide; NT-proBNP: N-terminal pro-brain natriuretic peptide; PAH: pulmonary arterial hypertension; PCA: prostacyclin analogue; PDE5i: phosphodiesterase type 5 inhibitor; PVR: pulmonary vascular resistance; UK: United Kingdom; WHO: World Health Organization

1.2 Description of the technology being evaluated

Table 2. Technology being evaluated

UK approved name and brand name	Sotatercept (brand name: Winrevair)
Mechanism of action	The imbalance of the activin and BMPR-II pathway contributes to disease pathogenesis in PAH. Sotatercept consists of a recombinant ActRIIA-Fc fusion protein, which acts as a ligand trap that scavenges excess activin A and other ActRIIA ligands to inhibit activin signalling. As a result, sotatercept is believed to help rebalance the pro-proliferative (ActRIIA-mediated) and antiproliferative (BMPR-II-mediated) signalling to modulate vascular proliferation (8).
Marketing authorisation/CE mark status	The marketing authorisation for sotatercept in the indication detailed in this submission was approved by the MHRA on 27-DEC-2024.
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	This submission is based on a subset of the company's full marketing authorisation which is: Winrevair, in combination with other PAH therapies, is indicated for the treatment of PAH in adult patients with WHO FC II to III, to improve exercise capacity.
Method of administration and dosage	Sotatercept at a starting dose of 0.3 mg/kg with a target dose of 0.7 mg/kg administered subcutaneously every 21 days.
Additional tests or investigations	Hgb and platelet count should be obtained prior to the first dose. Hgb and platelet count should be monitored for the first 5 doses, or longer if values are unstable. Thereafter, Hgb and platelet count should be verified every 3 to 6 months and the dose adjusted if necessary. Hgb and platelet count should be obtained again before reinitiating treatment.
List price and average cost of a course of treatment	The list price of sotatercept is £120.50 per mg, with an annual cost of [REDACTED], using expected patient weight distribution.
Patient access scheme (if applicable)	A patient access scheme (PAS) is available and has been approved by NHS England. The details of the PAS for sotatercept are provided in Appendix I.

ActRIIA: activin receptor type IIA; BMPR-II: bone morphogenetic protein receptor type 2; FC: functional class; Hgb: haemoglobin; MHRA: Medicines and Healthcare products Regulatory Agency; NHS: National Health Service; PAH: pulmonary arterial hypertension; UK: United Kingdom

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

1.3.1 Health condition

PAH is a highly progressive orphan disease that affects the blood vessels in the lungs, leading to right ventricular failure and ultimately death (1, 2). It is primarily caused by imbalance of anti-proliferative and pro-proliferative signals that leads to pulmonary vascular hyperproliferation (9). PAH is a chronic, life-limiting condition that affects people of all ages

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

and backgrounds, especially those who are young and female (10). Prognosis is poor, with rapid and inexorable disease progression resulting in a median survival of only 6-7 years after diagnosis despite available treatment options (1, 11, 12).

The common presenting symptoms of PAH include dyspnoea, fatigue and dizziness/syncope, which become increasingly debilitating (4). As the disease progresses, patients may develop more severe symptoms and ultimately right heart failure (11). The signs and symptoms of PAH are often non-specific, leading to detection and diagnosis at more advanced stages of the disease (13).

PAH is highly debilitating and significantly affects all aspects of daily life, including overall health and vitality (3). The clinical impact of PAH increases with disease progression and severity due to loss of functioning and greater disability (1, 14, 15). PAH imposes a considerable burden on patient quality of life and daily functioning, with decreased independence and feelings of distress and debilitation due to limited social functioning (3). PAH can have a substantial impact on the quality of life of patients, caregivers and family members. In the UK's largest survey into the lived experience of pulmonary hypertension conducted by Pulmonary Hypertension Association UK (PHA UK) in 2023, 96% of participants reported saying that pulmonary hypertension (PH) has had an impact on their overall quality of life, with 56% saying it has had a "major impact" (16). Furthermore, 85% reported an impact on their mental or emotional wellbeing, 79% reported an impact on their ability to work or attend education, and 78% reported saying it has had an impact on their relationships or family. Living with PAH involves a range of challenges and experiences that impact carers' and patients' daily lives. Patient based evidence research from people living with all PH types has identified several key areas of focus:

- **Quality of life/daily activity:** Patients with PAH experience symptoms such as shortness of breath, fatigue, chest pain, and dizziness, which significantly affect their quality of life and daily activities. Data from the 2019 NHS National Audit of Pulmonary Hypertension revealed that over 70% of people with PAH were at least WHO FC III, meaning that the majority have marked limitation of physical activity, with less than ordinary activity causing undue breathlessness and fatigue (17). Moreover, 96% of people living with PH surveyed by the PHA UK in 2023 stated that the disease has had an impact on their overall quality of life, with 50% saying it has had a "major impact" (16).
- **Psychological impact:** Living with PAH can lead to emotional distress, anxiety, and depression due to the chronic nature of the disease, uncertainty about the future, and limitations in physical activity. Data from the 2023 PHA UK survey recorded that 84%

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

of people with PH say the condition has an impact on their mental or emotional wellbeing, with 34% saying it has a “major impact” (16). Similarly, a survey conducted by PHA UK in 2019 identified that 53% of people with PH have experience or been diagnosed with anxiety or depression (18). In addition, 78% of respondents reported that PH has had an impact on their relationships or family (19).

- **Financial & social impact:** Living with PAH can also have a profound impact on personal finances. Data from the PHA UK 2023 survey recorded that 60% of people with PH reported that their condition had an impact on their financial situation, with 25% reporting a “major impact” (16). In addition, 79% of respondents reported that PH had an impact on their ability to attend work or education, with a number reporting that it caused them to have to stop working altogether (16). More specifically, 34% reported that their PH treatment and management had a major impact on their ability to work or attend education.

These limitations in patient functional capacity often result in burdensome emotional stress for caregivers, who experience an overwhelming workload and feelings of concern given the patient’s deteriorating health (3). Caregivers to patients with PAH help alleviate the stress of living with the condition as well as provide practical and emotional support to share the burden including (20):

- Day-to-day support in helping adapt the patient’s home, their lifestyles and new routines to suit the patient and their needs, as well as crisis management when patients come into difficulty (e.g. adverse effects from treatments).
- Attending appointments and preparation ahead of appointments to ensure symptoms and any toxicities are reported accurately, communicating with healthcare professionals (HCPs), and advocacy for patients when dealing with e.g., general practitioners (GPs) / accident and emergency (A&E) staff including providing accurate medical information.
- Treatment administration preparation and logistics by ensuring adherence to care plans, helping to coordinate delivery of medication to patients’ homes (which is often unreliable), and making up the intravenous (IV) treatment (which can be time consuming and which some patients struggle with).

Furthermore, patients who have to receive intravenous therapy (i.e., intravenous prostacyclin) via a pump face additional burdens to their quality of life (20).

Healthcare resource utilisation for PAH is significant, primarily driven by high rates of hospitalisations, investigations/ procedures, outpatient visits and emergency visits (3, 21-23).
Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Data from the UK NHS Digital Hospital Episode Statistics (HES) database show that PAH incurs a heavy burden of healthcare resource utilisation and associated costs; over the 5-year period 2013–2017, total cost was £43.2M, comprising 79% (£33.9M) for inpatient admissions, 19% (£8.3M) for outpatient visits and 2% (£0.9M) for A&E attendances (21).

PAH reduces the ability to work for both patients and caregivers, with major consequences for overall household incomes (14). The unpaid caring activities (described earlier) of caregivers to people with ill health can constitute a substantial quality-of-life burden (24).

Employment and work productivity in patients with PAH are significantly impaired, with studies showing a worsening of employment status over time due to disease progression and sub-optimal treatments (25). Survey data from Canada showed that many patients require caregiver assistance with daily activities as their functional capacity worsens, with caregivers experiencing an average loss of 8 hours worked per week (26).

1.3.2 Treatment pathway

The symptoms of PAH are non-specific but progressive in undiagnosed patients (27). Once patients are eventually diagnosed, they are assessed for disease severity and risk status to inform prognosis and treatment (27). The stratification of PAH patients into different risk categories based on risk algorithms, namely the ESC/ERS risk stratification system, plays a key role in the initial treatment and follow-up management (4).

The availability of treatments for PAH in the UK is currently defined via the NHS commissioning policies listed in Table 3, with the treatment regimens and the criteria for use of those regimens described in further detail in Appendix J.

Table 3 Treatments for PAH currently available for use in the UK as defined in NHS clinical commissioning policies

NHS clinical commissioning policy	Category of interventions relevant for WHO FC II/III PAH covered by the policy	Specific interventions relevant for WHO FC II/III PAH covered by the policy
Clinical Commissioning Policy: Targeted Therapies for Pulmonary Hypertension Functional Class II April 2013 Reference: NHSCB/A11/P/a	PDE5i (phosphodiesterase type 5 inhibitor)	• Sildenafil • Tadalafil
	PDE5i (phosphodiesterase type 5 inhibitor)	• Sildenafil • Tadalafil

NHS clinical commissioning policy	Category of interventions relevant for WHO FC II/III PAH covered by the policy	Specific interventions relevant for WHO FC II/III PAH covered by the policy
Targeted Therapies for use in Pulmonary Hypertension in Adults July 2015 Reference: A11/P/c	ERA (endothelin receptor antagonist)	• Ambrisentan • Bosentan • Macitentan
	PCA (IV/inhaled) (prostacyclin [including prostacyclin analogues and prostacyclin receptor analogues] that are administered intravenously or inhaled) ^a	• Epoprostenol • Iloprost
Riociguat for pulmonary arterial hypertension February 2017 Reference: NHS England: 16055/P	sGC (soluble guanylate cyclase)	• Riociguat
Selexipag for treating pulmonary arterial hypertension (adults) December 2018 NHS England Reference: 170104P	PRA (prostacyclin receptor agonist, oral)	• Selexipag

ERA: endothelin receptor antagonist; FC: functional class; IV: intravenous; NHS: National Health Service; PAH: pulmonary arterial hypertension; PCA: prostacyclin analogue; PDE5i: phosphodiesterase type 5 inhibitor; PRA: prostacyclin receptor agonist; sGC: soluble guanylate cyclase; UK: United Kingdom; WHO: World Health Organization

^a This policy noted that treprostinil will not be routinely commissioned for new patients but funding will continue for patients already established under the terms of the national policy. While it may have a role in the subcutaneous treatment of very small numbers of seriously ill patients with pulmonary hypertension who are not suitable for nebulised or intravenous administration of a prostanoid, commissioners believe that, at around 3-4 times the price of equivalent alternatives, its new pricing structure cannot be justified.

However, based on feedback from UK clinical experts, as well as comments from NHS England in the draft scope consultation comments for this appraisal, the current treatment pathway for PAH patients in the UK reflects that recommended in the 2022 ESC/ERS guidelines (4). These guidelines provide clinical recommendations for the diagnosis and treatment of PAH. Their treatment algorithm centres on a patient's risk status (predictive of one-year risk of mortality), with a goal of achieving and maintaining low risk status. The factors that are used to determine the risk status are the WHO FC, 6MWD and BNP/NT-proBNP levels. According to guidelines, patient risk status is assessed at diagnosis using a 3-strata system, with patients classified as low, intermediate, or high risk. PAH patients should be followed closely and re-evaluated after 3 to 6 months regardless of initial therapy choice. At follow-up, PAH patients should be assessed using a four-strata system: low, intermediate-low, intermediate-high, or high risk (described in Table 4).

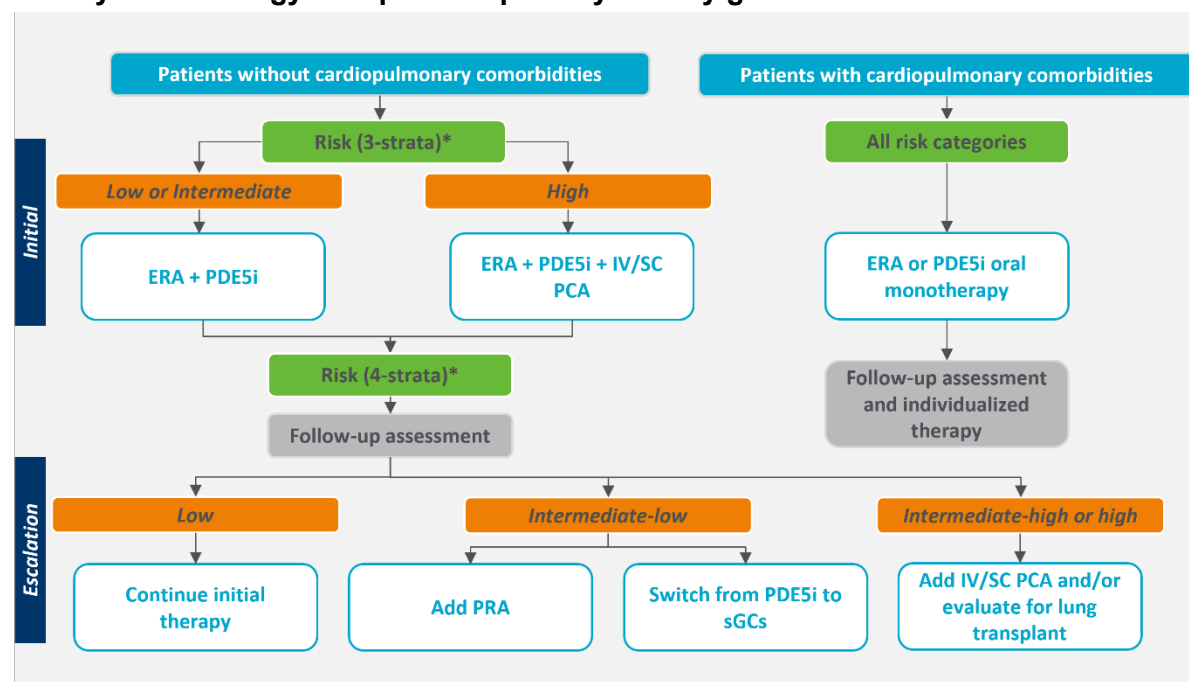
Table 4 2022 ESC/ERS simplified four-strata risk assessment tool for guidance of therapeutic decision-making

Determinants of prognosis	Low risk	Intermediate-low risk	Intermediate-high risk	High risk
Points assigned	1	2	3	4
WHO FC	I or II ^a	-	III	IV
6MWD, m	>440	320–440	165–319	<165
BNP or NT-proBNP, ng/L	<50 <300	50–199 300–649	200–800 650–1100	>800 >1100
Risk is calculated by dividing the sum of all grades by the number of variables and rounding to the next integer.				

6MWD: 6-minute walking distance; BNP: brain natriuretic peptide; ESC: European Society of Cardiology; ERS: European Respiratory Society; FC: functional class; NT-proBNP: N-terminal pro-brain natriuretic peptide; WHO: World Health Organization

^a WHO FC I and II are assigned 1 point as both are associated with good long-term survival.

Initiation of dual combination therapy of ERA and a PDE5i is recommended for patients in low and intermediate risk without cardiopulmonary comorbidities. Patients with inadequate response to treatment should be considered for additional therapy. Therapies patients receive are determined according to their risk status at follow-up visits, the treatment pathway recommended in the 2022 ESC/ERS guidelines (and followed in the UK) are summarised in Figure 1. The more recent 7th World Symposium on Pulmonary Hypertension (WSPH) treatment algorithm for pulmonary arterial hypertension is also aligned with this (summarised in Figure 2) (28).

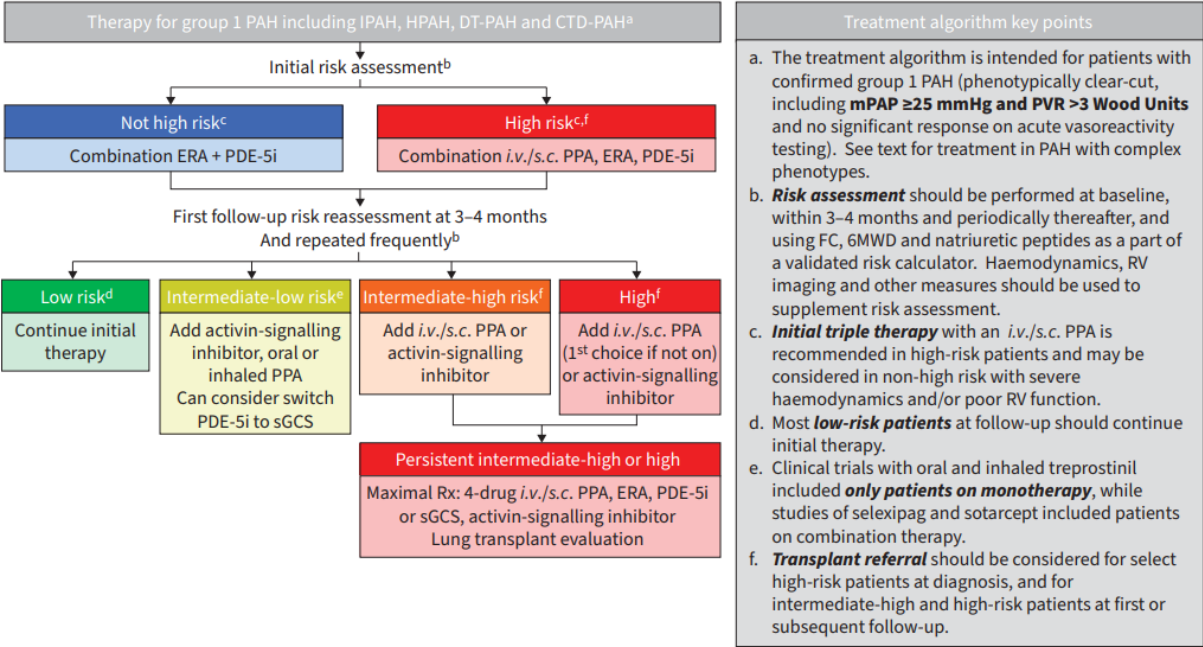
Figure 1 Pulmonary arterial hypertension treatment algorithm based on the European society of cardiology/European respiratory society guidelines

ERA: endothelin receptor antagonist; IV: intravenous; PCA: prostacyclin analogue; PDE5i: phosphodiesterase type-5 inhibitor; PRA: prostacyclin receptor antagonist; SC: subcutaneous; sGCs: soluble guanylate cyclase stimulator

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

*Risk assessment using 3-strata model to determine risk stratification at diagnosis; four-strata model based on determinants of prognosis.
Source: Humbert et al., 2022 (4)

Figure 2 Pulmonary arterial hypertension outline of overall treatment recommendations/treatment algorithm by the World Symposium on Pulmonary Hypertension



6MWD: 6-min walk distance; CTD: connective tissue disease; DT: drug and toxin; ERA: endothelin-1 receptor antagonist; FC: functional class; HPAH: hereditary PAH; IPAH: idiopathic PAH; *i.v.*: intravenous; mPAP: mean pulmonary artery pressure; PAH: pulmonary arterial hypertension; PDE-5i: phosphodiesterase-5 inhibitor; PPA: prostacyclin pathway agent; PVR: pulmonary vascular resistance; RV: right ventricle; Rx: prescription; *s.c.*: subcutaneous; sGCS: soluble guanylyl cyclase stimulator
Source: Chin et al., 2024 (28)

1.3.3 Positioning of sotatercept in the treatment pathway

This submission assesses the initiation of sotatercept in adults with PAH who are at intermediate-low risk according to the ESC/ERS four-strata risk rating system, as an alternative to the current standard of care (SoC) comprising PDE5i + ERA + selexipag. This positioning is based on:

- The current UK clinical practice treatment pathway for patients with PAH (as described in the preceding section 1.3.2).
- The UK marketing authorisation for sotatercept in this indication (as described in section 1.2).
- The clinical effectiveness evidence base for sotatercept (described in section 2).

The rationale for this is described in Table 5, in summary:

- Low risk patients:

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

- Patients in the low risk stratum are already achieving their treatment goal (remaining at low risk) with the current SoC.
- Intermediate-high risk patients:
 - The ESC/ERS treatment guidelines recommend that intermediate-high risk stratum patients be treated with PDE5i + ERA + PCA triple therapy. The pivotal STELLAR study does not provide a randomised comparison of the relevant sotatercept-based regimen in these patients (i.e., triple therapy of sotatercept in combination with background PDE5i + ERA) vs. triple therapy of PCA in combination with background PDE5i + ERA, and ITC for this was found not to be feasible (as described later in section 2.10 and Appendix Q), so evidence for the relative efficacy of sotatercept versus current SoC in these patients is unavailable.
- High risk patients:
 - The pivotal STELLAR clinical trial of sotatercept in this indication recruited [REDACTED] with a baseline high risk status (a consequence of the trial inclusion criterion being WHO FC II or III), [REDACTED], and so the clinical evidence base to support the use of sotatercept in high risk patients [REDACTED].
- Intermediate-low risk patients:
 - Patients in the intermediate-low risk stratum in current clinical practice currently receive SoC comprising PDE5i + ERA + selexipag, and instead should have the option to receive PDE5i + ERA + sotatercept to enable them to reach a low risk state. As mentioned previously, riociguat is currently administered for patients who require further pharmacological targeting of the nitrous oxide (NO) pathway. This position exists for a limited subset of patients prior to the initiation of a third pharmacological therapy, or for patients who cannot tolerate selexipag. The most recent National PAH Audit reports fewer than 100 PAH patients currently treated with riociguat (7), indicating the rarity of this potential comparator.
 - The STELLAR study included sufficient numbers of intermediate-low risk patients to inform analyses ([REDACTED] of the study population).
 - ITC using randomised controlled trial (RCT) data is feasible and has been performed for this comparison (described later in section 2.10). Furthermore, for the endpoint of risk status change, a post-hoc within-trial analysis for this comparison (described later in section 2.8.1) was conducted in order to

appropriately inform the cost-effectiveness analyses. Accordingly, there is comparative evidence to support the use of sotatercept in this group of patients.

- Clinical expert opinion has also indicated that this is an appropriate position for sotatercept in the treatment pathway, based on clinical experience which indicates a limited capacity for selexipag to enable these patients to achieve low risk status and to delay/prevent the need for escalation to parenteral therapies.

Table 5 Positioning of sotatercept in the treatment pathway

Position/PAH population based on ESC/ERS four-strata risk status	SoC regimen	Sotatercept-based regimen	Position to be considered as part of this technology appraisal?	Rationale for considering/not considering the position as part of this technology appraisal
Low	PDE5i + ERA	N/A	No	Patients currently achieving their treatment goal (i.e., low risk) with SoC
Intermediate-low (may include patients who failed to respond, or had sub-optimal response, to background dual therapy with PDE5i + ERA) ^a	PDE5i + ERA + selexipag	PDE5i + ERA + sotatercept	Yes	This submission covers the use of sotatercept in patients at intermediate-low risk to enable them to reach the low risk state or not progress to worse states. Clinical evidence to support this position is available from the STELLAR study as well as an indirect treatment comparison. Clinical expert opinion also supports the use of sotatercept in this position.
Intermediate-high (may include patients who failed to respond, or had sub-optimal response, to background triple therapy of PDE5i + ERA + selexipag)	PDE5i + ERA + PCA (IV/inhaled)	PDE5i + ERA + sotatercept	No	The STELLAR clinical trial does not provide direct evidence for this treatment comparison and an indirect treatment comparison was not found to be feasible, therefore comparative evidence to support this position is currently limited.
High (may include patients who failed to respond, or had sub-optimal response, to prior regimen)	PDE5i + ERA + PCA (IV/inhaled)	PDE5i + ERA + PCA (IV/inhaled) + sotatercept	No	The STELLAR clinical trial of sotatercept in this indication [REDACTED] with a baseline high risk status [REDACTED]; therefore, the evidence to support the use of sotatercept in high risk patients is [REDACTED]

ERA: endothelin receptor antagonist; ESC: European Society of Cardiology; ERS: European Respiratory Society; IV: intravenous; N/A: not applicable; PCA: prostacyclin analogue; PDE5i: phosphodiesterase type-5 inhibitor; SoC: standard of care

^a Existing intermediate-low risk patients who are currently treated with PDE5i + ERA + selexipag should have the option of receiving PDE5i + ERA + sotatercept to enable them to reach a low risk state.

1.3.4 Treatment setting

Treatment with sotatercept will be initiated and managed in specialist centres and administered at home.

1.4 Equality considerations

No equity or equality considerations are anticipated.

2 Clinical effectiveness

Summary of key clinical effectiveness information

- Relevant evidence on the clinical effectiveness of sotatercept in this appraisal were identified via a systematic literature review with searches run on 02-OCT-2024.
- Relevant evidence was identified from the STELLAR and PULSAR RCTs and the SOTERIA long-term follow-up study (5, 6, 8).
- As the Phase 3 STELLAR study supersedes the Phase 2 PULSAR study, and the SOTERIA long-term follow-up study also includes patients originally treated in the PULSAR study, clinical effectiveness evidence from the STELLAR and SOTERIA studies are presented in the submission.

STELLAR randomised controlled trial:

- STELLAR is a multicentre, double-blind, phase 3 trial in which adults with PAH (WHO functional class II or III) who were receiving stable background therapy were randomly assigned in a 1:1 ratio to receive subcutaneous sotatercept (starting dose, 0.3 mg per kilogram of body weight; target dose, 0.7 mg per kilogram) or placebo every 3 weeks (5). The primary endpoint was the change from baseline at week 24 in the 6-minute walk distance (5). Nine secondary end points were also tested (5).
- Results from the STELLAR study are provided from the 06-DEC-2022 database cutoff date; this is the latest available data cutoff date (29).
- Sotatercept provided statistically significant and clinically meaningful benefits to adult PAH patients in combination with background therapy, across key endpoints in the STELLAR trial (29).
- Participants treated with sotatercept achieved clinically meaningful and statistically significant improvement in 6MWD (40.4m) compared to participants treated with placebo in combination with background therapy (29).
- Participants receiving sotatercept showed statistically significant reduction in NT-proBNP from baseline compared to participants receiving placebo in combination with background therapy, indicating clinical improvement of cardiac functional capacity (29).
- Significantly more participants receiving sotatercept met the response criteria for multicomponent improvement (MCI), a novel stringent endpoint prognostic of PAH

improvement, compared to patients receiving placebo in combination with background therapy (29).

- Sotatercept demonstrated a profound reduction in risk of death, clinical worsening, hospitalisation, or transplantation compared to patients treated with placebo in combination with background therapy (29).
- Participants receiving sotatercept were significantly more likely to improve their WHO functional class compared to participants receiving placebo in combination with background therapy (29.4% vs. 13.8%) (29).
- A significantly greater proportion of participants receiving sotatercept achieved or maintained the guideline-aligned goal of low risk status on the French Risk Score compared to participants receiving placebo in combination with background therapy (40.1% vs. 19.5%) (29).
- Patients in the sotatercept treatment group showed significantly greater improvement from baseline to Week 24 than patients in the placebo group in terms of both Pulmonary arterial hypertension - symptoms and impact questionnaire (PAH-SYMPACT®) Physical Impacts ($p=0.012$) and Cardiopulmonary Symptoms ($p=0.037$) (29).
- Sotatercept has been found to be generally well tolerated with a manageable adverse event profile in PAH (29).

STELLAR within-trial analyses

- The comparison of PDE5i + ERA + sotatercept versus PDE5i + ERA + selexipag for the outcome of change from baseline in Risk Status at Week 24 can be informed via a post-hoc within-trial analysis of the STELLAR study including only the participants in each arm treated with these regimens specifically (i.e., PDE5i + ERA patients randomised to receive sotatercept and PDE5i + ERA + selexipag patients randomised to receive placebo).
- This analysis found that patients treated with PDE5i + ERA + sotatercept are more likely to have an improvement in risk status from baseline to Week 24 than patients treated with PDE5i + ERA + selexipag as well as less likely to have a worsening in risk status.

SOTERIA long-term follow-up study:

- SOTERIA is a long-term follow-up study wherein adults with PAH who completed prior sotatercept trials (SPECTRA, PULSAR, ZENITH, HYPERION, STELLAR)

receive open-label sotatercept (0.3 or 0.7 mg/kg once every three weeks [Q3W]) on background PAH therapy for up to 4 years. Safety/tolerability are assessed by adverse events (AEs) and efficacy by 6MWD, NT-proBNP levels, WHO FC, clinical worsening events and French risk score (FRS). Results from the study are available from the 08-NOV-2023 cumulative data cutoff date.

- Long-term treatment with sotatercept (922.2 patient-years of exposure) was generally well tolerated, with less than 5% of participants with TEAEs leading to treatment discontinuation or death.
- Improvements in clinical efficacy measures were maintained at 1 year during this open-label period.

Meta-analysis

- No relevant meta-analyses are feasible. At the current time, data on the effect of treatment with sotatercept in relevant patients are available only from the STELLAR and PULSAR studies, however meta-analysis of data from these two studies is not appropriate due to critical differences in study design.

Indirect and mixed treatment comparisons

- ITC was found to be feasible using the STELLAR trial to provide data on sotatercept and the GRIPHON and TRACE trials to provide data on selexipag.
- ITC was conducted using the Bucher method and found results in favour of sotatercept versus selexipag (where point estimates and both ends of the 95% confidence intervals [CIs] favoured sotatercept) in terms of the outcomes of WHO FC worsening, change in 6MWD and change in NT-proBNP.

2.1 Identification and selection of relevant studies

To identify and select relevant studies, a systematic literature review (SLR) search was carried out in accordance with National Institute for Health and Care Excellence (NICE) guidance, according to a pre-specified protocol to identify relevant studies that investigated sotatercept and any relevant comparator treatments for the indication of interest in this appraisal. Please refer to Appendix B for full details of the process and methods undertaken.

2.2 List of relevant clinical effectiveness evidence

A SLR was performed to identify all relevant published and unpublished RCTs relating to sotatercept as per the final NICE scope.

The SLR identified three studies that investigated sotatercept in the indication of interest to this appraisal (see Appendix B):

- Phase 2 PULSAR RCT (8)
- Phase 3 STELLAR RCT (5)
- Phase 3 SOTERIA long-term follow-up (LTFU) study (6)

As the Phase 3 STELLAR RCT supersedes the Phase 2 PULSAR RCT and the SOTERIA LTFU study includes patients originally treated in the PULSAR study, clinical effectiveness evidence from the STELLAR and SOTERIA studies are presented in the submission (Table 6).

Table 6. Clinical effectiveness evidence

Study	STELLAR (5)	SOTERIA (30)
Study design	Phase 3, randomised, double-blind, placebo-controlled, multicentre, parallel-group study	Phase 3, open-label, LTFU, interventional study
Population	Adults with PAH with WHO functional class II or III, who were receiving stable background therapy	Participants eligible to enrol in the study will have participated in and completed the relevant study requirements of the parent PAH sotatercept clinical studies (SPECTRA, PULSAR, ZENITH, HYPERION, STELLAR) and will transition to SOTERIA as each study concludes
Intervention(s)	Sotatercept	Sotatercept
Comparator(s)	Placebo	N/A
Indicate if study supports application for marketing authorisation	Yes	Yes
Indicate if study used in the economic model	Yes	No
Rationale if study not used in model	N/A	N/A
Reported outcomes specified in the decision problem	• 6MWD • MCI • PVR • NT-proBNP • WHO FC	• Long-term safety and tolerability • 6MWD • PVR • NT-proBNP • WHO FC

6MWD: 6-minute walk distance; FC: functional class; LTFU: long-term follow-up; MCI: multicomponent improvement; N/A: not applicable; NT-proBNP: N-terminal pro-brain natriuretic peptide; PAH: pulmonary arterial hypertension; PVR: pulmonary vascular resistance; WHO: World Health Organization

2.3 Summary of methodology of the relevant clinical effectiveness evidence

The methodologies of the STELLAR and SOTERIA studies are summarised in Table 7 and described in detail in the following subsections.

Table 7. Summary of trial methodology

Study name	STELLAR (5)	SOTERIA (6)
Trial design	Phase 3, multicentre, double-blind, randomised, placebo- controlled clinical trial	Phase 3, open-label, long-term follow-up, interventional study
Eligibility criteria for participants	Eligible patients had a confirmed diagnosis of pulmonary arterial hypertension (idiopathic, heritable, drug-induced, connective tissue disease-associated, or after shunt correction) in WHO functional class II or III, excluding subtypes associated with portopulmonary disease, schistosomiasis, human immunodeficiency virus infection, or veno-occlusive disease. All the patients were receiving stable background therapy for pulmonary arterial hypertension for at least 90 days before enrolment and continued receiving background therapy throughout the trial.	Participants eligible to enrol in the study will have participated in and completed the relevant study requirements of the parent PAH sotatercept clinical studies.
Settings and locations where the data were collected	Multinational trial with locations in: Argentina, Australia, Belgium, Brazil, Canada, Czechia, France, Germany, Israel, Italy, Korea, Republic of, Mexico, Netherlands, New Zealand, Poland, Serbia, Spain, Sweden, Switzerland, United Kingdom, United States	Multinational trial with locations in: Argentina, Australia, Austria, Belgium, Brazil, Canada, Colombia, Croatia, Czechia, Denmark, France, Germany, Greece, Israel, Italy, Korea, Republic of, Mexico, Netherlands, New Zealand, Poland, Portugal, Serbia, Spain, Sweden, Switzerland, Taiwan, United Kingdom, United States
Trial drugs	- Sotatercept at a starting dose of 0.3 mg/kg with a target dose of 0.7 mg/kg administered SC every 21 days plus background PAH therapy - Placebo administered SC every 21 days plus background PAH therapy	Sotatercept SC injection at a dose of 0.3 to 0.7 mg/kg.
Primary outcomes	6MWD	Long-term safety and tolerability
Other outcomes used in the economic	- MCI - PVR - NT-proBNP - WHO FC	6MWD - PVR - NT-proBNP - WHO FC

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Study name	STELLAR (5)	SOTERIA (6)
model/specified in the scope		
Pre-planned subgroups	- Sex (male and female) - PAH etiological subgroups (iPAH, hPAH, drug/toxin-induced PAH, connective tissue disease, congenital heart disease with s/p shunt repair) - Monotherapy vs. double vs. triple combination therapy at baseline - Prostacyclin infusion therapy vs. non-prostacyclin infusion at baseline - Baseline WHO functional class (II or III) - Baseline PVR (≤ 800 or > 800 dynes*sec/cm⁵)	None

6MWD: 6-minute walking distance; FC: functional class; hPAH: heritable PAH; iPAH: idiopathic PAH; MCI: multicomponent improvement; NT-proBNP: N-terminal pro-brain natriuretic peptide; PAH: pulmonary arterial hypertension; PVR: pulmonary vascular resistance; SC: subcutaneously; s/p: systemic to pulmonary; WHO: World Health Organization

2.3.1 STELLAR study - summary of methodology

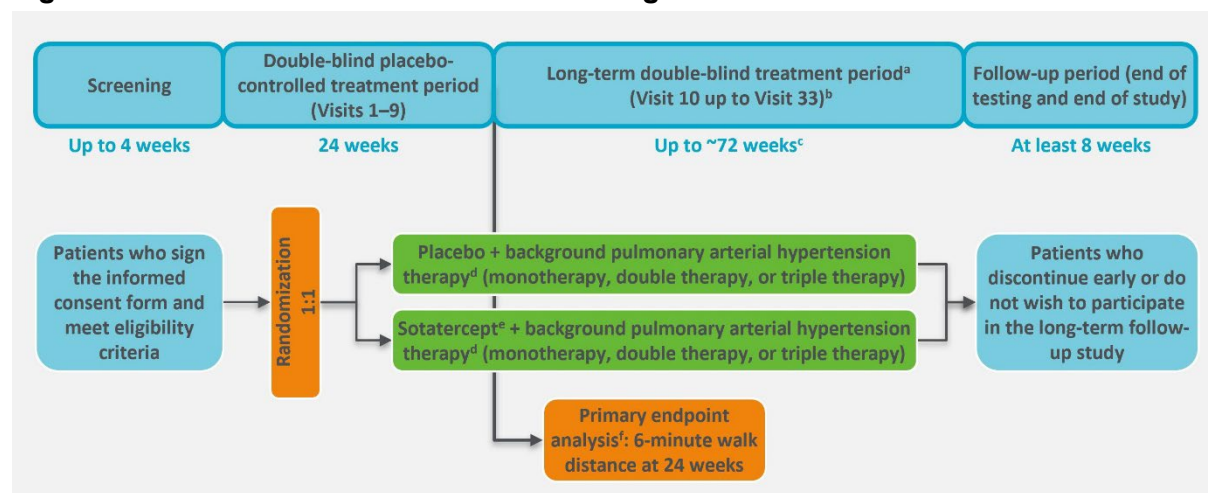
A summary of the methodology of the STELLAR study is provided below, additional details are provided in Appendix J.

Trial design summary

STELLAR (ClinicalTrials.gov ID: NCT04576988) was a phase 3, multicentre, double-blind, randomised, placebo- controlled clinical trial in participants with symptomatic PAH on stable background PAH therapy (5). The trial was conducted across 21 countries, including the UK, enrolling an adult patient population treated with background therapies that are available in UK practice (5). The objectives of the STELLAR trial were to evaluate the efficacy and safety of sotatercept treatment in addition to background therapy versus placebo in addition to background PAH therapy at Week 24 in adults with PAH in WHO FC II or III (5). The primary endpoint of the STELLAR trial was the change from baseline in 6MWD. 6MWD is a clinically validated tool and the most commonly used measure of exercise capacity in PAH (5). The STELLAR trial consisted of a screening period, a double-blind placebo- controlled treatment period, with up to approximately 72 weeks of long-term double-blind treatment period, and a follow-up period of at least 8 weeks (Figure 3) (5).

Participants were permitted to roll over into the ongoing open-label SOTERIA clinical trial (ClinicalTrials.gov ID: NCT04796337) if they experienced a clinical worsening event or completed the 24-week treatment period in the STELLAR trial (5).

Figure 3 Schematic of the STELLAR trial design



^a During the 24-week long-term double-blind treatment period, select study visits were performed as home health care visits.

^b Long-term double-blind treatment period lasted until the last participant randomised completed the double-blind placebo-controlled treatment period, at which point the study was unblinded and participants were invited to roll over into the long-term follow-up study (SOTERIA).

^c Long-term double-blind treatment period duration was estimated based on projected enrolment duration and time required for the last participant to complete the double-blind placebo-controlled treatment period.

^d Background pulmonary arterial hypertension therapy refers to approved pulmonary arterial hypertension-specific medications and consisted of monotherapy or combination therapy with endothelin receptor antagonists, phosphodiesterase-5 inhibitors, soluble guanylate cyclase stimulators, and/or prostacyclin analogues or receptor agonists.

^e Sotatercept at a starting dose of 0.3 mg/kg administered subcutaneously with a target dose of 0.7 mg/kg administered subcutaneously.

^f Primary endpoint analysis was performed after the last participant randomised completed the 24-week double-blind placebo-controlled treatment period.

Source: Adapted from Hoeper et al. 2023 (5)

Randomisation

After the screening period (Figure 3), participants were randomised to one of two treatment arms to receive either sotatercept or a placebo in addition to their background therapy (5). Patients who met all eligibility criteria were stratified by WHO FC class (II or III) and background PAH therapy (mono, double, or triple therapy) (5). A total of 323 participants were randomised in a 1:1 ratio to receive placebo or sotatercept every 3 weeks, in combination with stable background PAH therapy (5).

Dosing information

Sotatercept or placebo was administered as a subcutaneous (SC) injection every 21 days plus background PAH therapy (5). The starting dose of sotatercept or placebo was 0.3 mg/kg at

visit 1 and escalated to the target dose of 0.7 mg/kg at visit 2 and kept at this dose throughout the treatment period (unless haemoglobin or platelet changes warranted dose decrease) (5).

Study population

A total of 434 participants were screened according to inclusion and exclusion criteria (Table 8) (5). Eligible participants for the STELLAR trial were adults presenting with idiopathic or heritable PAH, PAH associated with connective tissue disease (CTD), drug or toxin-induced, post-shunt correction PAH, or PAH presenting at least 1 year following the correction of congenital heart defects (CHD), and currently on background PAH therapy. Participants were required to be classified as WHO FC II or III (5). All the patients were receiving stable background therapy for pulmonary arterial hypertension for at least 90 days before enrolment and continued receiving background therapy throughout the trial (5).

Table 8 Inclusion and exclusion criteria for the STELLAR trial

Inclusion criteria	Exclusion criteria
- Patients ≥ 18 years of age with confirmed diagnosis of PAH (IPAH, HPAH, drug-induced PAH, CTD-PAH and PAH associated with post-shunt correction) with WHO FC II or III. - Stable background PAH therapy at least 90 days prior to screening; for infusion prostacyclin, with dose adjustment within 10% of optimal dose is allowed per medical practice. - RHC performed at screening with PVR of $\geq 400 \text{ dyn}\cdot\text{sec}\cdot\text{cm}^{-5}$ (5 Wood units) and a PCWP or left ventricular end diastolic pressure of $\leq 15 \text{ mm Hg}$. 6MWD ≥ 150 and ≤ 500 metres repeated twice at screening and both values within 15% of each other. - Females of childbearing potential must not be pregnant until 8 weeks after the last dose of the study drug, must use contraception, and must refrain from breastfeeding until 16 weeks after the last dose of the study drug. - Males must agree to use condoms and refrain from sperm donation during the study until 16 weeks after the last dose of the study drug.	- Diagnosis of PH WHO groups 2, 3, 4, or 5. - Diagnosis of the following PAH group 1 subtype: - HIV-associated PAH and PAH associated with portal hypertension. - Exclusions in PAH group I should also include schistosomiasis-associated PAH. - Pulmonary veno-occlusive disease. - Haemoglobin at screening above gender- specific upper limit of normal range, per local laboratory test. - Uncontrolled systemic hypertension. - Baseline systolic blood pressure $<90 \text{ mm Hg}$. - Pulmonary function test values of forced vital capacity $<60\%$ predicted.

6MWD: 6-minute walking distance; CTD: connective tissue disease; FC: functional class; HIV: human immunodeficiency virus; HPAH: hereditary PAH; IPAH: idiopathic PAH; PAH: pulmonary arterial hypertension; PCWP: pulmonary capillary wedge pressure; PVR: pulmonary vascular resistance; PH: pulmonary hypertension; RHC: right heart catheterisation; WHO: World Health Organization
Source: Hoepfer et al. 2023 (5)

Baseline characteristics of trial participants

All study patients

Demographics and baseline disease characteristics were generally comparable in the sotatercept and placebo groups (Table 9), including disease history associated with PAH and baseline laboratory and clinical assessments (5, 29).

Participants in this study were adults with a median age of 48.0 years; 89.2% of participants were white and 79.3% were female (5, 29). The mean time from PAH diagnosis to screening was 8.8 years (5, 29). The three most common WHO PAH (Group 1) subtypes were idiopathic PAH (58.5%), heritable PAH (18.3%) and PAH associated with CTD (14.9%) (5, 29). The proportions of participants with WHO FC II (48.6%) and WHO FC III (51.4%) were similar in both groups (5, 29).

Table 9 Demographics and Baseline Characteristics - Full Analysis Set

	Sotatercept		Placebo		Total	
	n	(%)	n	(%)	n	(%)
Participants in population	163		160		323	
Age (years)						
Participants with data	163		160		323	
Mean	47.6		48.3		47.9	
SD	14.09		15.50		14.79	
Median	47.0		49.0		48.0	
Range	18.0 to 81.0		18.0 to 82.0		18.0 to 82.0	
Age Group, n (%)						
18 to <45	69	(42.3)	65	(40.6)	134	(41.5)
45 to <65	69	(42.3)	66	(41.3)	135	(41.8)
65 to <75	20	(12.3)	24	(15.0)	44	(13.6)
>=75	5	(3.1)	5	(3.1)	10	(3.1)
Sex, n (%)						
Male	34	(20.9)	33	(20.6)	67	(20.7)
Female	129	(79.1)	127	(79.4)	256	(79.3)
Race, n (%)						
White	147	(90.2)	141	(88.1)	288	(89.2)
Black or African American	2	(1.2)	5	(3.1)	7	(2.2)
Asian	1	(0.6)	6	(3.8)	7	(2.2)
American Indian or Alaska Native	0	(0.0)	1	(0.6)	1	(0.3)
Native Hawaiian or Other Pacific Islander	0	(0.0)	1	(0.6)	1	(0.3)
Multiple	2	(1.2)	0	(0.0)	2	(0.6)
Other	5	(3.1)	4	(2.5)	9	(2.8)
Missing	6	(3.7)	2	(1.3)	8	(2.5)
Ethnicity, n (%)						

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

	Sotatercept		Placebo		Total	
	n	(%)	n	(%)	n	(%)
Hispanic or Latino	27	(16.6)	31	(19.4)	58	(18.0)
Not Hispanic or Latino	132	(81.0)	124	(77.5)	256	(79.3)
Not reported	4	(2.5)	5	(3.1)	9	(2.8)
Height (cm)						
Participants with data	163		160		323	
Mean	165.1		165.3		165.2	
SD	8.86		9.14		8.99	
Median	165.0		165.0		165.0	
Range	144.0 to 193.0		147.3 to 195.0		144.0 to 195.0	
Weight (kg)						
Participants with data	163		160		323	
Mean	71.8		72.8		72.3	
SD	19.00		17.64		18.32	
Median	67.0		70.2		68.2	
Range	39.6 to 135.0		38.0 to 141.3		38.0 to 141.3	
BMI (kg/m²)						
Participants with data	163		160		323	
Mean	26.1		26.6		26.4	
SD	5.70		6.06		5.87	
Median	25.3		25.7		25.5	
Range	15.1 to 50.3		15.0 to 54.2		15.0 to 54.2	
Length of time since PAH diagnosis (years)						
Participants with data	163		160		323	
Mean	9.2		8.3		8.8	
SD	7.32		6.70		7.03	
Median	7.5		6.8		7.3	
Range	0.7 to 38.3		0.1 to 40.2		0.1 to 40.2	
WHO Diagnostic pulmonary hypertension Group I: PAH, n (%)						
Idiopathic PAH	83	(50.9)	106	(66.3)	189	(58.5)
Heritable PAH	35	(21.5)	24	(15.0)	59	(18.3)
Drug/toxin-induced PAH	7	(4.3)	4	(2.5)	11	(3.4)
PAH associated with connective tissue disease	29	(17.8)	19	(11.9)	48	(14.9)
PAH associated with simple, congenital systemic to pulmonary shunts at least 1 year following repair	9	(5.5)	7	(4.4)	16	(5.0)

	Sotatercept		Placebo		Total	
	n	(%)	n	(%)	n	(%)
WHO functional classification for symptomatic pulmonary hypertension, n (%)^a						
Class II	80	(49.1)	78	(48.8)	158	(48.9)
Class III	83	(50.9)	82	(51.3)	165	(51.1)
Type of background PAH therapy, n (%)^b						
Mono	9	(5.5)	4	(2.5)	13	(4.0)
Double	56	(34.4)	56	(35.0)	112	(34.7)
Triple	98	(60.1)	100	(62.5)	198	(61.3)
Exercise programme for cardiopulmonary rehabilitation, n (%)						
Yes	17	(10.4)	22	(13.8)	39	(12.1)
No	146	(89.6)	138	(86.3)	284	(87.9)

BMI: body mass index; CSR: clinical study report; FC: functional class; IRT: interactive response technology; PAH: pulmonary arterial hypertension; SD: standard deviation; WHO: World Health Organization
Study: STELLAR (MK-7962-003)

Database Cutoff Date: 06DEC2022

Number of participants: full analysis set

^a One participant, in the sotatercept arm, is counted in the WHO FC II category in this table (in alignment with the CSR Shift from baseline in WHO functional class by Visit entries). To note is that the same participant, in the CSR Disease History Associated with PAH entry is counted in the WHO FC III category. This discrepancy was caused by an incorrect entry in the WHO FC intake form in the IRT and was recorded to match the IRT entry by the site in the study database. For purposes of efficacy and safety analyses, this participant is counted as WHO FC II.

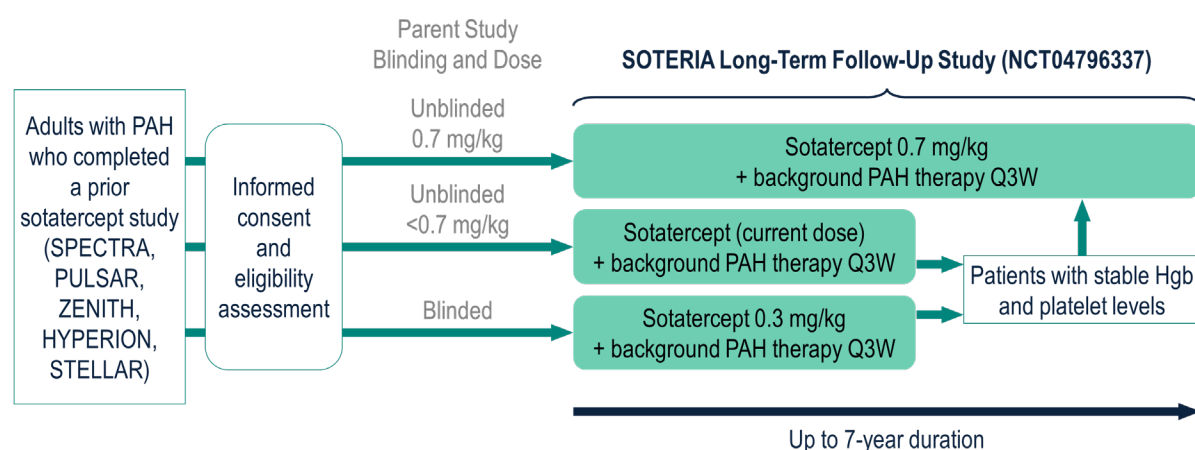
^b One participant, in the placebo arm, is counted in the double therapy category in this table (in alignment with the CSR Disease History Associated with PAH entry). To note is that the same participant, in the CSR Concomitant Standard of Care Medications is counted in the triple therapy category. The discrepancy was caused by an incomplete end date record in the study database for a concomitant PAH therapy. For purposes of efficacy and safety analyses, this participant is counted as being on double therapy.

Source: Hoeper et al. 2023 (5) and MSD 2024 STELLAR CSR (31)

2.3.2 SOTERIA study – summary of methodology

Trial design summary

SOTERIA (ClinicalTrials.gov ID: NCT04796337) is an ongoing phase 3, open-label, interventional, long-term follow-up study including adults with PAH receiving stable background PAH therapy (6). Eligible participants are those who completed a prior sotatercept study (SPECTRA, PULSAR, ZENITH, HYPERION, STELLAR, described in Table 10) without early discontinuation (Figure 4) (6). The primary objectives are to evaluate the long-term safety and tolerability, as measured by AEs, anti-drug antibodies, clinical laboratory assessments, vital signs and 12-lead electrocardiogram (ECG) (6). Secondary objectives are to assess the continued efficacy of sotatercept, as measured by 6MWD, NT-proBNP, WHO FC, pulmonary vascular resistance (PVR), overall survival (OS) and Simplified French Risk Score (6).

Figure 4 Schematic of SOTERIA trial design

Hgb: haemoglobin; PAH: pulmonary arterial hypertension; Q3W: once every three weeks

Source: Adapted from Preston et al. 2025(6)

Table 10 Studies from which participants may enter the SOTERIA study

Study	Description	Status
SPECTRA (NCT03738150)	A Phase 2a, single-arm, open-label, multicentre exploratory study to determine the effects of sotatercept plus SoC in adults with WHO FC III PAH.	Completed
PULSAR (NCT03496207)	a Phase 2, double-blind, randomised, placebo-controlled, parallel-group study of sotatercept plus SoC versus placebo plus SoC in participants with PAH of WHO Group 1, FC II-III.	Completed
ZENITH (NCT04896008)	A phase 3, randomised, double-blind, placebo-controlled study to evaluate sotatercept when added to maximum tolerated background PAH therapy on time to first event of all-cause death, lung transplantation, or PAH worsening-related hospitalisation of ≥ 24 hours, in participants with WHO FC III PAH or WHO FC IV PAH at high risk of mortality.	Stopped ^a .
HYPERION (NCT04811092)	A phase 3, randomised, double-blind, placebo-controlled study to evaluate sotatercept when added to background PAH therapy in newly diagnosed intermediate- or high risk PAH participants	Stopped ^b
STELLAR (NCT04576988)	A Phase 3, randomised, double-blind, placebo-controlled, multicentre, parallel-group study to evaluate the efficacy and safety of sotatercept treatment (plus background PAH therapy) versus placebo (plus background PAH therapy) at 24 weeks in adults with symptomatic PAH who present with idiopathic or heritable PAH, PAH associated with CTD, drug or toxin-induced, post-shunt correction PAH, or PAH presenting at least 1 year following the correction of CHDs, and currently on background PAH therapy.	Completed

CHD: congenital heart defect; CTD: connective tissue disease; FC: functional class; SoC: standard of care; WHO: World Health Organization

^a The ZENITH trial was stopped early based on the strength of its results as observed at the interim analysis of its primary endpoint. An independent data monitoring committee recommended ZENITH be stopped early and all participants be offered the opportunity to receive sotatercept through the SOTERIA open-label extension study.

^b The HYPERION trial lost clinical equipoise and was stopped early. The decision to stop the HYPERION study prior to its scheduled completion was based on the positive results from the interim analysis of the ZENITH trial and a review of the totality of data from the sotatercept clinical programme to date. Participants in the HYPERION trial will have the opportunity to receive sotatercept as part of the SOTERIA study.

Dosing information

Participants previously receiving sotatercept treatment rolling over from a unblinded parent study continued to receive sotatercept at their current dose and if at dose < 0.7 mg/kg SC were titrated up to 0.7 mg/kg SC and remained at 0.7 mg/kg SC for the remainder of the study (6). Participants previously on placebo rolling over from a blinded parent study began sotatercept at a dose of 0.3 mg/kg SC for Visit 1 (6). Dose escalated to 0.7 mg/kg SC at Visit 2 for the remainder of the study (6). Sotatercept was administered subcutaneously once every 21 days (6). Participants completed at least 5 consecutive in-person visits at the clinical site (6). Participants with stable Hgb levels (within 2 g/dL above gender-specific upper limit of normal (ULN) and stable platelet count levels (within normal range or within 10% of previous values) for the previous 3 consecutive visits were eligible for sotatercept self-administration at home from Day 106 (± 3 days) (6). Self-administration was accompanied by a phone call from clinical investigative site to confirm proper storage, dosing, administration, concomitant medication use, and to inquire about AEs (6). For participants opting for self-administration, quarterly on-site visits were conducted (6). Clinical site visits were continued for participants who were ineligible for or opted out of self-administration at home (6).

Baseline characteristics of trial participants

Table 11 SOTERIA study baseline characteristics of trial participants

		All Participants (N=426)
Age, years	Median (range)	49.0 (18, 83)
Sex, n (%)	Female	348 (81.7)
	Male	78 (18.3)
Race*, n (%)	White	380 (89.2)
	Black or African American	9 (2.1)
	Asian	9 (2.1)
	Other	20 (4.7)
Time since PAH diagnosis, years	Median (range)	7.88 (0.5, 39.0)
PAH classification*, n (%)	Idiopathic	247 (58.0)
	Hereditary	69 (16.2)
	Drug or toxin-induced	15 (3.5)
	Associated with connective tissue disease	75 (17.6)
	Associated with corrected congenital shunts	17 (4.0)
WHO functional class, n (%)	I	41 (9.6)
	II	242 (56.8)
	III	118 (27.7)
	IV	25 (5.9)

		All Participants (N=426)
Background PAH therapy, n (%)	Monotherapy	16 (3.8)
	Double therapy	170 (39.9)
	Triple therapy	238 (55.9)

PAH: pulmonary arterial hypertension; WHO: World Health Organization
Data cutoff date 08-NOV-2023
Source: Preston and al. 2025 (6)

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

Information on the statistical analyses and definitions of study groups of the STELLAR and SOTERIA studies are provided in the following subsections.

2.4.1 STELLAR study - statistical analysis plan summary

Details of the statistical analysis plan of the STELLAR study are provided in Appendix K, a summary of the key points is provided below (29).

The sample size was determined using estimates of treatment effect from previous studies reporting on the primary efficacy endpoint (change from baseline in 6-Minute Walk Distance, 6MWD, at Week 24), and a key secondary endpoint (improvement in WHO FC at Week 24) (29). In STELLAR, the Hodges-Lehmann location shift and associated 95% confidence interval (CI) were reported (29).

It is common for continuous endpoints to be summarised based on a measure of central tendency such as the mean or median. However:

- Means are highly influenced by outliers
- Methodology for comparing means typically relies on a normal distribution
- Medians are not sensitive to outliers, but they essentially ignore most of the data
- Nonparametric analyses are generally less sensitive to how imputations are performed for missing data.

In conversations with clinical experts, it was recommended to utilise a nonparametric analysis to address these considerations. The Hodges-Lehmann (H-L) location shift compares the entire distribution of one group to the entire distribution of the other. It does this by looking at the differences between every observation in one group (i.e., sotatercept) paired with every observation in the other group (placebo) (29). The median of these differences is the “Hodges-Lehmann location shift”. Benefits of the H-L location shift include (29):

- It does not require any assumptions about the distribution of the data
- A small number of extreme outliers (e.g., after assigning a very large value for deaths) has no impact on the location shift
- Note that this is not the same as the difference of the individual group medians.

The formal statistical hypothesis testing was conducted by using the Adjective Rating Scale of Withdrawal (ARSW) (29). The ARSW ranks all observations and compares the treatment groups based on the ranks and permits missing/non-existent data to be handled (29). The actual values assigned/imputed are irrelevant as long as the proper ranking is achieved (29). ARSW does not produce meaningful point estimates, hence the use of H-L.

The dichotomous efficacy end points were analysed using a Cochran-Mantel-Haenszel test with data stratified by randomisation factors (29). All continuous secondary end points were tested using the aligned rank stratified Wilcoxon test described above. The secondary end points were also analysed using an analysis of covariance (ANCOVA) model as a supportive analysis with treatment group, randomisation stratification factors and baseline measurement as covariates (29). Time to death or non-fatal clinical worsening event was analysed using a logrank test with randomisation factors as strata (29). Patients without death or non-fatal clinical worsening events were censored at the last visit while on the study (29). A Cox proportional hazards model stratified by randomisation factors was also used to estimate the hazard ratio, and the corresponding 95% CI was provided (29). Statistical analyses conducted on the STELLAR trial endpoints are summarised in Table 12. All efficacy analyses were performed in the intention-to-treat population, this is all the randomised participants (29).

Table 12 Statistical analyses for trial endpoints

Endpoints	Statistical analysis
Primary endpoint: change from baseline in 6MWD at Week 24	• Aligned rank stratified Wilcoxon test with randomisation factors as strata. • Hodges-Lehmann location shift estimate of the between-group difference with 95% CI, with p-value from the aligned rank stratified Wilcoxon test. • Sensitivity analyses were performed if there was a statistically significant treatment effect.
Other endpoints:	• Dichotomous efficacy endpoints were analysed using a Cochran-Mantel-Haenszel test with data stratified by randomisation factors. • ANCOVA model was used as supportive analysis. • Cox proportional hazards model stratified by randomisation factors was used to estimate the hazard ratio and corresponding 95% CI was provided.
	• Participants with missing data due to death were assigned a value to receive the worst rank score.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Endpoints	Statistical analysis
Multiple implications:	Participants with missing data due to non-fatal clinical worsening event were assigned another value to receive the next worst rank score.
Safety:	For selected data, Miettinen and Nurminen method was used to determine 95% CI for between-group difference.

6MWD: 6-minute walking distance; ANCOVA: analysis of covariance; CI: confidence interval.

Source: MSD 2024 STELLAR CSR (29)

2.4.2 SOTERIA study - statistical analysis plan summary

SOTERIA is an open-label, LTFU study to evaluate the safety and efficacy of sotatercept for up to 7 years in participants with PAH previously treated with sotatercept (6). Planned statistical analyses to be conducted are outlined below in Table 13.

Table 13 Summary of SOTERIA study statistical analysis plan

Sample size determination	There is no formal sample size determination for this study. The number of participants in this LTFU study is dependent upon the enrolment in the parent protocols. Approximately 700 participants are expected to enrol in the study.
Population for analysis	The population to be used for statistical analyses of safety and efficacy data is the Full Analysis Set, which is defined as all participants enrolled in the study who have received at least 1 dose of sotatercept in this study.
Statistical methods	In general, baseline is defined as either the EOT assessment from the parent study or Visit 1 assessment in this study depending on the parameter or what data is available for the individual participant. This definition is independent of the treatment the participant received from the parent study. In addition, for participants receiving sotatercept in the parent study, a second baseline will be determined and will be defined as the assessment from the parent study when they took sotatercept for the first time.
Analysis of safety	Adverse events - AEs will be coded using MedDRA. All AEs will be listed and will include, but not necessarily be limited to, the following: verbatim term, MedDRA preferred term, severity, relationship to study medication, relationship to non-study medication, action taken with respect to study medication, flag as to whether the AE is an AESI, and whether or not the AE is treatment-emergent, although not necessarily in that order. TEAEs are defined as those AEs that start or worsen in intensity on or after the first study drug administration up to 8 weeks after the last dose. AEs classified as TEAEs will be summarised. Anti-drug antibodies - Individual ADA data will be listed. The frequency and percentages of all participants testing positive for ADA at any point during the study will be summarised. In addition, for sotatercept ADA, a summary of the titers (median, minimum and maximum value) will be provided by scheduled visit and antibody follow-up visit (as applicable). Clinical laboratory assessments

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

	- Individual laboratory assessments coming from haematology, serum chemistry, FSH and urinalysis will be listed. Such data will be summarised by scheduled visit. Shift tables may be provided for selected laboratory parameters. Graphical summaries of mean $\pm$ SEM (standard error of the mean) may also be provided for selected laboratory parameters. Vital signs - Individual vital sign data will be listed. For each parameter and scheduled visit, the change from baseline will be summarised. Graphical summaries of mean $\pm$ SEM may also be provided for selected vital sign parameters. Electrocardiograms - Individual ECG data will be listed. For each parameter and scheduled visit, the change from baseline will be summarised. Graphical summaries of mean $\pm$ SEM may also be provided for selected ECG parameters.
Analysis of efficacy	The following analyses will be performed on the whole population and may also be performed, as appropriate, by SoC background PAH therapy (mono/double or triple therapy) and WHO FC at baseline. Six-Minute Walk Distance - Changes from baseline for 6MWD will be determined. A line graph of mean $\pm$ SEM will be provided for the raw change from baseline by scheduled time. In addition, the line graph will contain mean $\pm$ SEM estimates by the parent study from which the participant came from. - The change in 6MWD at 1 year in the study, which would generally correspond to Visit 17, will be analysed using an analysis of covariance (ANCOVA model with the parent study the participant belonged to prior to enrolment in this study and baseline 6MWD as covariates). Point estimates of the least squares mean change in 6MWD at 1 year with the corresponding 95% CI will be provided for each parent study and across all enrolled participants. - Similar analysis will also be performed for the change in 6MWD at 2, 3 and 4 years, which would generally correspond to Visits 33, 49 and 69, respectively, in the study. N-Terminal Pro-Hormone B-type Natriuretic Peptide - Changes from baseline for NT-proBNP will be determined. A line graph of mean $\pm$ SEM will be provided for the raw change from baseline by scheduled time. In addition, the line graph will contain mean $\pm$ SEM estimates by the parent study from which the participant came from. - The change in NT-proBNP at 1 year in the study, which would generally correspond to Visit 17, will be analysed using an ANCOVA model with the parent study the participant belonged to prior to enrolment in this study and baseline NT-proBNP as covariates. Point estimates of the least squares mean change in

NT-proBNP at 1 year with the corresponding 95% CI will be provided for each parent study and across all enrolled participants.

- Similar analysis will also be performed for the change in NT-proBNP at 2, 3 and 4 years, which would generally correspond to Visits 33, 49 and 69, respectively, in the study.

WHO Functional Class

- The proportion of participants who improve or show maintenance of WHO FC II at 1 year in the study, which would generally correspond to Visit 17, will be analysed using the binomial test with the normal approximation. In addition, this will also be analysed by parent study the participant belonged to using the binomial test (with normal approximation if applicable).
- Similar analysis will also be performed for the proportion of participants who improve or show maintenance of WHO FC II at 2, 3 and 4 years, which would generally correspond to Visits 33, 49 and 69, respectively, in the study.

Pulmonary Vascular Resistance

- Changes from baseline for PVR will be determined. A line graph of mean $\pm$ SEM will be provided for the raw change from baseline by scheduled time. In addition, the line graph will contain mean $\pm$ SEM estimates by the parent study from which the participant came from.
- The change in PVR at 4 years in the study, which would generally correspond to Visit 69, will be analysed using an ANCOVA model with the parent study and baseline PVR as covariates. Point estimates of the least squares mean change in PVR at 4 years with the corresponding 95% CI will be provided for each parent study and across all enrolled participants.

Overall survival

- Overall survival is defined as the time from the start of the first sotatercept treatment in the individual participant's parent study or in this study, if the participant was completely on placebo in the parent study, to the date of death in this study, regardless of the actual cause of the participant's death. For participants who are still alive at the time of the analysis or who are lost to follow-up, OS time will be censored at the last recorded date that the participant is known to be alive. Participants will be followed for survival information until the study is completed.
- The Kaplan-Meier curve will be generated to represent the OS and the median OS with 95% CI will be provided. This will be done for all enrolled participants as well as by parent study.

Simplified French Risk Score

- The proportion of participants who maintain or achieve a low risk score using the simplified French Risk score calculator at 1 year in the study, which would generally correspond to Visit 17, will be analysed using the binomial test with the normal approximation. In addition, this will also be analysed by parent study the participant belonged to using the binomial test (with normal approximation, if applicable).

	Similar analysis will also be performed for the proportion of participants who maintain or achieve a low risk score using the simplified French Risk score calculator at 2, 3 and 4 years, which would generally correspond to Visits 33, 49 and 69, respectively, in the study.
Pharmacokinetic analysis	PK data from this study will be listed and summarised. If needed, these data may be included into future population PK/PD modelling.
Interim analysis	No interim analyses are planned.
Data monitoring committee	- An external, independent DMC will monitor participant safety throughout the study. - A detailed charter will outline all activities of the DMC (including, but not limited to, the composition of the DMC, the type of data to be reviewed, the DMC responsibilities, and the frequency of meetings). - Internal data review of safety-related data will occur at a pre-planned frequency throughout the study duration.

6MWD: 6-minute walking distance; ADA: anti-drug antibody; AE: adverse event; AESI: adverse event of special interest; ANCOVA: analysis of covariance; CI: confidence interval; DMC: data monitoring committee; ECG: electrocardiogram; EOT: end of treatment; FC: functional class; FSH: follicle-stimulating hormone; LTFU: long-term follow-up; MedDRA: Medical Dictionary for Regulatory Activities; NT-proBNP: N-terminal pro-brain natriuretic peptide; OS: overall survival; PAH: pulmonary arterial hypertension; PD: pharmacodynamic; PK: pharmacokinetic; PVR: pulmonary vascular resistance; SEM: standard error of the mean; SoC: standard of care; TEAE: treatment-emergent adverse event; WHO: World Health Organization
Source: Preston et al. 2025 (6)

2.5 Critical appraisal of the relevant clinical effectiveness evidence

The quality assessment of the STELLAR and SOTERIA studies is provided in Appendix B1.3.

2.6 Clinical effectiveness results of the relevant studies

2.6.1 STELLAR study - clinical effectiveness results in the full study population

STELLAR study available data cutoff date results

Study findings are available from the most recent 06-DEC-2022 data cutoff date (29).

Patient disposition and follow-up duration

The patient disposition and follow-up duration of the STELLAR study at the 06-DEC-2022 last patient last visit database cutoff date are summarised in Table 14 and Table 15. The proportion of participants who completed Visit 9 (Week 24 of the double-blind placebo-controlled [DBPC] treatment period) was greater in the sotatercept group than in the placebo group (29). The proportion of participants who completed the overall study was also greater in the sotatercept group than in the placebo group (29). Fewer participants in the sotatercept group than in the placebo group discontinued study intervention (29). Two (1.2%) participants in the sotatercept group discontinued study intervention due to death and 4 (2.5%) discontinued due to an AE

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

(29). No other discontinuation category had >1 participant in the sotatercept group (29). The most common reasons for discontinuation of study intervention in the placebo group were clinical worsening event, death and other (29).

Table 14 STELLAR study - Disposition of participants – Full Analysis Set

	Sotatercept (N=163)	Placebo (N=160)	Total (N=323)
Total Number of Participants			
Randomised ^a	163 (100.0)	160 (100.0)	323 (100.0)
Completed Visit 9	159 (97.5)	148 (92.5)	307 (95.0)
Completed Study	155 (95.1)	142 (88.8)	297 (92.0)
Discontinued Study Treatment	10 (6.1)	40 (25.0)	50 (15.5)
Withdrew from Study	8 (4.9)	18 (11.3)	26 (8.0)
Primary Reason for Discontinuation of Study Treatment^b			
Adverse Event	4 (2.5)	2 (1.3)	6 (1.9)
Participant request (withdrawal of consent)	1 (0.6)	4 (2.5)	5 (1.5)
Participant's unwillingness or inability to comply with the protocol	1 (0.6)	2 (1.3)	3 (0.9)
Clinical worsening event	1 (0.6)	10 (6.3)	11 (3.4)
Pregnancy	0 (0.0)	1 (0.6)	1 (0.3)
Study terminated by sponsor	0 (0.0)	4 (2.5)	4 (1.2)
Death	2 (1.2)	6 (3.8)	8 (2.5)
Other	1 (0.6)	11 (6.9)	12 (3.7)
Primary Reason for Withdrawal from Study^b			
Adverse Event	3 (1.8)	1 (0.6)	4 (1.2)
Participant request (withdrawal of consent)	2 (1.2)	6 (3.8)	8 (2.5)
Participant's unwillingness or inability to comply with the protocol	1 (0.6)	2 (1.3)	3 (0.9)
Clinical worsening event	0 (0.0)	2 (1.3)	2 (0.6)
Death	2 (1.2)	6 (3.8)	8 (2.5)
Other	0 (0.0)	1 (0.6)	1 (0.3)

Database Cutoff Date: 06DEC2022

Number of participants: full analysis set

^a Excludes one participant (placebo group) who was inadvertently randomised and did not receive any study medication. This participant is also excluded from all analyses.

^b Participants may have only one primary reason for discontinuation and withdrawal.

Source: MSD 2024 STELLAR CSR (29)

Table 15 STELLAR study - Summary of follow-up time (Weeks) - Full Analysis Set

	Sotatercept	Placebo	Total
Participants in population	163	160	323
Theoretical (Maximum) Follow-up Time (Weeks)^a			
Participants with data	163	160	323
Mean	53.3	53.7	53.5
SE	0.8	0.9	0.6
Median	51.1	51.8	51.3

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

	Sotatercept	Placebo	Total
Q1 to Q3	44.1 to 59.9	43.9 to 60.7	44.0 to 60.1
Range	39.9 to 85.9	39.9 to 94.1	39.9 to 94.1
Actual Follow-up Time (Weeks)^b			
Participants with data	163	160	323
Mean	46.4	40.1	43.3
SE	0.9	1.1	0.7
Median	45.3	39.0	42.0
Q1 to Q3	37.3 to 54.1	32.4 to 48.1	36.1 to 51.1
Range	16.7 to 80.3	1.3 to 81.4	1.3 to 81.4

SE: standard error; Q1: first quartile; Q3: third quartile

Study: STELLAR (MK-7962-003)

^a Calculated from the date of randomisation until the database cutoff date, for all participants. This represents the length of the trial.

^b Calculated from the date of randomisation until the date of death, the date of last contact, or the database cutoff date if the participant is still alive.

Database Cutoff Date: 06DEC2022

Source: MSD 2024 STELLAR CSR (29)

Extent of exposure

Study drug duration of exposure in the STELLAR study at the time of the 06-DEC-2022 database cutoff date is shown in Table 16. The median duration of exposure to study intervention was longer in the sotatercept group (313.0 days) than in the placebo group (273.0 days; Table 17) (29).

Table 16 STELLAR – Study drug exposure by duration – Safety set

	Sotatercept			Placebo		
	(N=163)			(N=160)		
	n	(%)	Person-months	n	(%)	Person-months
Duration of Exposure						
>0 m	163	(100.0)	1,710.9	160	(100.0)	1,453.0
>=1 m	163	(100.0)	1,710.9	157	(98.1)	1,451.0
>=3 m	162	(99.4)	1,708.9	155	(96.9)	1,447.6
>=6 m	158	(96.9)	1,690.9	142	(88.8)	1,386.2
>=12 m	42	(25.8)	597.2	26	(16.3)	361.7
>=18 m	2	(1.2)	36.8	1	(0.6)	18.6
>=24 m	0	(0.0)	0.0	0	(0.0)	0.0

Study: STELLAR (MK-7962-003)

Database Cutoff Date: 06DEC2022

Number of participants: safety set

Each participant is counted once on each applicable duration category row.

Duration of exposure is the time from the first dose date to the last dose date, accounting for the treatment window period (last dose date - first dose date + 21).

The duration of exposure in-person time is calculated as the total of all individual participant duration values in days divided by days per month (1 Month = 30.4367 Days)

Source: MSD 2024 STELLAR CSR (29)

Table 17 STELLAR – Study drug exposure – Safety set – Cumulative results (DBPC+LTDB)

	Sotatercept (N=163)	Placebo (N=160)	Total (N=323)
Duration of exposure (days)^a			
n	163	160	323
Mean (SD)	319.5 (84.60)	276.4 (95.78)	298.1 (92.71)
Median	313.0	273.0	290.0
Q1, Q3	256, 377	218, 335	252, 357
Min, Max	61, 561	21, 566	21, 566
Number of treatment visits			
n	163	160	323
Mean (SD)	14.6 (3.99)	12.9 (4.45)	13.7 (4.31)
Median	14.0	13.0	13.0
Q1, Q3	12, 17	10, 16	11, 17
Min, Max	3, 27	1, 27	1, 27
Total (cumulative) dose administered (mg)			
n	163	160	323
Mean (SD)	694.7 (286.57)	0.0 (0.00)	350.6 (402.90)
Median	644.5	0.0	100.1
Q1, Q3	514.2, 848.4	0.0, 0.0	0.0, 646.2
Min, Max	90.9, 1587.2	0.0, 0.0	0.0, 1587.2
Number of subjects who reached 0.7 (mg/kg) dose, n (%)	162 (99.4)	0 (0.0)	162 (50.2)

DBPC: double-blind placebo-controlled period; LTDB: long-term double-blind period; SD: standard deviation, Q1: first quartile; Q3: third quartile.

Percentages are calculated as (n/N)*100.

The mg of dose administered is calculated as a function of the individual's weight.

^a Duration of exposure = last dose date - first dose date + 21.

Source: MSD 2024 STELLAR CSR (29)

Overall summary of results

An overall summary of the STELLAR full study population results is presented in Table 18 below. Results for the subgroup of participants of the STELLAR study who had ESC/ERS intermediate-low risk at baseline are summarised in Table 61 in Section 2.8.1 (29).

Table 18 STELLAR study - Overall summary of results - Full study population

	Sotatercept	Placebo
6-Minute Walk Distance		
<i>Change from baseline (metres) at Week 24</i>		
n	158	147
Mean (SD)	40.3 (64.18)	-0.6 (69.54)
Median	35.0	6.0
Min, Max	-245.0, 279.0	-291.5, 238.5
Missing	5	13

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

	Sotatercept	Placebo
Percentage change from baseline at Week 24		
n	158	147
Mean (SD)	11.6 (21.38)	-0.1 (18.45)
Median	8.5	1.2
Min, Max	-85.4, 138.1	-62.8, 80.7
Missing	5	13
Analysis of change from baseline in 6-Minute Walk Distance (metres) at Week 24		
Median estimate (Min, Max)	34.3 (33.0, 35.5)	1.0 (-1.0, 3.5)
Hodges-Lehmann location shift (ASE)	40.4 (6.70)	N/A
95% CI of Hodges-Lehmann location shift	(27.28, 53.53)	N/A
Wilcoxon p-value	<.001	N/A
Multicomponent improvement		
Improvement of WHO FC or maintenance of WHO FC II, n	163	159
Yes, n (%)	115 (70.6)	82 (51.6)
No, n (%)	48 (29.4)	77 (48.4)
p-value	<.001	N/A
Missing due to COVID-19	0	1
Improvement in NT-proBNP $\geq 30\%$ or maintenance/achievement of NT-proBNP level < 300 ng/L, n	162	159
Yes, n (%)	138 (85.2)	66 (41.5)
No, n (%)	24 (14.8)	93 (58.5)
p-value	<.001	N/A
Missing due to COVID-19	1	1
Improvement in 6MWD ≥ 30 metres, n	163	159
Yes, n (%)	88 (54.0)	35 (22.0)
No, n (%)	75 (46.0)	124 (78.0)
p-value	<.001	N/A
Missing due to COVID-19	0	1
Multicomponent improvement of all three criteria above, n	162	159
Yes, n (%)	64 (39.5)	17 (10.7)
No, n (%)	98 (60.5)	142 (89.3)
p-value	<.001	N/A
Missing due to COVID-19	1	1
Pulmonary vascular resistance		
Change from baseline at Week 24		
n	154	141
Mean (SD)	-234.5 (335.98)	41.9 (220.62)
Median	-172.0	24.0
Min, Max	-1792.0, 656.0	-608.0, 848.0

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

	Sotatercept	Placebo
Missing	9	19
Percentage change from baseline at Week 24		
n	154	141
Mean (SD)	-25.0 (27.17)	7.9 (34.32)
Median	-26.7	2.8
Min, Max	-81.5, 92.1	-60.3, 165.6
Missing	9	19
Analysis of change from baseline in PVR (dynes*sec/cm⁵) at Week 24 (Standard multiple imputation)		
Median estimate (Min, Max)	-160.6 (-176.0, -152.0)	32.8 (24.0, 40.0)
Hodges-Lehmann location shift (ASE)	-231.8 (27.61)	N/A
95% CI of Hodges-Lehmann location shift	(-285.88, -177.63)	N/A
Wilcoxon p-value	<.001	N/A
Analysis of change from baseline in PVR (dynes*sec/cm⁵) at Week 24 (Standard multiple imputation with ANCOVA model)		
LS mean (SE)	-214.540 (24.168)	81.356 (24.753)
95% CI of LS Mean	(-261.909, -167.171)	(32.840, 129.872)
LS mean diff (SE)	-295.896 (34.120)	N/A
95% CI of LS mean diff	(-362.771, -229.020)	N/A
Shapiro-Wilk p-value	<.001	N/A
NT-proBNP		
Change from baseline at Week 24		
n	148	139
Mean (SD)	-722.3 (2479.36)	346.6 (1276.45)
Median	-224.0	44.0
Min, Max	-28290.0, 660.0	-1633.0, 9055.0
Missing	15	21
Percentage change from baseline at Week 24		
n	148	139
Mean (SD)	-46.0 (58.74)	51.7 (180.06)
Median	-62.6	17.8
Min, Max	-99.8, 392.3	-78.8, 1929.4
Missing	15	21
WHO FC improvement		
Improvement of WHO FC, n	163	159
Yes, n (%)	48 (29.4)	22 (13.8)
No, n (%)	115 (70.6)	137 (86.2)
p-value	<.001	N/A
Missing due to COVID-19	0	1
Time to first clinical worsening or death		

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

	Sotatercept	Placebo
Total number of participants who experienced at least one clinical worsening event or death, n (%)	11 (6.7)	42 (26.3)
Total number of clinical worsening events or death	11	45
<i>Time to first occurrence of clinical worsening events or death (weeks)</i>		
Logrank test p-value	<.001	N/A
Hazard ratio (95% CI)	0.23 (0.12, 0.44)	N/A
<i>Assessment of first occurrence of clinical worsening events or death, n (%)</i>		
Death	2 (1.2)	6 (3.8)
Worsening-related listing for lung and/or heart transplant	1 (0.6)	1 (0.6)
Need to initiate rescue therapy with an approved PAH therapy or the need to increase the dose of infusion prostacyclin by 10% or more	2 (1.2)	17 (10.6)
Need for atrial septostomy	0 (0.0)	0 (0.0)
PAH-specific hospitalisation (≥24 hours)	1 (0.6)	7 (4.4)
Deterioration of PAH	5 (3.1)	15 (9.4)
Simplified French Risk Score		
Any maintain or achieve a low risk score using the simplified French Risk Score calculator at Week 24 vs. baseline, n	162	159
Yes	65 (40.1)	31 (19.5)
No	97 (59.9)	128 (80.5)
p-value	<.001	N/A
Missing due to COVID-19	1	1
PAH-SYMPACT® – Physical Impacts Domain		
<i>Change from baseline in Physical Impacts Domain Score of PAH-SYMPACT at Week 24</i>		
n	86	92
Mean (SD)	-0.3 (0.67)	0.0 (0.60)
Median	-0.1	0.0
Min, Max	-2.7, 1.0	-2.4, 2.3
Missing	77	68
<i>Analysis of change from baseline in Physical Impacts Domain Score of PAH-SYMPACT® at Week 24 (Standard multiple imputation)</i>		
Median estimate (Min, Max)	-0.12 (-0.15, 0.00)	0.01 (0.00, 0.14)
Hodges-Lehmann location shift (ASE)	-0.26 (0.114)	N/A
95% CI of Hodges-Lehmann location shift	(-0.482, -0.034)	N/A

	Sotatercept	Placebo
Wilcoxon p-value	0.012	N/A
<i>Analysis of change from baseline in Physical Impacts Domain Score of PAH-SYMPACT® at Week 24 (Standard multiple imputation with ANCOVA model)</i>		
LS mean (SE)	-0.198 (0.113)	0.192 (0.109)
95% CI of LS mean	(-0.419, 0.023)	(-0.022, 0.406)
LS mean diff (SE)	-0.390 (0.157)	N/A
95% CI of LS mean diff	(-0.697, -0.083)	N/A
p-value	0.013	N/A
Shapiro-Wilk p-value	0.002	N/A
PAH-SYMPACT® – Cardiopulmonary Symptoms Domain		
<i>Change from baseline in Cardiopulmonary Symptoms Domain Score of PAH-SYMPACT® at Week 24</i>		
n	96	93
Mean (SD)	-0.1 (0.44)	0.0 (0.39)
Median	-0.1	-0.0
Min, Max	-1.6, 1.0	-1.4, 1.5
Missing	67	67
<i>Analysis of change from baseline in Cardiopulmonary Symptoms Domain Score of PAH-SYMPACT® at Week 24 (Standard multiple imputation)</i>		
Median estimate (Min, Max)	-0.11 (-0.14, -0.06)	-0.01 (-0.03, 0.02)
Hodges-Lehmann location shift (ASE)	-0.13 (0.062)	N/A
95% CI of Hodges-Lehmann location shift	(-0.249, -0.007)	N/A
Wilcoxon p-value	0.037	N/A
<i>Analysis of change from baseline in Cardiopulmonary Symptoms Domain Score of PAH-SYMPACT® at Week 24 (Standard multiple imputation with ANCOVA model)</i>		
LS mean (SE)	-0.051 (0.076)	0.162 (0.077)
95% CI of LS mean	(-0.200, 0.097)	(0.010, 0.314)
LS mean diff (SE)	-0.213 (0.104)	N/A
95% CI of LS mean diff	(-0.418, -0.008)	N/A
Shapiro-Wilk p-value	<.001	N/A
PAH-SYMPACT® – Cognitive/Emotional Impacts Domain		
<i>Change from baseline in Cognitive/Emotional Impacts Domain Score of PAH-SYMPACT® at Week 24</i>		
n	86	92
Mean (SD)	-0.1 (0.55)	0.1 (0.75)

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

	Sotatercept	Placebo
Median	0.0	0.0
Min, Max	-2.0, 1.8	-2.3, 2.5
Missing	77	68
Analysis of change from baseline in Cognitive/Emotional Impacts Domain Score of PAH-SYMPACT® at Week 24 (Standard multiple imputation)		
Median estimate (Min, Max)	0.000 (0.000, 0.000)	0.000 (0.000, 0.001)
Hodges-Lehmann location shift (ASE)	-0.15 (0.121)	N/A
95% CI of Hodges-Lehmann location shift	(-0.383, 0.091)	N/A
Wilcoxon p-value	0.186	N/A
Analysis of change from baseline in Cognitive/Emotional Impacts Domain Score of PAH-SYMPACT® at Week 24 (Standard multiple imputation with ANCOVA Model)		
LS mean (SE)	0.004 (0.115)	0.253 (0.112)
95% CI of LS mean	(-0.222, 0.230)	(0.034, 0.472)
LS mean diff (SE)	-0.249 (0.159)	N/A
95% CI of LS mean diff	(-0.561, 0.063)	N/A
Shapiro-Wilk p-value	<.001	N/A
EQ-5D-3L Index Score		
Analysis of change from baseline in EQ-5D-3L Index Score based on the UK crosswalk (from EQ-5D-5L to EQ-5D-3L) mapping algorithm at Week 24		
n	90	89
Mean (SD)	0.02 (0.18)	-0.01 (0.18)
Median	0.00	0.00
Min, Max	-0.82, 0.46	-0.71, 0.75
Missing	73	71
Analysis of change from baseline in EQ-5D-3L Index Score Based on the UK crosswalk (from EQ-5D-5L to EQ-5D-3L) mapping algorithm at Week 24 (ANCOVA)		
LS mean (SE)	0.02 (0.02)	-0.01 (0.02)
95% CI of LS mean	(-0.02, 0.05)	(-0.05, 0.02)
LS mean diff (SE)	0.03 (0.02)	N/A
95% CI of LS mean diff	(-0.02, 0.08)	N/A
Shapiro-Wilk p-value	<.001	N/A
EQ-5D-5L VAS		
Change from baseline in EQ-5D-5L VAS at Week 24		
n	90	89
Mean (SD)	4.33 (16.58)	-0.87 (19.00)
Median	5.00	0.00

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

	Sotatercept	Placebo
Min, Max	-36.00, 71.00	-65.00, 74.00
Missing	73	71
Change from baseline in EQ-5D-5L VAS at Week 24 (ANCOVA)		
LS mean (SE)	4.37 (1.74)	-0.91 (1.73)
95% CI of LS mean	(0.94, 7.80)	(-4.33, 2.51)
LS mean diff (SE)	5.28 (2.39)	N/A
95% CI of LS mean diff	(0.56, 10.00)	N/A
p-value	0.029	N/A
Shapiro-Wilk p-value	0.002	N/A

6MWD: 6-minute walking distance; ANCOVA: analysis of covariance; ASE: asymptotic standard error; CI: confidence interval; COVID-19: coronavirus disease 2019; EQ-5D-3L: EuroQol 5-Dimension 3-Level; EQ-5D-5L: EuroQol 5-Dimension 5-Level; FC: functional class; LS: least squares; N/A: not applicable; NT-proBNP: N-terminal pro-brain natriuretic peptide; PAH: pulmonary arterial hypertension; PAH-SYMPACT: Pulmonary Arterial Hypertension-Symptoms and Impact Questionnaire; PVR: pulmonary vascular resistance; SD: standard deviation; SE: standard error; VAS: visual analogue scale; WHO: World Health Organization
Source: MSD 2024 STELLAR CSR (29)

Six-minute walk distance

In adults with PAH (WHO Group 1, FC II or III) on background PAH therapy, the addition of sotatercept, compared with the addition of placebo, led to a significant improvement from baseline in 6MWD at Week 24 (29). The median treatment difference (H-L location shift) between the sotatercept and placebo groups was 40.4 metres ($p < 0.001$) (Table 19) (29). At Week 24, there was a significantly greater proportion of patients with an improvement in 6MWD of at least 30 metres in the patients in the sotatercept arms than in the placebo arm (54% compared to 22%, Table 20) (29). Descriptive statistics of the observed mean change from baseline in 6MWD at Week 24 was higher in the sotatercept group (40.3 metres) compared with the placebo group (-0.6 metres) (Table 21), a trend that continued to be apparent in visits up to Week 84 (Figure 5) (29).

Table 19 Analysis of change from baseline in 6-Minute Walk Distance (metres) at Week 24 - Full Analysis Set - (Standard Multiple Imputation)

	Sotatercept (N=163)	Placebo (N=160)
Median estimate (Min, Max) ^a	34.3 (33.0, 35.5)	1.0 (-1.0, 3.5)
Hodges-Lehmann location shift (ASE) ^b	40.4 (6.70)	N/A
95% CI of Hodges-Lehmann location shift	(27.28, 53.53)	N/A
Wilcoxon p-value ^c	<.001	N/A

ASE: asymptotic standard error; CI: confidence interval; N/A: not applicable

Study: STELLAR (MK-7962-003)

Database Cutoff Date: 06DEC2022

Number of participants: full analysis set

^a Average, minimum and maximum of the medians across the imputed datasets if missing data imputed.

^b Hodges-Lehmann location shift from placebo estimate (median of all paired differences).

^c Wilcoxon p-value refers to p-value from the aligned rank stratified Wilcoxon test with randomisation factors as strata.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Change from baseline in 6MWD at Week 24 for participants who died was assigned a value of -2000 metres to receive the worst rank. Change from baseline in 6MWD at Week 24 for participants who have missing data due to a non-fatal clinical worsening event was imputed to -1000 metres to receive the next worst rank.
Source: MSD 2024 STELLAR CSR (29)

Table 20 Analysis of change from baseline in 6-Minute Walk Distance (by at least 30 metres) at Week 24 - Full Analysis Set

	Sotatercept (N=163)	Placebo (N=160)
Improvement in 6MWD by at least 30 metres, m	163	159
Yes, n (%)	88 (54.0)	35 (22.0)
No, n (%)	75 (46.0)	124 (78.0)
p-value [1]	<.001	N/A
Missing due to COVID-19	0	1

6MWD: 6-minute walking distance; COVID-19: coronavirus disease 2019; N/A: not applicable
Source: MSD 2024 STELLAR CSR (29)

Table 21 Change from Baseline in 6-Minute Walk Distance (metres) - Full Analysis Set

	Sotatercept (N=163)	Placebo (N=160)
Baseline		
Baseline		
n	163	160
Mean (SD)	397.6 (84.28)	404.7 (80.59)
Median	417.0	427.1
Min, Max	160.5, 497.5	151.5, 514.5
Missing	0	0
Week 24		
Week 24		
n	158	147
Mean (SD)	439.9 (97.16)	409.2 (100.54)
Median	454.0	425.0
Min, Max	42.0, 691.0	63.0, 647.0
Missing	5	13
Change from baseline at Week 24		
n	158	147
Mean (SD)	40.3 (64.18)	-0.6 (69.54)
Median	35.0	6.0
Min, Max	-245.0, 279.0	-291.5, 238.5
Missing	5	13
Percentage change from baseline at Week 24		
n	158	147
Mean (SD)	11.6 (21.38)	-0.1 (18.45)
Median	8.5	1.2
Min, Max	-85.4, 138.1	-62.8, 80.7
Missing	5	13
Week 84^a		

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

	Sotatercept (N=163)	Placebo (N=160)
Week 84		
n	2	1
Mean (SD)	367.3 (34.44)	328.0 (NE)
Median	367.3	328.0
Min, Max	342.9, 391.6	328.0, 328.0
Missing	161	159
Change from baseline at Week 84		
n	2	1
Mean (SD)	41.6 (31.86)	-161.5 (NE)
Median	41.6	-161.5
Min, Max	19.1, 64.1	-161.5, -161.5
Missing	161	159
Percentage Change from baseline at Week 84		
n	2	1
Mean (SD)	12.7 (9.68)	-33.0 (NE)
Median	12.7	-33.0
Min, Max	5.9, 19.6	-33.0, -33.0
Missing	161	159

NE: Not Evaluable; SD: standard deviation

Study: STELLAR (MK-7962-003)

Database Cutoff Date: 06DEC2022

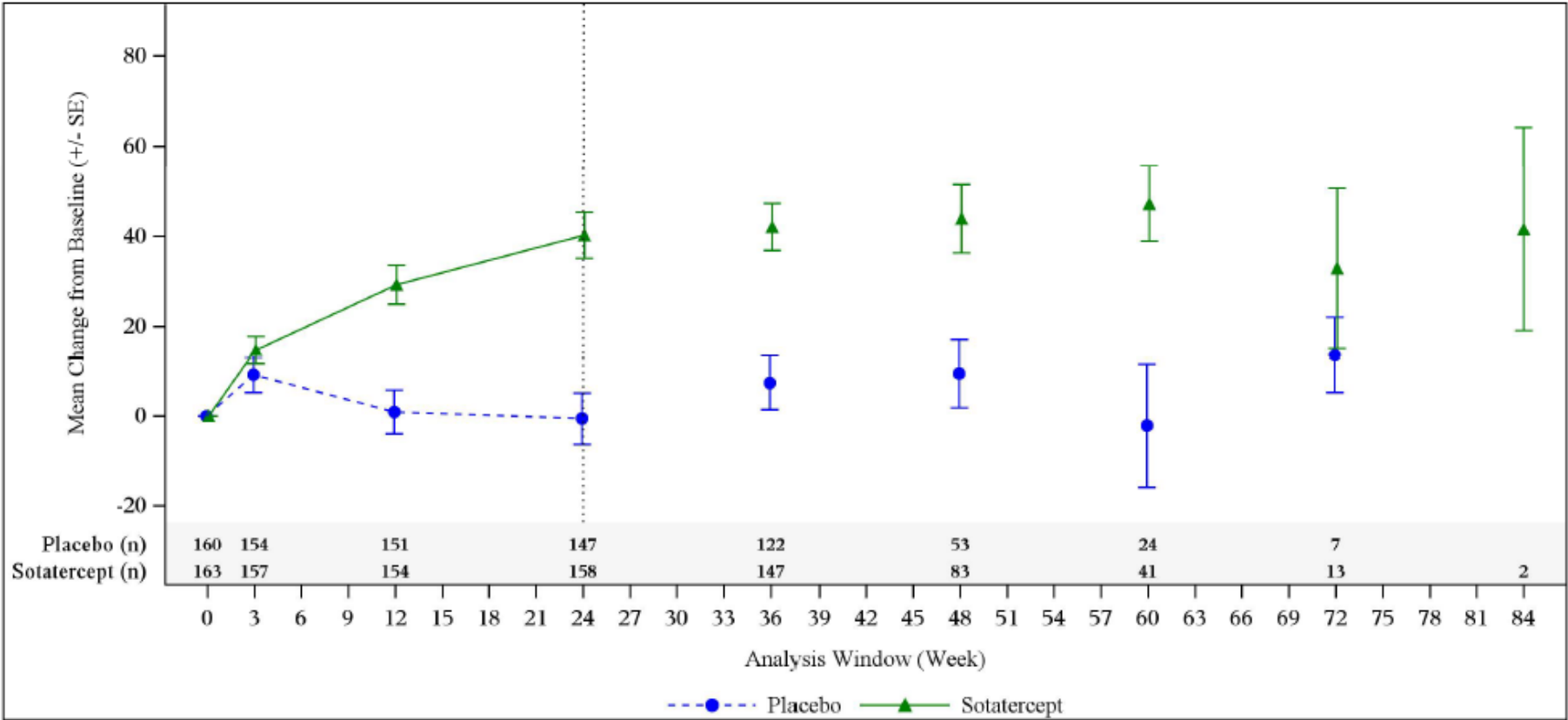
Number of participants: full analysis set

Baseline is the average of the two screening visits.

^a Latest visit for which results for this outcome are available.

Source: MSD 2024 STELLAR CSR (29)

Figure 5 Change from Baseline in 6-Minute Walk Distance (metres) by visit - Full Analysis Set



N: number of subjects in the analysis window; SE: standard error
Only visits with at least 2 subjects are shown.
Source: MSD 2024 STELLAR CSR (29)

Multicomponent improvement

The proportion of participants who met all 3 criteria of the MCI endpoint was significantly greater in the sotatercept group (39.5%) than in the placebo group (10.7%) ($p < 0.001$) at Week 24 (Table 22) (29). The proportion of participants who met each criterion were also significantly greater in the sotatercept group than in the placebo group (29).

Table 22 Analysis of Multicomponent Improvement endpoint at Week 24 - Full Analysis Set

	Sotatercept (N=163)	Placebo (N=160)
Individual component:		
Improvement of WHO FC or maintenance of WHO FC II, N	163	159
Yes, n (%)	115 (70.6)	82 (51.6)
No, n (%)	48 (29.4)	77 (48.4)
p-value ^a	<.001	N/A
Missing due to COVID-19	0	1
Improvement in NT-proBNP $\geq 30\%$ or maintenance/achievement of NT-proBNP level < 300 ng/L, N	162	159
Yes, n (%)	138 (85.2)	66 (41.5)
No, n (%)	24 (14.8)	93 (58.5)
p-value ^a	<.001	N/A
Missing due to COVID-19	1	1
Improvement in 6MWD by at least 30 metres, N	163	159
Yes, n (%)	88 (54.0)	35 (22.0)
No, n (%)	75 (46.0)	124 (78.0)
p-value ^a	<.001	N/A
Missing due to COVID-19	0	1
Multicomponent improvement of all three criteria above, N	162	159
Yes, n (%)	64 (39.5)	17 (10.7)
No, n (%)	98 (60.5)	142 (89.3)
p-value ^a	<.001	N/A
Missing due to COVID-19	1	1

6MWD: 6-minute walking distance; CMH: Cochran-Mantel-Haenszel; COVID-19: coronavirus disease 2019; FC: functional class; N/A: not applicable; NT-proBNP: N-terminal pro-brain natriuretic peptide; PAH: pulmonary arterial hypertension; WHO: World Health Organization

Study: STELLAR (MK-7962-003)

Number of participants: full analysis set

^a Comparison with placebo uses CMH method stratified by WHO functional class (class II vs. III) and background PAH therapy (mono/double vs. triple therapy).

Percentages are calculated as $(n/N) \times 100$. A missing result at week 24 not due to COVID-19 was considered a non-responder. Participants who missed assessments due to COVID-19 were taken out of the denominator.

Database Cutoff Date: 06DEC2022

Source: MSD 2024 STELLAR CSR (29)

Pulmonary vascular resistance

A decrease (i.e., improvement) from baseline in PVR was observed in the sotatercept group whereas an increase (i.e., worsening) from baseline in PVR was observed in the placebo

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

group at Week 24 (Visit 9) (29). The mean PVR values at baseline were 781.3 dynes*sec/cm⁵ in the sotatercept group and 745.8 dynes*sec/cm⁵ in the placebo group (29). Descriptive statistics of the observed mean change from baseline in PVR at Week 24 showed a decrease in the sotatercept group (-234.5 dynes*sec/cm⁵) compared with an increase in the placebo group (41.9 dynes*sec/cm⁵) (Table 23) (29). The median treatment difference between the sotatercept and placebo groups was -231.8 dynes*sec/cm⁵ (p<0.001) (Table 24) (29). Results from the supportive analysis using the ANCOVA model were consistent with those from the primary analysis (Table 25) (29).

Table 23 Change from Baseline in PVR (dynes*sec/cm⁵) by Visit 24 - Full Analysis Set

	Sotatercept (N=163)	Placebo (N=160)
Baseline		
n	163	160
Mean (SD)	781.3 (398.47)	745.8 (313.53)
Median	624.0	684.0
Min, Max	400.0, 2688.0	400.0, 2624.0
Missing	0	0
Week 24		
n	154	141
Mean (SD)	545.8 (265.15)	779.3 (335.13)
Median	488.0	744.0
Min, Max	144.0, 1864.0	272.0, 2448.0
Missing	9	19
Change from baseline at Week 24		
n	154	141
Mean (SD)	-234.5 (335.98)	41.9 (220.62)
Median	-172.0	24.0
Min, Max	-1792.0, 656.0	-608.0, 848.0
Missing	9	19
Percentage change from baseline at Week 24		
n	154	141
Mean (SD)	-25.0 (27.17)	7.9 (34.32)
Median	-26.7	2.8
Min, Max	-81.5, 92.1	-60.3, 165.6
Missing	9	19

PVR: pulmonary vascular resistance; SD: standard deviation

Study: STELLAR (MK-7962-003)

Database Cutoff Date: 06DEC2022

Number of participants: full analysis set

Source: MSD 2024 STELLAR CSR (29)

Table 24 Analysis of change from baseline in PVR (dynes*sec/cm⁵) at Week 24 - Full Analysis Set (Standard multiple imputation)

	Sotatercept (N=163)	Placebo (N=160)
Median estimate (Min, Max) ^a	-160.6 (-176.0, -152.0)	32.8 (24.0, 40.0)
Hodges-Lehmann location shift (ASE) ^b	-231.8 (27.61)	N/A
95% CI of Hodges-Lehmann location shift	(-285.88, -177.63)	N/A
Wilcoxon p-value ^c	<.001	N/A

ASE; asymptomatic standard error; CI: confidence interval; N/A: not applicable; PVR: pulmonary vascular resistance

Database Cutoff Date: 06DEC2022

Number of participants: full analysis set

^a Average, minimum and maximum of the medians across the imputed datasets if missing data imputed.

^b Hodges-Lehmann location shift from placebo estimate (median of all paired differences).

^c Wilcoxon p-value refers to p-value from the aligned rank stratified Wilcoxon test with randomisation factors as strata.

Change from baseline in PVR at Week 24 for participants who died was assigned as 20000 to receive the worst rank. Change from baseline in PVR at Week 24 for participants who had missing data due to a non-fatal clinical worsening event was imputed as 15000 to receive the next worst rank.

Source: MSD 2024 STELLAR CSR (29)

Table 25 Analysis of change from baseline in PVR (dynes*sec/cm⁵) at Week 24 - Full Analysis Set (Standard multiple imputation with ANCOVA model)

	Sotatercept (N=163)	Placebo (N=160)
Change from Baseline at Week 24^a		
LS Mean (SE)	-214.540 (24.168)	81.356 (24.753)
95% CI of LS Mean	(-261.909, -167.171)	(32.840, 129.872)
LS Mean Diff (SE)	-295.896 (34.120)	N/A
95% CI of LS Mean Diff	(-362.771, -229.020)	N/A
Shapiro-Wilk p-value ^b	<.001	N/A

ANCOVA: analysis of covariance; CI: confidence interval; FC: functional class; LS: least squares; N/A: not applicable; PVR: pulmonary vascular resistance; SE: standard error; WHO: World Health Organization

Study: STELLAR (MK-7962-003)

Database Cutoff Date: 06DEC2022

Number of participants: full analysis set

^a ANCOVA with treatment, randomisation stratification factors of baseline (Mono/Double, Triple) and baseline WHO FC (II, III) as fixed factors, baseline PVR as covariate. Non-existent data due to death and missing data due to non-fatal clinical worsening was assigned/imputed to 1 unit worse than the worst observed change from baseline from both treatment groups.

^b P-value is from Shapiro-Wilk test on the residuals from the ANCOVA model. If the normality assumption is violated (p-value <0.001), p-value results from ANCOVA are not presented.

Source: MSD 2024 STELLAR CSR (29)

Table 26 PVR responder analysis - Full Analysis Set

	Sotatercept (N=163)	Placebo (N=160)
Achieve PVR <3 WU at Week 24, N	154	141
Yes, n (%)	8 (5.2)	0
No, n (%)	146 (94.8)	141 (100.0)
p-value ^a	0.005	N/A

CMH: Cochran-Mantel-Haenszel; N/A: not applicable; PVR: pulmonary vascular resistance; WU: Wood Unit

Study: STELLAR (MK-7962-003)

Database Cutoff Date: 06DEC2022

Number of participants: full analysis set

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

^a Comparison with placebo uses CMH method stratified by WHO functional class (class II vs. III) and background PAH therapy (mono/double vs. triple therapy).
Percentages are calculated as (n/N)*100.
Source: MSD 2024 STELLAR CSR (29)

NT-proBNP

A decrease (i.e., improvement) from baseline in NT-proBNP was observed in the sotatercept group whereas an increase (i.e., worsening) from baseline in NT-proBNP was observed in the placebo group at Week 24 (Visit 9) (29). The mean NT-proBNP values at baseline were 1037.5 pg/mL in the sotatercept group and 1194.3 pg/mL in the placebo group (29). Descriptive statistics of the observed mean change from baseline in NT-proBNP at Week 24 showed a decrease in the sotatercept group (-722.3 pg/mL) compared with an increase in the placebo group (346.6 pg/mL; Table 27) (29).

The median treatment difference between the sotatercept and placebo groups was -438.9 pg/mL ($p < 0.001$; Table 28) (29). Results from the supportive analysis using the ANCOVA model were consistent with those from the primary analysis (Table 29) (29).

Table 27 Change from baseline in NT-proBNP (pg/mL) by visit - Full Analysis Set

	Sotatercept (N=163)	Placebo (N=160)
Baseline		
n	162	159
Mean (SD)	1037.5 (2498.62)	1194.3 (2671.39)
Median	422.0	440.0
Min, Max	6.5, 28354.0	20.0, 29037.0
Missing	1	1
Week 24		
n	149	140
Mean (SD)	303.4 (507.77)	1237.4 (2382.35)
Median	118.0	486.5
Min, Max	6.0, 3261.0	26.0, 15179.0
Missing	14	20
Change from baseline at Week 24		
n	148	139
Mean (SD)	-722.3 (2479.36)	346.6 (1276.45)
Median	-224.0	44.0
Min, Max	-28290.0, 660.0	-1633.0, 9055.0
Missing	15	21
Percentage change from baseline at Week 24		
n	148	139
Mean (SD)	-46.0 (58.74)	51.7 (180.06)
Median	-62.6	17.8
Min, Max	-99.8, 392.3	-78.8, 1929.4
Missing	15	21

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

	Sotatercept (N=163)	Placebo (N=160)
Week 84^a		
n	2	1
Mean (SD)	187.0 (128.69)	316.0 (NE)
Median	187.0	316.0
Min, Max	96.0, 278.0	316.0, 316.0
Missing	161	159
Change from baseline at Week 84		
n	2	1
Mean (SD)	-44.0 (131.52)	96.0 (NE)
Median	-44.0	96.0
Min, Max	-137.0, 49.0	96.0, 96.0
Missing	161	159
Percentage change from baseline at Week 84		
n	2	1
Mean (SD)	35.6 (97.06)	43.6 (NE)
Median	35.6	43.6
Min, Max	-33.0, 104.3	43.6, 43.6
Missing	161	159

SD: standard error; NE: Not Evaluable; NT-proBNP: N-terminal pro-brain natriuretic peptide

Study: STELLAR (MK-7962-003)

Database Cutoff Date: 06DEC2022

Number of participants: full analysis set

Baseline is defined as the last measurement prior to first dose.

^a Latest visit for which results for this outcome are available.

Source: MSD 2024 STELLAR CSR (29)

Table 28 Analysis of change from baseline in NT-proBNP (pg/mL) at Week 24 - Full Analysis Set (Standard multiple imputation)

	Sotatercept (N=163)	Placebo (N=160)
Median estimate (Min, Max) ^a	-228.4 (-236.0, -223.0)	64.8 (49.0, 85.0)
Hodges-Lehmann location shift (ASE) ^b	-438.9 (66.70)	N/A
95% CI of Hodges-Lehmann location shift	(-569.64, -308.19)	N/A
Wilcoxon p-value ^c	<.001	N/A

ASE: asymptomatic standard error; CI: confidence interval; N/A: not applicable; NT-proBNP: N-terminal pro-brain natriuretic peptide

Study: STELLAR (MK-7962-003)

Database Cutoff Date: 06DEC2022

Number of participants: full analysis set

^a Average, minimum and maximum of the medians across the imputed datasets if missing data imputed.

^b Hodges-Lehmann location shift from placebo estimate (median of all paired differences).

^c Wilcoxon p-value refers to p-value from the aligned rank stratified Wilcoxon test with randomisation factors as strata.

Change from baseline in NT-proBNP at Week 24 for participants who died was assigned as 200000 to receive the worst rank.

Change from baseline in NT-proBNP at Week 24 for participants who had missing data due to a non-fatal clinical worsening event was imputed as 150000 to receive the next worst rank.

Source: MSD 2024 STELLAR CSR (29)

Table 29 Analysis of change from baseline in NT-proBNP (pg/mL) at Week 24 - Full Analysis Set (Standard multiple imputation with ANCOVA model)

	Sotatercept (N=163)	Placebo (N=160)
Change from Baseline at Week 24^a		
LS Mean (SE)	-494.309 (425.093)	1818.058 (436.003)
95% CI of LS Mean	(-1327.477, 338.859)	(963.503, 2672.613)
LS Mean Diff (SE)	-2312.367 (600.454)	N/A
95% CI of LS Mean Diff	(-3489.238, -1135.497)	N/A
Shapiro-Wilk p-value ^b	<.001	N/A

ANCOVA: analysis of covariance; CI: confidence interval; FC: functional class; LS: least squares; N/A: not applicable; NT-proBNP: N-terminal pro-brain natriuretic peptide; SE: standard error; WHO: World Health Organization

Study: STELLAR (MK-7962-003)

Database Cutoff Date: 06DEC2022

Number of participants: full analysis set

^a ANCOVA with treatment, randomisation stratification factors of baseline (Mono/Double, Triple) and baseline WHO FC (II, III) as fixed factors, baseline NT-proBNP as covariate.

Non-existent data due to death and missing data due to non-fatal clinical worsening was assigned/imputed to 1 unit worse than the worst observed change from baseline from both treatment groups.

^b P-value is from Shapiro-Wilk test on the residuals from the ANCOVA model. If the normality assumption is violated (p-value <0.001), p-value results from ANCOVA are not presented.

Source: MSD 2024 STELLAR CSR (29)

WHO FC improvement

The proportion of participants with improvement from baseline in WHO FC at Week 24 was significantly greater in the sotatercept group (29.4%) than in the placebo group (13.8%) (p<0.001; Table 30), regardless of the baseline WHO FC (Table 31, Figure 6) (29). The time to first WHO FC improvement was shorter in the sotatercept group than in the placebo group (hazard ratio [HR]: 2.42, 95% CI: 1.58, 3.72, p<0.001) (29).

Table 30 Analysis of change from baseline in WHO FC at Week 24 - Full Analysis Set

	All participants		WHO FC II at baseline		WHO FC III at baseline	
	Sotatercept (N=163)	Placebo (N=160)	Sotatercept (N=80)	Placebo (N=78)	Sotatercept (N=83)	Placebo (N=82)
Improvement of WHO FC, N	163	159	80	78	83	81
Yes, n (%)	48 (29.4)	22 (13.8)	8 (10.0)	3 (3.8)	40 (48.2)	19 (23.5)
No, n (%)	115 (70.6)	137 (86.2)	72 (90.0)	75 (96.2)	43 (51.8)	62 (76.5)
p-value ^a	<.001	N/A	0.125	N/A	0.001	N/A
Missing due to COVID-19	0	1	0	0	0	1
Maintenance of WHO FC, N	163	159	80	78	83	81
Yes, n (%)	104 (63.8)	106 (66.7)	67 (83.8)	60 (76.9)	37 (44.6)	46 (56.8)
No, n (%)	59 (36.2)	53 (33.3)	13 (16.3)	18 (23.1)	46 (55.4)	35 (43.2)
p-value ^a	0.605	N/A	0.301	N/A	0.123	N/A
Missing due to COVID-19	0	1	0	0	0	1
Worsening of WHO FC, N	163	159	80	78	83	81
Yes, n (%)	7 (4.3)	19 (11.9)	5 (6.3)	12 (15.4)	2 (2.4)	7 (8.6)
No, n (%)	156 (95.7)	140 (88.1)	75 (93.8)	66 (84.6)	81 (97.6)	74 (91.4)
p-value ^a	0.013	N/A	0.069	N/A	0.090	N/A
Missing due to COVID-19	0	1	0	0	0	1

CMH: Cochran-Mantel-Haenszel; COVID-19: coronavirus disease 19; FC: functional class; N/A: not applicable; PAH: pulmonary arterial hypertension; WHO: World Health Organization

Database Cutoff Date: 06DEC2022

Number of participants: full analysis set

^a Comparison with placebo uses CMH method stratified by WHO functional class (class II vs. III) and background PAH therapy (mono/double vs. triple therapy).

Percentages are calculated as (n/N)*100. A missing result at week 24 not due to COVID-19 was considered a non-responder. Participants who missed assessments due to COVID-19 were taken out of the denominator.

Source: MSD 2024 STELLAR CSR (29)

Table 31 Shift from baseline in WHO functional class by Visit 24 – Full Analysis Set

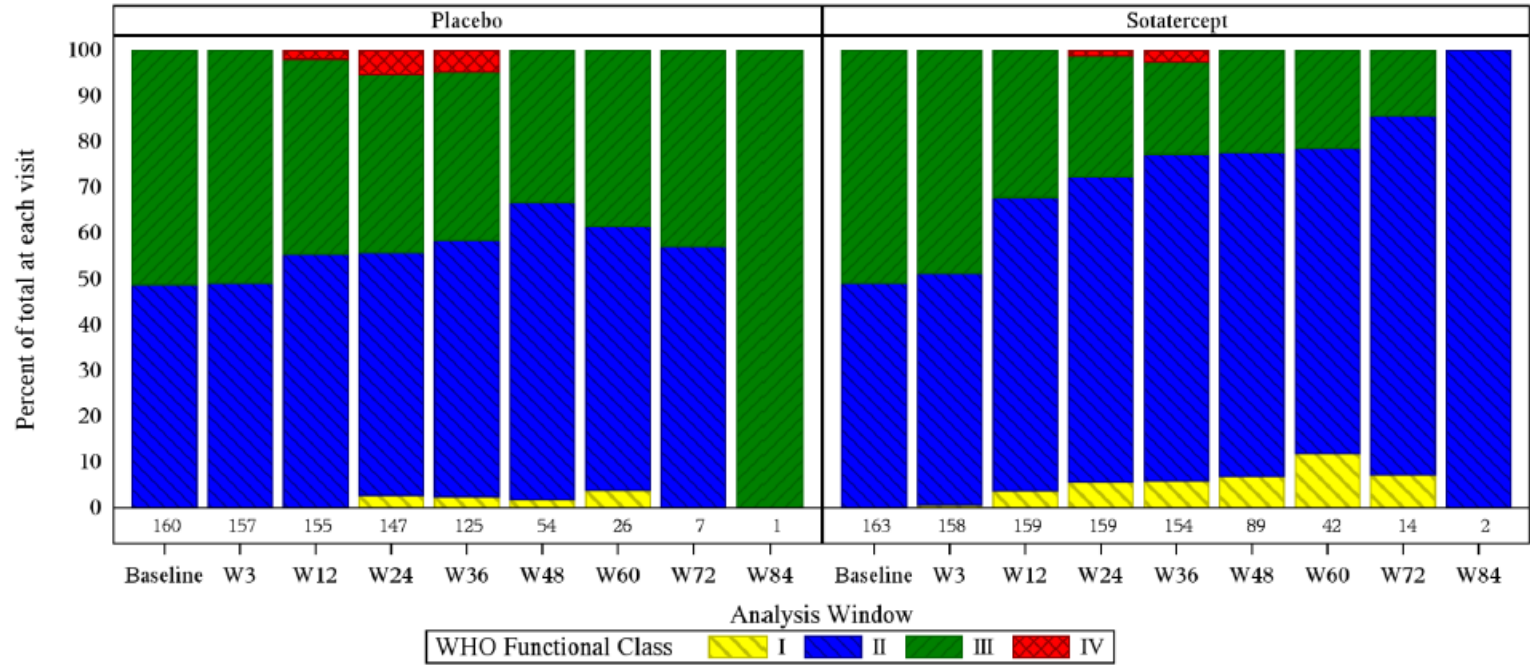
WHO Functional Class	Sotatercept				Placebo			
	N	Class II	Class III	Total	N	Class II	Class III	Total
Baseline	163				160			
Total		80 (49.1)	83 (50.9)	163 (100.0)		78 (48.8)	82 (51.3)	160 (100.0)

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

	Sotatercept				Placebo			
WHO Functional Class	N	Class II	Class III	Total	N	Class II	Class III	Total
Week 24	159				147			
Class I		8 (5.0)	1 (0.6)	9 (5.7)		3 (2.0)	1 (0.7)	4 (2.7)
Class II		67 (42.1)	39 (24.5)	106 (66.7)		60 (40.8)	18 (12.2)	78 (53.1)
Class III		5 (3.1)	37 (23.3)	42 (26.4)		11 (7.5)	46 (31.3)	57 (38.8)
Class IV		0 (0.0)	2 (1.3)	2 (1.3)		1 (0.7)	7 (4.8)	8 (5.4)
Total		80 (50.3)	79 (49.7)	159 (100.0)		75 (51.0)	72 (49.0)	147 (100.0)

WHO: World Health Organization
Database Cutoff Date: 06DEC2022
Source: MSD 2024 STELLAR CSR (29)

Figure 6 WHO functional class by visit – Full Analysis Set – Cumulative results



W: week; WHO: World Health Organization

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

© Merck Sharp & Dohme (UK) Limited (2025). All rights reserved

Database Cutoff Date: 06DEC2022
Source: MSD 2024 STELLAR CSR (29)

Table 32 Time to first WHO FC improvement

Treatment	N	Number of events (%)	Person-weeks	Event rate/ 100 person-weeks (%)	Median WHO FC improvement ^a (weeks) (95% CI)	Patients without WHO FC improvement at Week 24 in % ^a (95% CI)
Sotatercept	163	72 (44.2)	5676.57	1.27	59.14 (44.00, NR)	74.20 (66.75, 80.23)
Placebo	160	30 (18.8)	5570.71	0.54	NR (NR, NR)	85.15 (78.50, 89.88)
Pairwise comparisons					Hazard ratio^b (95% CI)^b	p-value^c
Sotatercept vs. placebo					2.42 (1.58, 3.72)	<0.001

CI: confidence interval; FC: functional class; NR: not reached; PAH: pulmonary arterial hypertension; WHO: World Health Organization

Database Cutoff Date: 06DEC2022

^a From product-limit (Kaplan-Meier) method for censored data.

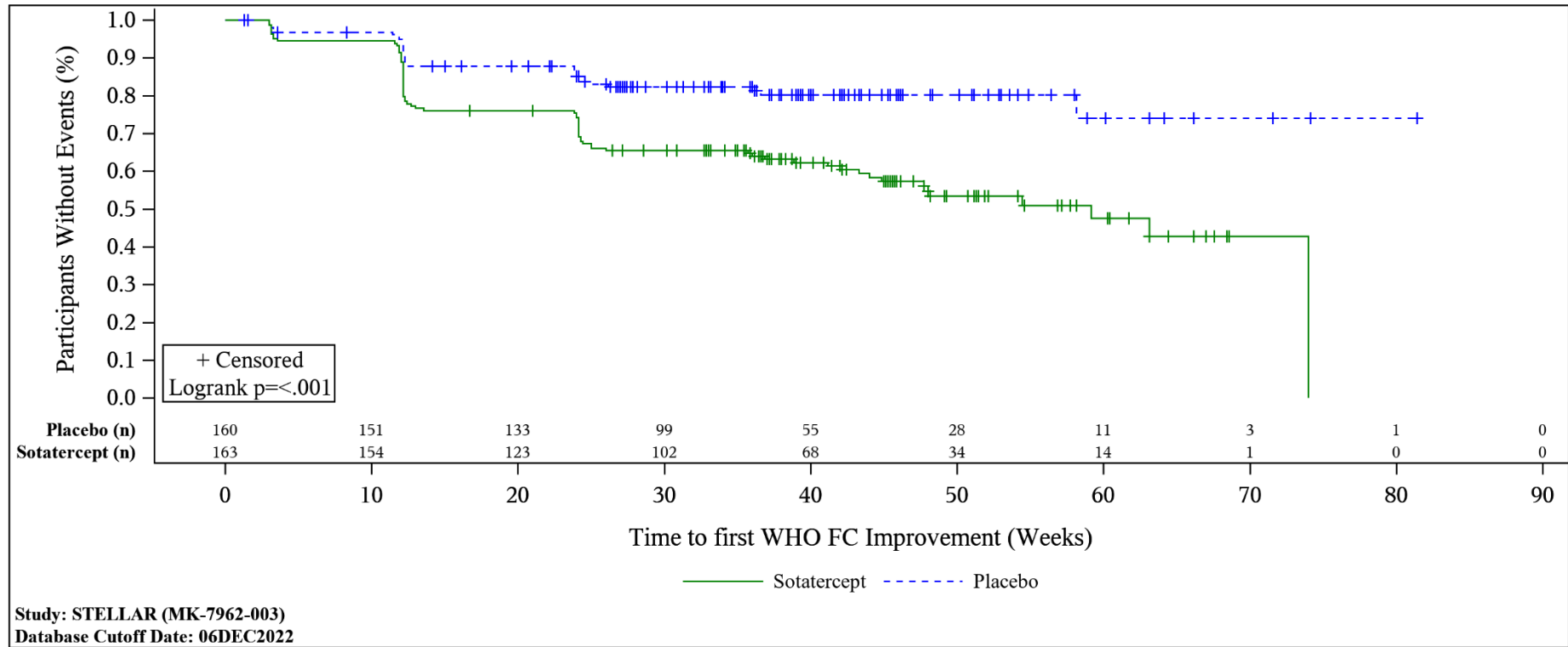
^b Based on Cox regression model with Efron's method of tie handling with treatment as a covariate stratified by the following covariates: WHO functional class (class II vs. III) and background PAH therapy (mono/double vs. triple therapy).

^c Two-sided p-value based on logrank test stratified by the following covariates: WHO functional class (class II vs. III) and background PAH therapy (mono/double vs. triple therapy).

Person-Weeks refers to an estimate of the actual time-at-risk in weeks that all participants in the given treatment arm contributed to the study.

Source: MSD 2024 STELLAR CSR (29)

Figure 7 WHO FC improvement Kaplan-Meier plot – Full Analysis Set



FC: functional class; WHO: World Health Organization
Source: MSD 2024 STELLAR CSR (29)

Time to death or the first occurrence of a worsening event

Analyses of time to death or first occurrence of a clinical worsening event were performed using data up to the data cutoff (06-DEC-2022; Table 33) (29). Fewer participants in the sotatercept group (11 [6.7%]) than in the placebo group (42 [26.3%]) died or had at least 1 clinical worsening event as of the data cutoff (29). The risk of death or a first clinical worsening event was 77% lower in the sotatercept group compared with the placebo group (HR: 0.23; 95% CI, 0.12, 0.44) (29). The proportion of participants with a clinical worsening event was lower in the sotatercept group than in the placebo group in each clinical worsening category, except for worsening-related listing for lung and/or heart transplant (1 participant in both groups) and need for atrial septostomy (0 participants in each group) (29). Kaplan-Meier curves for time to death or first occurrence of a clinical worsening showed an early separation (approximately Week 10) of the sotatercept and placebo curves that continued for the remainder of the study (Figure 8) (29). This was confirmed by the stratified logrank test ($p < 0.001$) (29).

Table 33 Time to first clinical worsening or death - Full Analysis Set

	Sotatercept (N=163)	Placebo (N=160)
Total number of participants who experienced at least one clinical worsening event or death, n (%)	11 (6.7)	42 (26.3)
Total number of clinical worsening events or death	11	45
Time to first occurrence of clinical worsening events or death (weeks)		
• Logrank test p-value ^a	<.001	N/A
• Hazard Ratio (95% CI) ^b	0.23 (0.12, 0.44)	N/A
Assessment of first occurrence of clinical worsening events or death, n (%)^c		
• Death	2 (1.2)	6 (3.8)
• Worsening-related listing for lung and/or heart transplant	1 (0.6)	1 (0.6)
• Need to initiate rescue therapy with an approved PAH therapy or the need to increase the dose of infusion prostacyclin by 10% or more	2 (1.2)	17 (10.6)
• Need for atrial septostomy	0 (0.0)	0 (0.0)
• PAH-specific hospitalisation (≥ 24 hours)	1 (0.6)	7 (4.4)
• Deterioration of PAH ^d	5 (3.1)	15 (9.4)

6MWD: 6-minute walking distance; 6MWT: 6-minute walking test; CI: confidence interval; N/A: not applicable;

PAH: pulmonary arterial hypertension; WHO: World Health Organization

Database Cutoff Date: 06DEC2022

Number of participants: full analysis set

^a Logrank test comparison with placebo stratified by the randomisation factors.

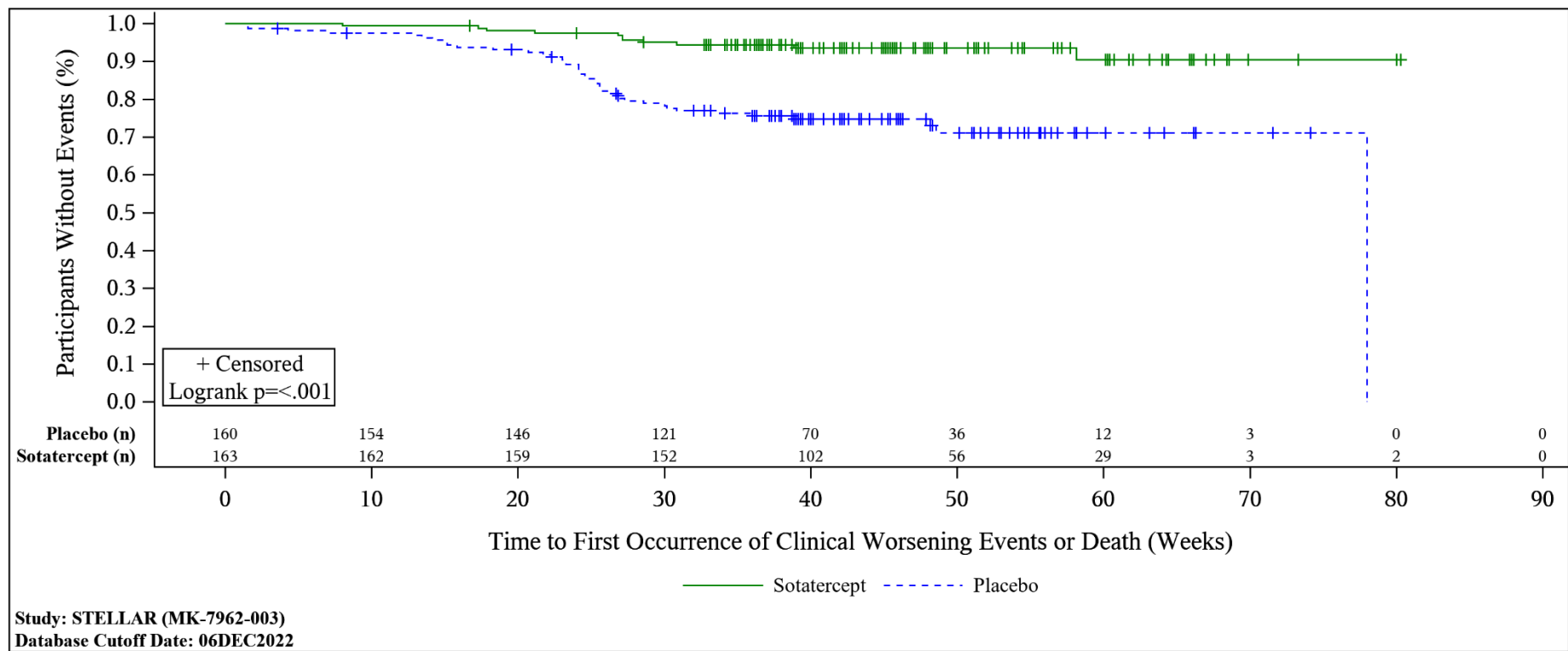
^b The hazard ratio (sotatercept / placebo) is derived from a Cox proportional hazard model with treatment group as the covariate stratified by the randomisation factors.

^c A participant can have more than one assessment recorded for their first event of clinical worsening or death.

^d Deterioration of PAH therapy is defined by both of the following events occurring at any time, even if they began at different times, as compared to their baseline values: (a) Worsened WHO functional class (II to III, III to IV, II to IV, etc.), (b) Decrease in 6MWD by $\geq 15\%$ (confirmed by two 6MWTs at least 4 hours apart but no more than one week).

Source: MSD 2024 STELLAR CSR (29)

Figure 8 Clinical worsening or death Kaplan-Meier plot



Source: MSD 2024 STELLAR CSR (29)

French Risk score

The proportion of participants who maintained or achieved a low risk score (French Risk score calculator) relative to baseline at Week 24 was significantly greater in the sotatercept group (40.1%) than in the placebo group (19.5%) ($p < 0.001$) (Table 35) (29).

Table 34 Low risk score (French Risk Score calculator) by visit - Full Analysis Set

Low Risk Score	Sotatercept (N=163)	Placebo (N=160)
Baseline		
m	163	159
Yes, n (%)	28 (17.2)	27 (17.0)
No, n (%)	135 (82.8)	132 (83.0)
Missing due to public health concerns (e.g., COVID-19)	0	0
Week 24		
m	154	145
Yes, n (%)	65 (42.2)	31 (21.4)
No, n (%)	89 (57.8)	114 (78.6)
Missing due to public health concerns (e.g., COVID-19)	1	1

6MWD: 6-minute walking distance; COVID-19: coronavirus disease 2019; FC: functional class; m: number of subjects with available values at a visit; NT-proBNP: N-terminal pro-brain natriuretic peptide; WHO: World Health Organization

A subject is considered to achieve a low risk score using the simplified French Risk score calculator if the following criteria are satisfied: WHO FC I or II, 6MWD > 440 m and NT-proBNP < 300 ng/L.

Percentages are calculated as $(n/m) \times 100$.

Source: MSD 2024 STELLAR CSR (29)

Table 35 Analysis of change from baseline in low risk score (French Risk Score calculator) at Week 24 - Full Analysis Set

	Sotatercept (N=163) 162	Placebo (N=160) 159
Any maintain or achieve a low risk score using the Simplified French Risk Score calculator at Week 24 vs. baseline, m		
Yes	65 (40.1)	31 (19.5)
No	97 (59.9)	128 (80.5)
p-value ^a	$< .001$	N/A
Missing due to COVID-19	1	1

6MWD: 6-minute walking distance; CMH: Cochran-Mantel-Haenszel; COVID-19: coronavirus disease 2019; PAH: pulmonary arterial hypertension; NT-proBNP: N-terminal pro-brain natriuretic peptide; WHO: World Health Organization

Database Cutoff Date: 06DEC2022

Number of participants: full analysis set

^a Comparison with placebo uses CMH method stratified by WHO functional class (class II vs. III) and background PAH therapy (mono/double vs. triple therapy).

A participant is considered to achieve a low risk score using the simplified French Risk score calculator if the following criteria are satisfied: WHO FC I or II, 6MWD > 440 m and NT-proBNP < 300 ng/L.

Percentages are calculated as $(n/m) \times 100$. A missing result at week 24 not due to COVID-19 was considered a non-responder. Participants who missed assessments due to COVID-19 were taken out of the denominator.

Source: MSD 2024 STELLAR CSR (29)

ESC/ERS (COMPERA 2.0) four-strata risk status

Change from baseline to Week 24 was also assessed in detail according to the ESC/ERS (COMPERA 2.0) four-strata risk scoring system (4, 32). Patients in the sotatercept treatment group were significantly more likely to experience a risk status improvement at Week 24 than patients in the placebo treatment group (Table 36 to Table 39) (33).

Table 36 Four-strata risk status assessment by treatment at baseline and Week 24 – Full Analysis Set

	Sotatercept (N^a=163)	Placebo (N^a=160)	Total (N^a=323)
Baseline^b	162 (100.0%)	159 (100.0%)	321 (100.0%)
Low Risk	██████	██████	██████
Intermediate-Low Risk	██████	██████	██████
Intermediate-High Risk	██████	██████	██████
High Risk	██████	██████	██████
Week 24^b	148 (100.0%)	140 (100.0%)	288 (100.0%)
Low Risk	██████	██████	██████
Intermediate-Low Risk	██████	██████	██████
Intermediate-High Risk	██████	██████	██████
High Risk	██████	██████	██████

Study: STELLAR (MK-7962-003)

Database Cutoff Date: 06DEC2022

^a Number of participants: full analysis set population

^b Number of participants: full analysis set population, participants with non-missing risk assessment stratum at each visit. Column percentages are based on this number

Four-stratum risk assessment is based on COMPERA version 2.0

Source: MSD 2024 Confidential STELLAR post-hoc analyses results (33)

Table 37 Four-strata risk status assessment transitions from baseline to Week 24 – Full Analysis Set

Treatment Group	From Baseline	n ^b	To Week 24			
			Low Risk	Intermediate-Low Risk	Intermediate-High Risk	High Risk
Placebo (N ^a =160)	Low Risk	■	■	■	■	■
	Intermediate-Low Risk	■	■	■	■	■
	Intermediate-High Risk	■	■	■	■	■
	High Risk	■	■	■	■	■
Sotatercept (N ^a =163)	Low Risk	■	■	■	■	■
	Intermediate-Low Risk	■	■	■	■	■
	Intermediate-High Risk	■	■	■	■	■
	High Risk	■	■	■	■	■

Study: STELLAR (MK-7962-003)

Database Cutoff Date: 06DEC2022

^a Number of participants: full analysis set population^b Number of participants: full analysis set population, participants with non-missing risk assessment stratum at both analysis visits. Row percentages are based on this number
Four-stratum risk assessment is based on COMPERA version 2.0

Source: MSD 2024 Confidential STELLAR post-hoc analyses results (33)

Table 38 Analysis of change in four-strata risk status from baseline at Week 24 based on complete cases - Full Analysis Set

Study: STELLAR (MK-7962-003) ^a		Sotatercept		Placebo		Sotatercept vs. placebo		
	N ^b	Participants with event n (%)	N ^b	Participants with event n (%)	Odds ratio [95% CI] ^c	Relative risk [95% CI] ^c	p-value for interaction test ^d	
Risk status improvement at Week 24								
Risk status at baseline								
Low risk								
Intermediate-low risk								
Intermediate-high risk								
Risk status maintenance at Week 24								
Risk status at baseline								
Low risk								

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Study: STELLAR (MK-7962-003) ^a	Sotatercept		Placebo		Sotatercept vs. placebo		p-value for interaction test ^d
	N ^b	Participants with event n (%)	N ^b	Participants with event n (%)	Odds ratio [95% CI] ^c	Relative risk [95% CI] ^c	
Intermediate-low risk	■	■■■■■	■	■■■■■	■■■■■	■■■■■	
Intermediate-high risk	■	■■■■■	■	■■■■■	■■■■■	■■■■■	
Risk status worsening at Week 24							
Risk status at baseline							
Low risk	■	■■■■■	■	■■■■■	■■■■■	■■■■■	■
Intermediate-low risk	■	■■■■■	■	■■■■■	■■■■■	■■■■■	
Intermediate-high risk	■	■	■	■■■■■	■■■■■	■■■■■	

CI: confidence interval
^a Database Cutoff Date: 06DEC2022
^b Number of participants: full analysis set. Participants with missing assessment at either baseline or Week 24 are taken out of the population
^c Based on 2x2 contingency table, using Wald-type confidence interval.
For Odds ratio, in case of no participants in any contingency table cell, a continuity correction factor (+0.5) is applied and Wald modified confidence interval is applied
For Relative risk, in case of no participants with event in at least one treatment arm, a continuity correction factor (+0.5) is applied and Wald modified confidence interval is applied
For both odds ratio and relative risk, in case of no participants or all participants with event in both treatment arms, report 'n.a.'
^d Based on Breslow-Day test for homogeneity of odds ratios stratified by subgroup
Risk status assessment is based on COMPERA version 2.0
Source: MSD 2024 Confidential STELLAR post-hoc analyses results (33)

Table 39 Analysis of change in four-strata risk status from baseline at Week 24 based on imputed cases - Full Analysis Set

Study: STELLAR (MK-7962-003) ^a	Sotatercept		Placebo		Sotatercept vs. placebo		p-value for interaction test ^d
	N ^b	Participants with event n (%)	N ^b	Participants with event n (%)	Odds ratio [95% CI] ^c	Relative risk [95% CI] ^c	
Risk status improvement at Week 24							
Risk status at baseline:							
Low risk	■	■	■	■	■	■	■
Intermediate-low risk	■	■	■	■	■	■	
Intermediate-high risk	■	■	■	■	■	■	

Study: STELLAR (MK-7962-003) ^a	Sotatercept		Placebo		Sotatercept vs. placebo		p-value for interaction test ^d
	N ^b	Participants with event n (%)	N ^b	Participants with event n (%)	Odds ratio [95% CI] ^c	Relative risk [95% CI] ^c	
Risk status maintenance at Week 24							
Risk status at baseline:							
Low Risk	■	■	■	■	■	■	■
Intermediate-low risk	■	■	■	■	■	■	
Intermediate-high risk	■	■	■	■	■	■	
Risk status worsening at Week 24							
Risk status at baseline:							
Low risk	■	■	■	■	■	■	■
Intermediate-low risk	■	■	■	■	■	■	
Intermediate-high risk	■	■	■	■	■	■	

CI: confidence interval; COVID-29: coronavirus disease 2019

^a Database Cutoff Date: 06DEC2022

^b Number of participants: full analysis set. Participants with missing assessment at week 24 not due to COVID-19 are considered a non-responder. Participants who missed assessment at Week 24 due to COVID-19 are taken out of the population

^c Based on 2x2 contingency table, using Wald-type confidence interval.

For Odds ratio, in case of no participants in any contingency table cell, a continuity correction factor (+0.5) is applied and Wald modified confidence interval is applied

For Relative risk, in case of no participants with event in at least one treatment arm, a continuity correction factor (+0.5) is applied and Wald modified confidence interval is applied

For both odds ratio and relative risk, in case of no participants or all participants with event in both treatment arms, report 'n.a.'

Confidence interval entries with value 0.00 have unrounded entries > 0.

^d Based on Breslow-Day test for homogeneity of odds ratios stratified by subgroup

Risk status assessment is based on COMPERA version 2.0

Source: MSD 2024 Confidential STELLAR post-hoc analyses results (33)

PAH-SYMPACT®**PAH-SYMPACT® Physical Impacts**

The change from baseline in Physical Impacts domain score of PAH-SYMPACT® showed improvement that was significantly greater in the sotatercept group than in the placebo group at Week 24 (Visit 9) (Table 40) (29). The median of the treatment difference between the sotatercept and placebo groups was -0.26. (p=0.012; Table 41) (29). Results from the supportive analysis using the ANCOVA model were consistent with those from the primary analysis (Table 42) (29).

Table 40 Change from baseline in Physical Impacts Domain Score of PAH-SYMPACT® by visit - Full Analysis Set

	Sotatercept (N=163)	Placebo (N=160)
Baseline		
n	117	123
Mean (SD)	1.2 (0.87)	1.2 (0.79)
Median	1.1	1.1
Min, Max	0.0, 3.6	0.0, 3.6
Missing	46	37
Week 24		
n	115	110
Mean (SD)	0.9 (0.73)	1.1 (0.85)
Median	0.9	0.9
Min, Max	0.0, 3.4	0.0, 3.7
Missing	48	50
Change from baseline at Week 24		
n	86	92
Mean (SD)	-0.3 (0.67)	0.0 (0.60)
Median	-0.1	0.0
Min, Max	-2.7, 1.0	-2.4, 2.3
Missing	77	68

PAH-SYMPACT: Pulmonary Arterial Hypertension-Symptoms and Impact Questionnaire; SD: standard deviation
Database Cutoff Date: 06DEC2022

Baseline is the average of the two screening visits.

Source: MSD 2024 STELLAR CSR (29)

Table 41 Analysis of change from baseline in Physical Impacts Domain Score of PAH-SYMPACT® at Week 24 - Full Analysis Set (Standard multiple imputation)

	Sotatercept (N=163)	Placebo (N=160)
Median estimate (Min, Max) ^a	-0.12 (-0.15, 0.00)	0.01 (0.00, 0.14)
Hodges-Lehmann location shift (ASE) ^b	-0.26 (0.114)	N/A
95% CI of Hodges-Lehmann location shift	(-0.482, -0.034)	N/A
Wilcoxon p-value ^c	0.012	N/A

ASE: asymptomatic standard error; CI: confidence interval; N/A: not applicable; PAH-SYMPACT: Pulmonary Arterial Hypertension-Symptoms and Impact Questionnaire

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Database Cutoff Date: 06DEC2022

^a Average, minimum and maximum of the medians across the imputed datasets if missing data imputed.

^b Hodges-Lehmann location shift from placebo estimate (median of all paired differences).

^c Wilcoxon p-value refers to p-value from the aligned rank stratified Wilcoxon test with randomisation factors as strata.

Change from baseline in SYMPACT® scores at Week 24 for participants who died was assigned as 200 to receive the worst rank.

Change from baseline in SYMPACT® scores at Week 24 for participants who had missing data due to a non-fatal clinical worsening event was imputed as 150 to receive the next worst rank.

Source: MSD 2024 STELLAR CSR (29)

Table 42 Analysis of change from baseline in Physical Impacts Domain Score of PAH-SYMPACT® at Week 24 - Full Analysis Set (Standard multiple imputation with ANCOVA model)

	Sotatercept (N=163)	Placebo (N=160)
Change from baseline at Week 24^a		
LS mean (SE)	-0.198 (0.113)	0.192 (0.109)
95% CI of LS mean	(-0.419, 0.023)	(-0.022, 0.406)
LS mean diff (SE)	-0.390 (0.157)	N/A
95% CI of LS mean diff	(-0.697, -0.083)	N/A
p-value	0.013	N/A
Shapiro-Wilk p-value ^b	0.002	N/A

ANCOVA: analysis of covariance; CI: confidence interval; FC: functional class; LS: least squares; N/A: not applicable; PAH-SYMPACT: Pulmonary Arterial Hypertension-Symptoms and Impact Questionnaire; SE: standard error; WHO: World Health Organization

^a ANCOVA with treatment, randomisation stratification factors of baseline (mono/double, triple) and baseline WHO FC (II, III) as fixed factors, baseline in selected domain scores from PAH-SYMPACT® as covariate. Non-existent data due to death and missing data due to non-fatal clinical worsening was assigned/imputed to 1 unit worse than the worst observed change from baseline from both treatment groups.

^b P-value is from Shapiro-Wilk test on the residuals from the ANCOVA model. If the normality assumption is violated (p-value <0.001), p-value result from ANCOVA are not presented.

Database Cutoff Date: 06DEC2022

Source: MSD 2024 STELLAR CSR (29)

PAH-SYMPACT® Cardiopulmonary Symptoms

The change from baseline in Cardiopulmonary Symptoms domain score of PAH-SYMPACT® showed improvement that was significantly greater in the sotatercept group than in the placebo group at Week 24 (Visit 9; Table 43) (29). The median of the treatment difference between the sotatercept and placebo groups was -0.13 (p=0.037; Table 44) (29). Results from the supportive analysis using the ANCOVA model were consistent with those from the primary analysis (Table 45) (29).

Table 43 Change from baseline in Cardiopulmonary Symptoms Domain Score of PAH-SYMPACT® by visit - Full Analysis Set

	Sotatercept (N=163)	Placebo (N=160)
Baseline		
n	115	117
Mean (SD)	0.9 (0.56)	0.8 (0.54)
Median	0.9	0.8

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

	Sotatercept (N=163)	Placebo (N=160)
Min, Max	0.0, 2.3	0.0, 2.4
Missing	48	43
Week 24		
n	122	113
Mean (SD)	0.8 (0.53)	0.8 (0.59)
Median	0.7	0.7
Min, Max	0.0, 2.3	0.0, 2.8
Missing	41	47
Change from baseline at Week 24		
n	96	93
Mean (SD)	-0.1 (0.44)	0.0 (0.39)
Median	-0.1	-0.0
Min, Max	-1.6, 1.0	-1.4, 1.5
Missing	67	67

PAH-SYMPACT: Pulmonary Arterial Hypertension-Symptoms and Impact Questionnaire; SD: standard deviation
Database Cutoff Date: 06DEC2022

Baseline is the average of the two screening visits.

Source: MSD 2024 STELLAR CSR (29)

Table 44 Analysis of change from baseline in Cardiopulmonary Symptoms Domain Score of PAH-SYMPACT® at Week 24 - Full Analysis Set (Standard multiple imputation)

	Sotatercept (N=163)	Placebo (N=160)
Median estimate (Min, Max) ^a	-0.11 (-0.14, -0.06)	-0.01 (-0.03, 0.02)
Hodges-Lehmann location shift (ASE) ^b	-0.13 (0.062)	N/A
95% CI of Hodges-Lehmann location shift	(-0.249, -0.007)	N/A
Wilcoxon p-value ^c	0.037	N/A

ASE: asymptomatic standard error; CI: confidence interval; N/A: not applicable; PAH-SYMPACT: Pulmonary Arterial Hypertension-Symptoms and Impact Questionnaire
Database Cutoff Date: 06DEC2022

^a Average, minimum and maximum of the medians across the imputed datasets if missing data imputed.

^b Hodges-Lehmann location shift from placebo estimate (median of all paired differences).

^c Wilcoxon p-value refers to p-value from the aligned rank stratified Wilcoxon test with randomisation factors as strata.

Change from baseline in SYMPACT® scores at Week 24 for participants who died was assigned as 200 to receive the worst rank.

Change from baseline in SYMPACT® scores at Week 24 for participants who had missing data due to a non-fatal clinical worsening event was imputed as 150 to receive the next worst rank.

Source: MSD 2024 STELLAR CSR (29)

Table 45 Analysis of change from baseline in Cardiopulmonary Symptoms Domain Score of PAH-SYMPACT® - Full Analysis Set (Standard multiple imputation with ANCOVA model)

	Sotatercept (N=163)	Placebo (N=160)
Change from Baseline at Week 24^a		
LS Mean (SE)	-0.051 (0.076)	0.162 (0.077)
95% CI of LS Mean	(-0.200, 0.097)	(0.010, 0.314)

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

	Sotatercept (N=163)	Placebo (N=160)
LS Mean Diff (SE)	-0.213 (0.104)	N/A
95% CI of LS Mean Diff	(-0.418, -0.008)	N/A
Shapiro-Wilk p-value ^b	<.001	N/A

ANCOVA: analysis of covariance; CI: confidence interval; FC: functional class; LS: least squares; N/A: not applicable; PAH-SYMPACT: Pulmonary Arterial Hypertension-Symptoms and Impact Questionnaire; SE: standard error; WHO: World Health Organization
Database Cutoff Date: 06DEC2022

^a ANCOVA with treatment, randomisation stratification factors of baseline (Mono/Double, Triple) and baseline WHO FC (II, III) as fixed factors, baseline in selected domain scores from PAH-SYMPACT[®] as covariate. Non-existent data due to death and missing data due to non-fatal clinical worsening was assigned/imputed to 1 unit worse than the worst observed change from baseline from both treatment groups.

^b P-value is from Shapiro-Wilk test on the residuals from the ANCOVA model. If the normality assumption is violated (p-value <0.001), p-value results from ANCOVA are not presented.

Source: MSD 2024 STELLAR CSR (29)

PAH-SYMPACT[®] Cognitive/Emotional Impacts

The difference in the change from baseline in Cognitive/Emotional Impacts domain score of PAH-SYMPACT[®] between the sotatercept and placebo groups at Week 24 (Visit 9) was not statistically significant (Table 46, Table 47) (29). Results from the supportive analysis using the ANCOVA model were consistent with those from the primary analysis (Table 48) (29).

While a directional benefit was demonstrated, there are several reasons why the PAH-SYMPACT[®] may not have reported a statistically significant difference in the cognitive and emotional impact component (29). First, it may have been difficult to observe a substantial change in cognitive/emotional quality of life (QoL) functioning over a 24-week period, as cognitive and emotional responses tend to develop over a longer period and consequently are not usually quickly changed (29).

In addition, other, non-clinical factors within a person's environment and personality can affect how they respond to their disease. Given the chronic and debilitating nature of PAH, it would be expected that improvements in clinical outcomes might not have a strong impact on patients' cognitive and emotional outcomes in the short-term and the cognitive/emotional domain does not appear to be sensitive to clinical improvements (34, 35). Further, the study was conducted during the coronavirus disease 2019 (COVID-19) pandemic, which serves as a potential confounder of the true impact of treatment on cognitive and emotional status in both clinical trial arms. These findings are further evidence that patients' cognitive and emotional status does not appear to be driven by treatment-or clinical-related changes in disease status but rather may be more indicative of underlying responses to the condition itself.

Table 46 Change from baseline in Cognitive/Emotional Impacts Domain Score of PAH-SYMPACT® by visit - Full Analysis Set

	Sotatercept (N=163)	Placebo (N=160)
Baseline		
n	117	123
Mean (SD)	0.9 (0.76)	0.9 (0.80)
Median	0.8	0.8
Min, Max	0.0, 2.8	0.0, 3.5
Missing	46	37
Week 24		
n	115	110
Mean (SD)	0.7 (0.67)	0.9 (0.81)
Median	0.5	0.8
Min, Max	0.0, 2.5	0.0, 3.3
Missing	48	50
Change from baseline at Week 24		
n	86	92
Mean (SD)	-0.1 (0.55)	0.1 (0.75)
Median	0.0	0.0
Min, Max	-2.0, 1.8	-2.3, 2.5
Missing	77	68

PAH-SYMPACT: Pulmonary Arterial Hypertension-Symptoms and Impact Questionnaire; SD: standard deviation
Database Cutoff Date: 06DEC2022

Baseline is the average of the two screening visits.

Source: MSD 2024 STELLAR CSR (29)

Table 47 Analysis of change from baseline in Cognitive/Emotional Impacts Domain Score of PAH-SYMPACT® at Week 24 - Full Analysis Set (Standard multiple imputation)

	Sotatercept (N=163)	Placebo (N=160)
Median estimate (Min, Max) ^a	0.000 (0.000, 0.000)	0.000 (0.000, 0.001)
Hodges-Lehmann location shift (ASE) ^b	-0.15 (0.121)	N/A
95% CI of Hodges-Lehmann location shift	(-0.383, 0.091)	N/A
Wilcoxon p-value ^c	0.186	N/A

ASE: asymptomatic standard error; CI: confidence interval; N/A: not applicable; PAH-SYMPACT: Pulmonary Arterial Hypertension-Symptoms and Impact Questionnaire

Database Cutoff Date: 06DEC2022

^a Average, minimum and maximum of the medians across the imputed datasets if missing data imputed.

^b Hodges-Lehmann location shift from placebo estimate (median of all paired differences).

^c Wilcoxon p-value refers to p-value from the aligned rank stratified Wilcoxon test with randomisation factors as strata.

Change from baseline in SYMPACT® scores at Week 24 for participants who died was assigned as 200 to receive the worst rank.

Change from baseline in SYMPACT® scores at Week 24 for participants who had missing data due to a non-fatal clinical worsening event was imputed as 150 to receive the next worst rank.

Source: MSD 2024 STELLAR CSR (29)

Table 48 Analysis of change from baseline in Cognitive/Emotional Impacts Domain Score of PAH-SYMPACT® at Week 24 - Full Analysis Set (Standard multiple imputation with ANCOVA model)

	Sotatercept (N=163)	Placebo (N=160)
Change from baseline at Week 24^a		
LS mean (SE)	0.004 (0.115)	0.253 (0.112)
95% CI of LS mean	(-0.222, 0.230)	(0.034, 0.472)
LS mean diff (SE)	-0.249 (0.159)	N/A
95% CI of LS mean diff	(-0.561, 0.063)	N/A
Shapiro-Wilk p-value ^b	<.001	N/A

ANCOVA: analysis of covariance; CI: confidence interval; FC: functional class; LS: least squares; N/A: not applicable; PAH-SYMPACT: Pulmonary Arterial Hypertension-Symptoms and Impact Questionnaire; SE: standard error; World Health Organization

Database Cutoff Date: 06DEC2022

^a ANCOVA with treatment, randomisation stratification factors of baseline (Mono/Double, Triple) and baseline WHO FC (II, III) as fixed factors, baseline in selected domain scores from PAH-SYMPACT® as covariate. Non-existent data due to death and missing data due to non-fatal clinical worsening was assigned/imputed to 1 unit worse than the worst observed change from baseline from both treatment groups.

^b P-value is from Shapiro-Wilk test on the residuals from the ANCOVA model. If the normality assumption is violated (p-value <0.001), p-value result from ANCOVA are not presented.

Source: MSD 2024 STELLAR CSR (29)

EQ-5D

EQ-5D-3L index score

EQ-5D was an exploratory endpoint in the STELLAR study (29). There was no notable difference in the change from baseline in EQ-5D-3L index score between the sotatercept and placebo groups at Week 24 (Table 49) (29).

Table 49 Analysis of change from baseline in EQ-5D-3L Index Score based on the UK crosswalk (from EQ-5D-5L to EQ-5D-3L) mapping algorithm at Week 24 - Full Analysis Set

	Sotatercept (N=163)	Placebo (N=160)
Baseline		
n	125	126
Mean (SD)	0.75 (0.19)	0.76 (0.19)
Median	0.78	0.79
Min, Max	-0.03, 0.99	-0.10, 0.99
Missing	38	34
Week 24		
n	109	104
Mean (SD)	0.78 (0.20)	0.75 (0.19)
Median	0.80	0.79
Min, Max	-0.22, 0.99	0.05, 0.99
Missing	54	56
Change from baseline at Week 24		
n	90	89
Mean (SD)	0.02 (0.18)	-0.01 (0.18)

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

	Sotatercept (N=163)	Placebo (N=160)
Median	0.00	0.00
Min, Max	-0.82, 0.46	-0.71, 0.75
Missing	73	71
ANCOVA^a		
LS mean (SE)	0.02 (0.02)	-0.01 (0.02)
95% CI of LS mean	(-0.02, 0.05)	(-0.05, 0.02)
LS mean diff (SE)	0.03 (0.02)	N/A
95% CI of LS mean diff	(-0.02, 0.08)	N/A
Shapiro-Wilk p-value ^b	<.001	N/A

ANCOVA: analysis of covariance; CI: confidence interval; EQ-5D-3L: EuroQol 5-Dimension 3-Level; EQ-5D-5L: EuroQol 5-Dimension 5-Level; FC: functional class; LS: least squares; N/A: not applicable; SD: standard deviation; UK: United Kingdom; WHO: World Health Organization

Database Cutoff Date: 06DEC2022

^a ANCOVA with treatment, randomisation stratification factors of baseline (mono/double, triple) and baseline WHO FC (II, III) as fixed factors, baseline of EQ-5D-3L index score as covariate.

^b P-value is from Shapiro-Wilk test on the residuals from the ANCOVA model. If the normality assumption is violated (p-value <0.001), p-value result from ANCOVA are not presented.

No imputation was performed for missing data.

EQ-5D UK crosswalk (from EQ-5D-5L to EQ-5D-3L) algorithm applied based on reference: Van Hout B et al. (2012) "Interim Scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L Value Sets", Value in Health Volume 15, Issue 5, July-August 2012, Pages 708-715.

EQ-5D-3L index scores (mapped crosswalk) derived based on reference: Alava, H. et al. "Estimating EQ-5D by age and sex for the UK", DSU, January 2022.

Source: MSD 2024 STELLAR CSR (29)

EQ-5D-5L VAS

The increase from baseline in EQ-5D-5L VAS was greater in the sotatercept group than in the placebo group at Week 24 (Table 50) (29).

Table 50 Change from baseline in EQ-5D-5L VAS at Week 24 - Full Analysis Set

	Sotatercept (N=163)	Placebo (N=160)
Baseline		
n	125	126
Mean (SD)	68.53 (17.82)	68.94 (19.01)
Median	70.00	70.50
Min, Max	23.00, 100.00	9.00, 100.00
Missing	38	34
Week 24		
n	109	104
Mean (SD)	72.85 (18.10)	67.18 (20.14)
Median	76.00	70.00
Min, Max	25.00, 100.00	5.00, 100.00
Missing	54	56
Change from baseline at Week 24		
n	90	89
Mean (SD)	4.33 (16.58)	-0.87 (19.00)
Median	5.00	0.00

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

	Sotatercept (N=163)	Placebo (N=160)
Min, Max	-36.00, 71.00	-65.00, 74.00
Missing	73	71
ANCOVA^a		
LS mean (SE)	4.37 (1.74)	-0.91 (1.73)
95% CI of LS mean	(0.94, 7.80)	(-4.33, 2.51)
LS mean diff (SE)	5.28 (2.39)	N/A
95% CI of LS mean diff	(0.56, 10.00)	N/A
p-value	0.029	N/A
Shapiro-Wilk p-value ^b	0.002	N/A

ANCOVA: analysis of covariance; CI: confidence interval; EQ-5D-5L: EuroQol 5-Dimension 5-Level; FC: functional class; LS: least squares; N/A: not applicable; SD: standard deviation; SE: standard error; VAS: visual analogue scale; WHO: World Health Organization

Database Cutoff Date: 06DEC2022

^a ANCOVA with treatment, randomisation stratification factors of baseline (mono/double, triple) and baseline WHO FC (II, III) as fixed factors, baseline of EQ-5D-5L VAS as covariate.

^b P-value is from Shapiro-Wilk test on the residuals from the ANCOVA model. If the normality assumption is violated (p-value <0.001), p-value result from ANCOVA are not presented.

No imputation was performed for missing data.

Source: MSD 2024 STELLAR CSR (29)

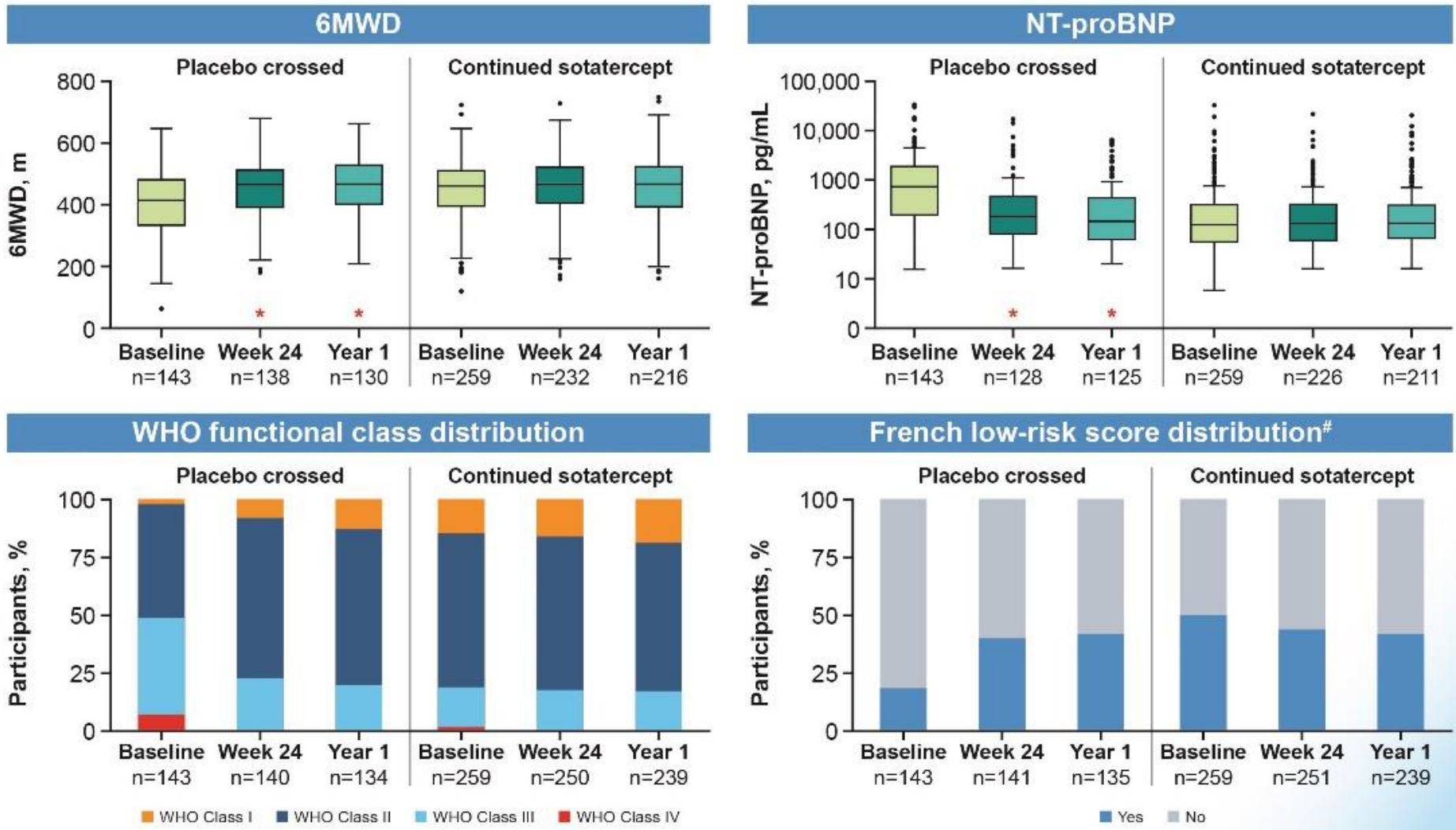
2.6.2 SOTERIA study – clinical effectiveness results

Results from the SOTERIA study are available from a cumulative cutoff date of 08-NOV-2023 and are summarised in this section (6, 30, 36, 37).

Clinical efficacy measures

At one year from baseline in SOTERIA, improvements in efficacy endpoints appear maintained (6, 37). Figure 9 and Table 51 show mean change in 6MWD, NT-proBNP, WHO FC II and low French Risk Score from SOTERIA baseline (at SOTERIA week 24) and at one year (6, 37).

Figure 9 SOTERIA study clinical efficacy results summary



6MWD: 6-minute walking distance; FC: functional class; NT-proBNP: N-terminal pro-brain natriuretic peptide; SD: standard deviation; WHO: World Health Organization
#Simplified French low risk score was defined as the presence of three low risk criteria: WHO functional class I-II, 6MWD >440 m, and NT-proBNP <300pg/mL.

Baseline is the measurement at Visit 1 or the last measure from the parent study before rollover. Week 24 and 1 year correspond to Visits 9 and 17, respectively. If a visit was missing, then another visit within a 12-week window was used.
SOTERIA cumulative data cutoff date: 08 November 2023 (30).

Table 51 SOTERIA: change from Baseline^a at Week 24 and Year 1 in clinical parameters

	Placebo-Crossed			Continued Sotatercept			Blinded				Total	
	Baseline	Week 24	Year 1	Baseline	Week 24	Year 1	Baseline	Week 24	Year 1	Baseline	Week 24	Year 1
6MWD, m												
N	143	138	130	259	232	214	24	10	3	426	380	347
Mean (SD)	404.5 (103.3)	448.3 (93.1)	455.3 (91.8)	446.3 (104.6)	455.9 (100.7)	452.8 (106.1)	235.0 (129.1)	300.2 (99.6)	364.0 (63.5)	420.4 (116.4)	449.0 (100.8)	453.1 (100.7)
Mean (SD) change from baseline	—	41.3 (74.3)	47.3 (80.7)	—	-0.02 (37.2)	-1.7 (47.6)	—	92.9 (94.2)	122.0 (201.3)	—	17.4 (59.9)	17.8 (68.7)
p-value	—	<0.0001	<0.0001	—	0.80	0.75	—	0.02	0.08	—	<0.0001	<0.0001
NT-proBNP, pg/mL												
N	143	128	125	259	226	211	24	9	3	426	363	339
Mean (SD)	1245.5 (2519.8)	437.9 (1124.7)	339.9 (642.3)	364.8 (1279.1)	314.7 (953.0)	326.6 (981.9)	4926.8 (6668.9)	828.7 (1397.0)	189.7 (147.1)	917.5 (2579.1)	370.9 (1029.7)	330.3 (866.4)
Mean (SD) change from baseline	—	-828.3 (1861.5)	-826.9 (2099.1)	—	-31.0 (987.4)	-32.8 (1145.3)	—	-3436.0 (4388.3)	-1806.3 (2165.3)	—	-396.6 (1620.6)	-341.3 (1620.1)
p-value	—	<0.0001	<0.0001	—	0.39	0.52	—	<0.0001	0.01	—	<0.0001	<0.0001
WHO FC II/I												
N	143	140	134	259	250	239	24	10	3	426	400	376
n (%)	73 (51.0)	108 (77.1)	107 (79.9)	209 (80.7)	205 (82.0)	197 (82.4)	1 (4.2)	3 (30)	1 (33.3)	283 (66.4)	316 (79.0)	305 (81.1)

	Placebo-Crossed			Continued Sotatercept			Blinded				Total	
	Baseline	Week 24	Year 1	Baseline	Week 24	Year 1	Baseline	Week 24	Year 1	Baseline	Week 24	Year 1
Low French risk score												
N	143	141	135	259	251	239	24	10	3	426	402	377
Yes^b, n (%)	27 (18.9)	57 (40.4)	57 (42.2)	130 (50.2)	110 (43.8)	101 (42.3)	0	0	0	157 (36.9)	167 (41.5)	158 (41.9)

^a Baseline is the measurement at Visit 1 of SOTERIA or the last measure from the parent study before rollover.

^b "Yes" means participants met all three criteria for simplified French low risk: WHO FC I-II, 6MWD >440 m, and NT-proBNP <300 pg/mL. Participants who did not meet all three criteria for simplified French low risk are not shown.

6MWD, 6-minute walk distance; N, number of participants that completed the visit and the assessment; NT-proBNP, N-terminal pro-B-type natriuretic peptide; SD, standard deviation; WHO FC, World Health Organization functional class.

Source: SOTERIA cumulative data cutoff date: 08 November 2023 (6, 37)

Analysis of long-term effect of sotatercept in STELLAR through SOTERIA

The results presented in this subsection refer specifically to those participants in the SOTERIA study who were originally in the sotatercept treatment arm of the STELLAR study (36).

Clinical worsening events

The number of patients who had at least one clinical worsening event was low in SOTERIA at the time of the analysis: 20 patients (4.7%) experienced 22 clinical worsening events (8 PAH-related hospitalisation, 6 deaths, 4 deteriorations of PAH, 2 worsening-related listings for lung and/or heart transplant and 2 initiations of rescue therapy or increase of prostacyclin dose by $\geq 10\%$; summarised in Table 52) (6, 37).

Table 52 SOTERIA study summary of clinical worsening events

	Placebo crossed (N=143)	Continued sotatercept (N=259)	Blinded (N=24)	All participants (N=426)
Participants with at least one clinical worsening event, n (%)	4 (2.8%)	14 (5.4%)	2 (8.3%)	20 (4.7%)
Total no. of clinical worsening events	4	16	2	22
Median (range) time to first event, weeks	33.4 (0.1, 36.3)	38.6 (6.7, 88.3)	0.14 (0.1, 0.1)	35.6 (0.1, 88.3)
Assessment of first clinical worsening event^a, n (%)				
Death	2 (1.4%)	3 (1.2%)	1 (4.2%)	6 (1.4%)
Worsening-related listing for lung and/or heart transplant	2 (1.4%)	NR	NR	2 (0.5%)
Initiation of rescue therapy or increase infusion prostacyclin dose by $\geq 10\%$	NR	2 (0.8%)	NR	2 (0.5%)
PAH-related hospitalisation (≥ 24 h)	1 (0.7%)	6 (2.3%)	1 (4.2%)	8 (1.9%)
Deterioration of PAH [†]	NR	3 (1.2%)	1 (4.2%)	4 (0.9%)

6MWD: 6-minute walking distance; FC: functional class; NR: not reported; PAH: pulmonary arterial hypertension; WHO: World Health Organization

SOTERIA cumulative data cutoff date: 08 November 2023.

^a No participant needed to have atrial septostomy as assessment of first clinical worsening event. A participant could have more than one assessment as the first clinical worsening event.

[†] Defined by both worsened WHO FC and decrease in 6MWD by $\geq 15\%$ (confirmed by two 6-minute walk tests ≥ 4 hours apart but ≤ 1 week), compared with baseline values, even if these two events began at different times.

Sources: Preston 2024 (37) and Preston et al. 2025 (6)

WHO FC improvement or maintenance

Analyses were performed for the outcomes of WHO FC improvement or maintenance (Table 53 and Table 54) (38). The findings of these analyses show that long-term treatment with sotatercept resulted in an enduring treatment effect for these outcomes (38).

Table 53 Analysis of the proportion of WHO FC improvement or maintenance relative to STELLAR baseline - Participants who enrolled to SOTERIA from STELLAR sotatercept arm and continued receiving sotatercept

Endpoint	Sotatercept (N ^a = ■)		
	n/N ^b	%	(95% CI) ^c
WHO FC improvement or maintenance relative to STELLAR baseline			
Week 3	■	■	■
Week 12	■	■	■
Week 24	■	■	■
Week 36	■	■	■
Week 52 (Year 1)	■	■	■
Week 108 (Year 2)	■	■	■

CI: confidence interval; FC: functional class; n: number of participants with improvement or maintenance of the endpoint assessment at a given timepoint relative to baseline; WHO: World Health Organization

Database Cutoff Date: STELLAR (MK-7962-003) - 06DEC2022 / SOTERIA (MK-7962-004) - 08NOV2023

^a Participants who enrolled to SOTERIA from STELLAR sotatercept arm and continued receiving sotatercept. 154 participants from STELLAR sotatercept arm enrolled to SOTERIA, but one participant discontinued before receiving study dose in SOTERIA.

^b Participants with non-missing endpoint assessment at both given timepoint and STELLAR baseline. Row percentages are based on this number.

^c Derived based on the Clopper-Pearson method (exact binomial distribution).

Risk status assessment is based on COMPERA version 2.0.

Source: MSD 2024 [Confidential Data on File] (38)

Table 54 Analysis of the proportion of WHO FC improvement or maintenance relative to STELLAR Week 24 - Participants who enrolled to SOTERIA from STELLAR sotatercept arm and continued receiving sotatercept

Endpoint	Sotatercept (N ^a = ■)		
	n/N ^b	%	(95% CI) ^c
WHO FC improvement or maintenance relative to STELLAR Week 24			
Week 36	■	■	■
Week 52 (Year 1)	■	■	■
Week 108 (Year 2)	■	■	■

CI: confidence interval; FC: functional class; n: number of participants with improvement or maintenance of the endpoint assessment at a given timepoint relative to week 24; WHO: World Health Organization

Database Cutoff Date: STELLAR (MK-7962-003) - 06DEC2022 / SOTERIA (MK-7962-004) - 08NOV2023

^a Participants who enrolled to SOTERIA from STELLAR sotatercept arm and continued receiving sotatercept. 154 participants from STELLAR sotatercept arm enrolled to SOTERIA, but one participant discontinued before receiving study dose in SOTERIA.

^b Participants with non-missing endpoint assessment at both given timepoint and STELLAR week 24. Row percentages are based on this number.

^c Derived based on the Clopper-Pearson method (exact binomial distribution).

Source: MSD 2024 [Confidential Data on File] (38)

Risk Status improvement or maintenance

Analyses were performed for the outcomes of Risk Status improvement or maintenance ([Table 55](#) and [Table 56](#)) (38). The findings of these analyses show that long-term treatment with sotatercept resulted in an enduring treatment effect for these outcomes (38).

Table 55 Analysis of the proportion of risk status improvement or maintenance relative to STELLAR baseline - Participants who enrolled to SOTERIA from STELLAR sotatercept arm and continued receiving sotatercept

Endpoint	Sotatercept (N ^a = ■)		
	n/N ^b	%	(95% CI) ^c
Risk status improvement or maintenance relative to STELLAR baseline			
Week 3	■	■	■
Week 12	■	■	■
Week 24	■	■	■
Week 36	■	■	■
Week 52 (Year 1)	■	■	■
Week 108 (Year 2)	■	■	■

CI: confidence interval; n: number of participants with improvement or maintenance of the endpoint assessment at a given timepoint relative to baseline

Database Cutoff Date: STELLAR (MK-7962-003) - 06DEC2022 / SOTERIA (MK-7962-004) - 08NOV2023

^a Participants who enrolled to SOTERIA from STELLAR sotatercept arm and continued receiving sotatercept. 154 participants from STELLAR sotatercept arm enrolled to SOTERIA, but one participant discontinued before receiving study dose in SOTERIA.

^b Participants with non-missing endpoint assessment at both given timepoint and STELLAR baseline. Row percentages are based on this number.

^c Derived based on the Clopper-Pearson method (exact binomial distribution).

Risk status assessment is based on COMPERA version 2.0.

Source: MSD 2024 [Confidential Data on File] (38)

Table 56 Analysis of the proportion of risk status improvement or maintenance relative to STELLAR Week 24 - Participants who enrolled to SOTERIA from STELLAR sotatercept arm and continued receiving sotatercept

Endpoint	Sotatercept (N ^a = ■)		
	n/N ^b	%	(95% CI) ^c
Risk status improvement or maintenance relative to STELLAR Week 24			
Week 36	■	■	■
Week 52 (Year 1)	■	■	■
Week 108 (Year 2)	■	■	■

CI: confidence interval; N: number of participants with improvement or maintenance of the endpoint assessment at a given timepoint relative to week 24

Database Cutoff Date: STELLAR (MK-7962-003) - 06DEC2022 / SOTERIA (MK-7962-004) - 08NOV2023

^a Participants who enrolled to SOTERIA from STELLAR sotatercept arm and continued receiving sotatercept. 154 participants from STELLAR sotatercept arm enrolled to SOTERIA, but one participant discontinued before receiving study dose in SOTERIA.

^b Participants with non-missing endpoint assessment at both given timepoint and STELLAR week 24. Row percentages are based on this number.

^c Derived based on the Clopper-Pearson method (exact binomial distribution).

Source: MSD 2024 [Confidential Data on File] (38)

Overall survival

Analyses were also performed for the outcome of overall survival (6). When patients who were randomised to sotatercept in the STELLAR study (and continued sotatercept treatment during the SOTERIA study) are compared to patients randomised to placebo in the STELLAR study (who then received treatment with sotatercept during the SOTERIA study), the hazard ratio is ■; Table 57) (38).

However, following the intent-to-treat principle where participants are analysed in the treatment arms they were originally randomised to without taking into account treatment switching does not identify the true comparative effectiveness of the treatments under investigation and underestimates true treatment effect as treatment switch introduces bias (i.e., outcomes in "placebo" arm are affected by these patients receiving the experimental drug in the SOTERIA study phase). Overall survival was therefore also analysed where treatment switching in the placebo arm is adjusted for using the Rank-Preserving Structural Failure time (RPSFT) method.

While other methodologies that accommodate treatment switching exist, such as Two-stage adjustment and Inverse Probability of Censoring Weighting (IPCW) but these are not applicable to the STELLAR study design for the following key reasons:

- Two-stage and IPCW: There is high proportion of treatment switchers out of the 148 eligible participants (who completed DBPC period) from the placebo arm in the STELLAR study to sotatercept in the SOTERIA study ($143/148 = 96.6\%$). If there are very few placebo arm participants who did not switch treatments, some participants may be assigned very high weights. In this case, the method becomes prone to substantial error, because there is a higher likelihood that these remaining participants may not be representative of the wider trial population.
- Two-stage: With this method, switching must only occur soon after a pre-defined secondary baseline (i.e., specific disease-related time point) otherwise differences between switchers and non-switchers may not be adequately adjusted for. In STELLAR, treatment switching is optional and not defined by a disease-related reason (the only condition is to complete at least the Week 24 assessment of the DBPC period).

When treatment switching is adjusted for, the hazard ratio improves to [REDACTED]; Table 58, Figure 10) showing even better overall survival in the patients treated with sotatercept from the beginning of STELLAR (38).

Table 57 Analysis of overall survival - Period on STELLAR and SOTERIA - Without adjusting for treatment switch to sotatercept in placebo arm (STELLAR Full Analysis Set population)

Study: STELLAR (MK-7962-003) ^a	Sotatercept			Placebo			Sotatercept vs. placebo	
	N ^b	Participants with event n (%)	Median time ^c in Weeks [95% CI]	N ^b	Participants with event n (%)	Median time ^c in Weeks [95% CI]	Hazard ratio [95% CI] ^d	p-value ^{d,e}
Overall Survival	■	■	■	■	■	■	■	■

CI: Confidence Interval; PAH: Pulmonary Arterial Hypertension; WHO: World Health Organization

^a Database Cutoff Date: STELLAR (MK-7962-003) - 06DEC2022/SOTERIA (MK-7962-004) - 08NOV2023

^b Number of participants: STELLAR full analysis set population. Participants alive are censored at the date of last contact in STELLAR or SOTERIA (if enrolled in SOTERIA) without adjusting for treatment switch from placebo in STELLAR to sotatercept in SOTERIA

^c From product-limit (Kaplan-Meier) method for censored data

^d Based on Cox regression model with treatment as a covariate stratified by the following covariates: WHO functional class (class II vs. III) and background PAH therapy (mono/double vs. triple therapy)

^e Two-sided p-value using Wald test (Score test in case of zero event in one treatment group)

Source: MSD 2024 [Confidential Data on File] (38)

Table 58 Analysis of overall survival - Period on STELLAR and SOTERIA - Adjusting for treatment switch to sotatercept in placebo arm using rank-preserving structural failure time model (STELLAR Full Analysis Set population)

Treatment	N	Number of events (%)	Person-weeks	Event rate/100 person-weeks (%)	Median OS ^a (Week) (95% CI)	Sotatercept vs. placebo	
						Hazard ratio ^b (95% CI) ^c	p-value ^d
Placebo	■	■	■	■	■	■	■
Placebo (RPSFT Adjusted)	■	■	■	■	■		
Sotatercept	■	■	■	■	■		

CI: Confidence Interval; OSa: overall survival analysis; PAH: Pulmonary Arterial Hypertension; RPSFT: Rank-Preserving Structural Failure Time; WHO: World Health Organization

Database Cutoff Date: STELLAR (MK-7962-003) - 06DEC2022/SOTERIA (MK-7962-004) - 08NOV2023

RPSFT model is used to adjust for the effect of cross-over from placebo in STELLAR to sotatercept in SOTERIA

Number of participants: STELLAR full analysis set population. Participants alive and did not enrol to SOTERIA are censored at the date of last contact in STELLAR.

Participants in the STELLAR sotatercept arm alive and enrolled to SOTERIA are censored at the date of last contact in SOTERIA. Participants in the STELLAR placebo arm

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

alive and enrolled to SOTERIA are censored at the date of last contact in SOTERIA but adjusted for the date they switched from placebo in STELLAR to sotatercept in SOTERIA

^a From product-limit (Kaplan-Meier) method for censored data

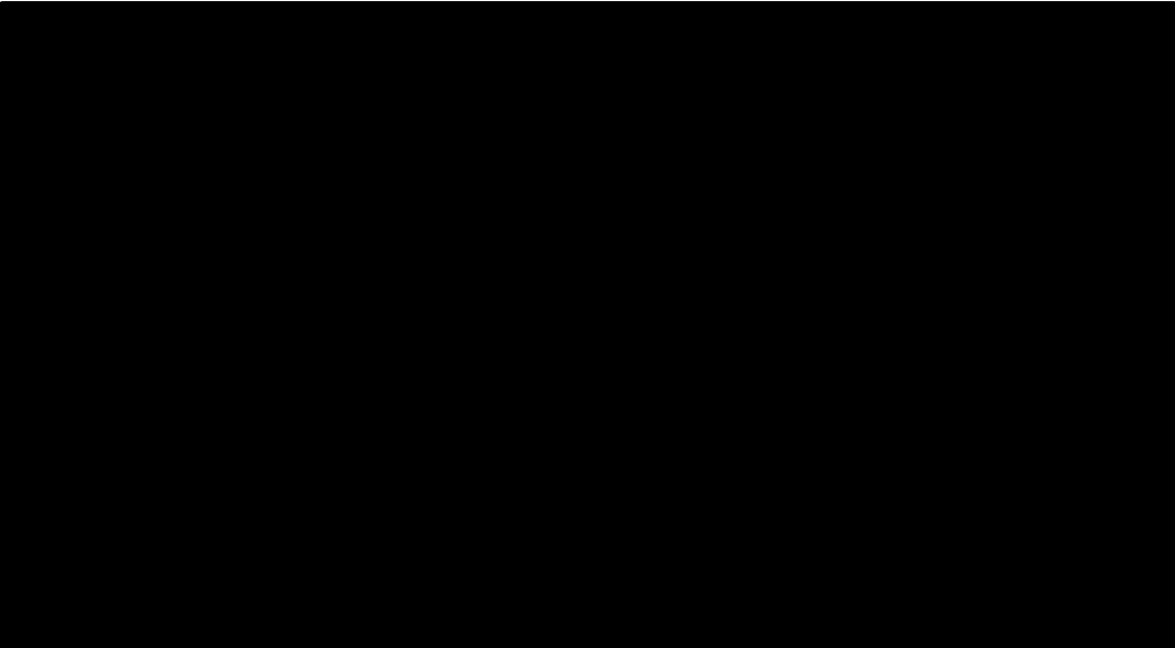
^b Based on Cox regression model corrected by RPSFT with treatment as a covariate stratified by the following covariates: WHO functional class (class II vs. III) and background PAH therapy (mono/double vs. triple therapy)

^c Obtained by inflating the standard error of the log-hazard ratio to preserve the full analysis set population p-value from the Cox model

^d Two-sided p-value using Wald test

Source: MSD 2024 [Confidential Data on File] (38)

Figure 10 Kaplan-Meier curves of overall survival - Period on STELLAR and SOTERIA - Adjusting for treatment switch to sotatercept in placebo arm using rank-preserving structural failure time model (STELLAR Full Analysis Set population)



RPSFT: Rank-Preserving Structural Failure Time
Source: MSD 2024 [Confidential Data on File] (38)

2.7 Subsequent treatments used in the relevant studies

As described in section 2.3.1 and in Appendix K, the STELLAR study consisted of a 24-week DBPC treatment period followed by an approximately 72-week long-term double-blind treatment period. Participants were permitted to roll over into the SOTERIA study if they experienced a clinical worsening event or completed the 24-week treatment period in the STELLAR trial (results for the population of patients of the SOTERIA study who entered from the STELLAR study are shown in section 2.6.2).

As described in section 1.3, if patients' condition worsens to the intermediate-high risk status (according to the ESC/ERS four-strata risk assessment tool), then patients in UK clinical practice are expected to be treated with PDE5i + ERA + PCA triple therapy.

2.8 Subgroup analysis

STELLAR study

Subgroup analyses and interaction tests

- Subgroup analyses and interaction tests were conducted to investigate whether, or to what extent, the clinical benefit of treatment with sotatercept in combination with stable background therapy versus placebo in combination with stable background therapy may differ between patients differentiated according to particular baseline characteristics (33).
- These analyses show that the results in subgroups are generally representative of those seen in the full population of the STELLAR study (33).
- Significant differences (as determined by p-value for interaction test <0.005) were observed in certain groups of patients (patients with baseline WHO FC II versus III and baseline PVR ≤ 800 dynes*sec/cm⁵ vs. >800 dynes*sec/cm⁵), however these findings are hypothesis-generating rather than ones from which firm conclusions to be used for clinical decision-making can be rigorously drawn (33).

Within-trial analyses of PDE5i + ERA + sotatercept versus PDE5i + ERA + selexipag

- For the outcome of change from baseline in risk status at Week 24, it is not possible to inform this comparison with the ITC using external trial data, described in section 2.10, as the required outcomes are not reported for those trials (33). The comparison for this outcome can be informed via a post-hoc within-trial analysis of the STELLAR study including only the participants in each arm treated with these regimens specifically (i.e., PDE5i + ERA patients randomised to receive sotatercept and PDE5i + ERA + selexipag patients randomised to receive placebo) (33).

- This analysis found that patients treated with PDE5i + ERA + sotatercept are more likely to have an improvement in risk status from baseline to Week 24 than patients treated with PDE5i + ERA + selexipag [REDACTED] as well as less likely to have a worsening in risk status [REDACTED] (33). Sample size for the analysis is low, reflecting the rare nature of PAH (33). The results of this within-trial analysis are used to inform the cost-effectiveness analyses described in section 3.

2.8.1 STELLAR study

Subgroup analyses and interaction tests

Given the rare nature of PAH, efforts have been taken to preserve sample size in the trial analyses where appropriate. Subgroup analyses and interaction tests were performed to investigate whether, or to what extent, clinical benefit of treatment with sotatercept in combination with stable background therapy versus placebo in combination with stable background therapy may differ between patients with certain characteristics. Forest plots of subgroup analyses for outcomes are shown in Figure 11 to Figure 13, with details of subgroup analyses for other outcomes provided in Appendix C (33). A summary of the results of the comprehensive array of interaction tests for several patient baseline characteristics categories across several study outcomes is shown in Table 59 with further details (including statistical methodological details) provided in Appendix C (33).

These analyses show that the findings observed in the specific subgroups are generally reflective of results in the overall population (33). However, significant differences in relative treatment effect as determined by p-value for interaction test <0.005 were observed in patients with baseline WHO FC II versus III, baseline PVR ≤ 800 dynes*sec/cm⁵ vs. >800 dynes*sec/cm⁵, and in baseline risk status, for certain outcomes (see Table 59) (33). As the STELLAR study was not designed or statistically powered to investigate whether there are true underlying differences for these, such findings are hypothesis-generating as opposed to ones from which firm conclusions to be used for clinical decision-making can be rigorously drawn (33). However, it indicates it is appropriate to use the full trial results when comparing sotatercept to appropriate comparators for certain subpopulations.

Table 59 Summary of STELLAR study subgroup analyses interaction test results

Characteristic	p-value for interaction test ^a												
	Change from baseline in 6-Minute Walk Distance (metres) at Week 24	Change from baseline in NT-proBNP (pg/mL) at Week 24	Change from baseline in PVR (dynes*sec/cm ⁵) at Week 24	Time to first clinical worsening or death	Change from baseline in WHO functional classification at Week 24			Change from baseline in risk status at Week 24			Change from baseline in French low risk criteria at Week 24		
					Improvement	Maintenance	Worsening	Improvement	Maintenance	Worsening	Improvement	Maintenance	Worsening
Sex ^b	■	■	■	■	■	■	■	■	■	■	■	■	■
PAH diagnostic group ^c	■	■	■	■	■	■	■	■	■	■	■	■	■
Background PAH therapy at baseline ^d	■	■	■	■	■	■	■	■	■	■	■	■	■
Prostacyclin infusion therapy at baseline ^e	■	■	■	■	■	■	■	■	■	■	■	■	■
WHO functional class at baseline ^f	■	■	■	■	■	■	■	■	■	■	■	■	■
PVR at baseline ^g	■	■	■	■	■	■	■	■	■	■	■	■	■
Risk status at baseline ^h	■	■	■	■	■	■	■	■	■	■	■	■	■

6MWD: 6-minute walking distance; FC: functional class; NT-proBNP: N-terminal pro-brain natriuretic peptide; PVR: pulmonary vascular resistance; WHO: World Health Organization

^a Based on Cochran's Q (chi-squared) two-sided p-value for heterogeneity test for 6MWD, NT-proBNP and PVR. Based on Cox regression model with treatment and subgroup as covariates, and treatment-by-subgroup interaction (p-value of likelihood ratio test for interaction term) for Time to First Clinical Worsening or Death. Based on Breslow-Day test for homogeneity of odds ratios stratified by subgroup for WHO FC, risk status and French risk.

^b Male vs. female.

^c Idiopathic vs. heritable vs. drug/toxin-induced vs. connective tissue disease vs. congenital heart disease with systemic to pulmonary (s/p) shunt repair.

^d Monotherapy vs. double vs. triple.

^e Yes vs. no.

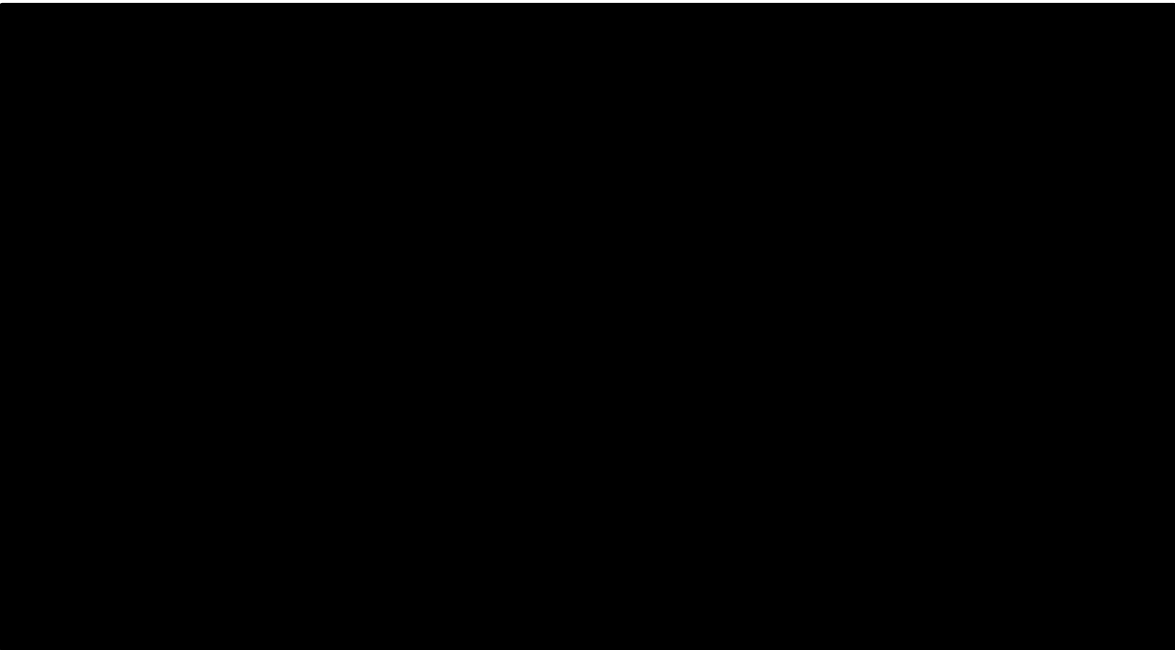
^f II vs. III.

^g ≤800 dynes*sec/cm⁵ vs. >800 dynes*sec/cm⁵.

^h Low risk vs. intermediate-low risk vs. intermediate-high risk.

Source: MSD 2024 [Confidential Data on File] (33)

Figure 11 Subgroups analyses of change from baseline in 6-Minute Walk Distance (metres) at Week 24 - Full Analysis Set (Standard multiple imputation)



6MWD: 6-minute walking distance; ASE: asymptomatic standard error; CI: confidence interval; HL: Hodges-Lehmann; PAH: pulmonary arterial hypertension; PVR: pulmonary vascular resistance; WHO: World Health Organization

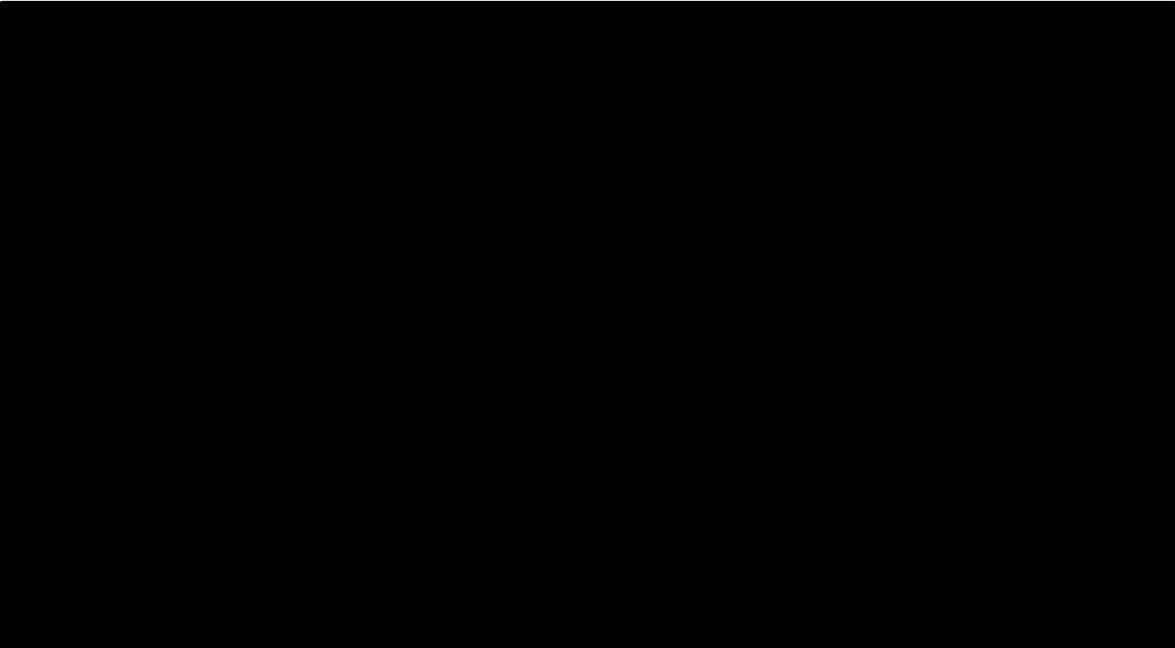
Study: STELLAR (MK-7962-003)

Participants who had missing 6MWD at Week 24 for reasons other than death or a clinical worsening event, a standard multiple imputation method was used to impute missing data used in computing change from baseline. Change from baseline in 6MWD at Week 24 for participants who died was assigned a value of -2000 metres to receive the worst rank. Change from baseline in 6MWD at Week 24 for participants who have missing data due to a non-fatal clinical worsening event was imputed to -1000 metres to receive the next worst rank

[1] Hodges-Lehmann Location Shift from the placebo estimate, median of all paired differences between the two treatment arms and its ASE estimate

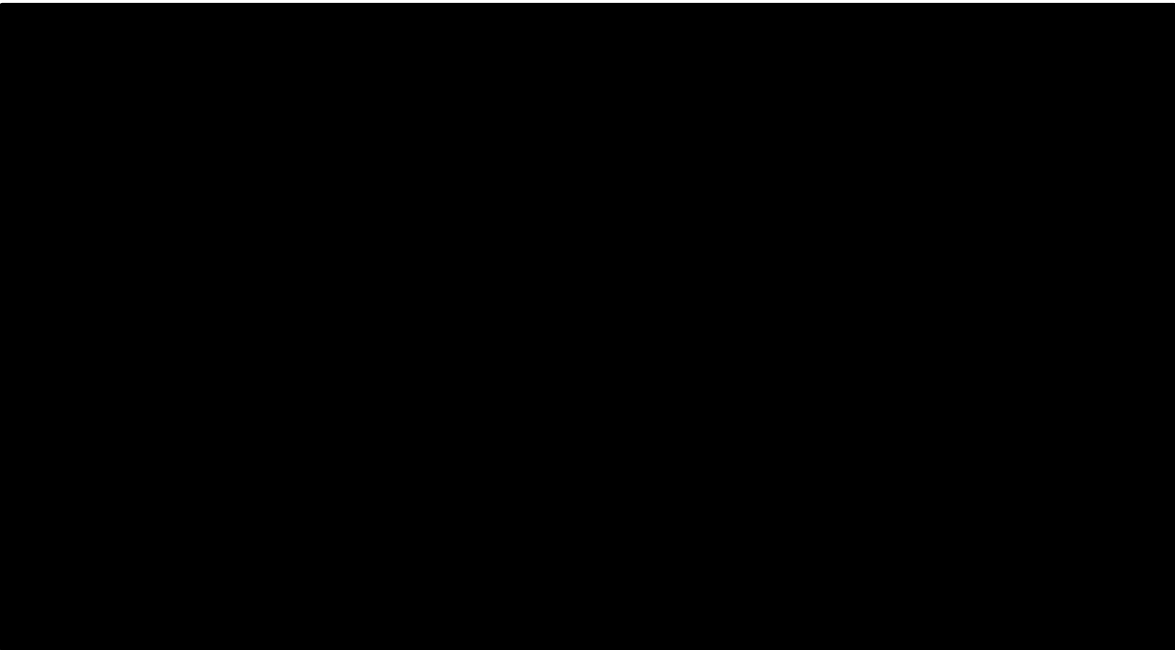
Source: MSD 2024 [Confidential Data on File] (33)

Figure 12 Subgroups analyses of change from baseline in NT-proBNP (pg/mL) at Week 24 - Full Analysis Set (Standard multiple imputation)



ASE: asymptomatic standard error; CI: confidence interval; HL: Hodges-Lehmann; NT-proBNP: N-terminal pro-brain natriuretic peptide; PAH: pulmonary arterial hypertension; PVR: pulmonary vascular resistance; WHO: World Health Organization
Study: STELLAR (MK-7962-003)
Participants who had missing NT-proBNP at Week 24 for reasons other than death or a clinical worsening event, a standard multiple imputation method was used to impute missing data used in computing change from baseline. Change from baseline in NT-proBNP at Week 24 for participants who died was assigned as 200000 to receive the worst rank. Change from baseline in NT-proBNP at Week 24 for participants who had missing data due to a non-fatal clinical worsening event was imputed as 150000 to receive the next worst rank
[1] Hodges-Lehmann Location Shift from the placebo estimate, median of all paired differences between the two treatment arms and its ASE estimate
Source: MSD 2024 [Confidential Data on File] (33)

Figure 13 Subgroups analyses of change from baseline in PVR (dynes*sec/cm⁵) at Week 24 - Full Analysis Set (Standard multiple imputation)



ASE: asymptomatic standard error; CI: confidence interval; HL: Hodges-Lehman; PAH: pulmonary arterial hypertension; PVR: pulmonary vascular resistance; WHO: World Health Organization

Participants who had missing PVR at Week 24 for reasons other than death or a clinical worsening event, a standard multiple imputation method was used to impute missing data used in computing change from baseline. Change from baseline in PVR at Week 24 for participants who died was assigned as 20000 to receive the worst rank. Change from baseline in PVR at Week 24 for participants who had missing data due to a non-fatal clinical worsening event was imputed as 15000 to receive the next worst rank

[1] Hodges-Lehmann Location Shift from the placebo estimate, median of all paired differences between the two treatment arms and its ASE estimate

Study: STELLAR (MK-7962-003) Source: MSD 2024 [Confidential Data on File] (33)

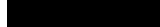
Subgroup results – Participants with baseline intermediate-low risk

Results for the subgroup of participants of the STELLAR study who had ESC/ERS intermediate-low risk at baseline are summarised in Table 60, full details of subgroup analyses methods and results are provided in Appendix C (33). This was not a pre-defined subgroup analysis of the STELLAR study.

Table 60 Summary of subgroup results for participants with baseline ESC/ERS intermediate-low risk

	Sotatercept	Placebo
6 Minute Walk Distance		
Change from baseline in 6-Minute Walk Distance (metres) at Week 24		
N		
Sotatercept vs. placebo: Hodges-Lehmann location shift (ASE) [95% CI]		
NT-proBNP		
Change from baseline in NT-proBNP (pg/mL) at Week 24		
N		
Sotatercept vs. placebo: Hodges-Lehmann location shift (ASE) [95% CI]		
PVR		
Change from baseline in PVR (dynes*sec/cm5) at Week 24		
N		
Sotatercept vs. placebo: Hodges-Lehmann location shift (ASE) [95% CI]		
Time to first clinical worsening or death		
Time to first clinical worsening or death		
N		
Participants with event n (%)		
Median time in months [95% CI]		
Sotatercept vs. placebo: Hazard ratio [95% CI]		
Change from baseline in WHO functional classification at Week 24 (imputed cases)		
WHO FC improvement at Week 24		
N		
Participants with event n (%)		
Sotatercept vs. placebo: Odds ratio [95% CI]		
Sotatercept vs. placebo: Relative risk [95% CI]		
WHO FC maintenance at Week 24		
N		
Participants with event n (%)		
Sotatercept vs. placebo: Odds ratio [95% CI]		
Sotatercept vs. placebo: Relative risk [95% CI]		
WHO FC worsening at Week 24		

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

	Sotatercept	Placebo
N		
Participants with event n (%)		
Sotatercept vs. placebo: Odds ratio [95% CI]		
Sotatercept vs. placebo: Relative risk [95% CI]		
Change from baseline in risk status at Week 24		
Risk status improvement at Week 24		
N		
Participants with event n (%)		
Sotatercept vs. placebo: Odds ratio [95% CI]		
Sotatercept vs. placebo: Relative risk [95% CI]		
Risk status maintenance at Week 24		
N		
Participants with event n (%)		
Sotatercept vs. placebo: Odds ratio [95% CI]		
Sotatercept vs. placebo: Relative risk [95% CI]		
Risk status worsening at Week 24		
N		
Participants with event n (%)		
Sotatercept vs. placebo: Odds ratio [95% CI]		
Sotatercept vs. placebo: Relative risk [95% CI]		
Change from baseline in French low risk criteria at Week 24		
French low risk criteria improvement at Week 24		
N		
Participants with event n (%)		
Sotatercept vs. placebo: Odds ratio [95% CI]		
Sotatercept vs. placebo: Relative risk [95% CI]		
French low risk criteria maintenance at Week 24		
N		
Participants with event n (%)		
Sotatercept vs. placebo: Odds ratio [95% CI]		
Sotatercept vs. placebo: Relative risk [95% CI]		
French low risk criteria worsening at Week 24		
N		
Participants with event n (%)		
Sotatercept vs. placebo: Odds ratio [95% CI]		
Sotatercept vs. placebo: Relative risk [95% CI]		

ASE: asymptomatic standard error; CI: confidence interval; ESC: European Society of Cardiology; ERS: European Respiratory Society; FC: functional class; NT-proBNP: N-terminal pro-brain natriuretic peptide; PVR: pulmonary vascular resistance; WHO: World Health Organization
Database Cutoff Date: 06DEC2022
Source: MSD 2024 [Confidential Data on File] (33)

Within-trial analyses of PDE5i + ERA + sotatercept versus PDE5i + ERA + selexipag

The expected position for sotatercept initiation in the treatment pathway in UK clinical practice is in patients with PAH with ESC/ERS four-strata intermediate-low risk, where the relevant sotatercept-based regimen would be PDE5i + ERA + sotatercept and the current SoC is PDE5i + ERA + selexipag (33).

For the outcome of change from baseline in risk status at Week 24, it is not possible to inform this comparison with the ITC using external trial data, described in section 2.10, as the required outcomes are not reported for those trials (33). In order to appropriately inform the cost-effectiveness analyses, the comparison for this outcome can be informed via a post-hoc within-trial analysis of the STELLAR study including only the participants in each arm treated with these regimens specifically (i.e., PDE5i + ERA patients randomised to receive sotatercept and PDE5i + ERA + selexipag patients randomised to receive placebo) (33). The findings of this analysis are informative despite the fact that this analysis breaks randomisation (as this is not a pre-specified subgroup analysis, and while randomisation was stratified based on background PAH therapy [mono/double or triple therapy], randomisation was not stratified based on these specific regimens) and is not powered (see Appendix L for details of sample size/statistical power considerations in the STELLAR statistical analysis plan) (33). Despite these limitations, this analysis is necessary to understand the relative effects of sotatercept and selexipag on risk status change when added to background dual therapy (33).

The patients included in this within-trial analysis were (33):

- PDE5i + ERA + sotatercept arm: patients randomised to the sotatercept (+ background PAH therapy) arm, who received dual therapy of PDE5i + ERA as the background PAH therapy.
- PDE5i + ERA + selexipag arm: patients randomised to the placebo (+ background PAH therapy) arm, who received triple therapy with PDE5i + ERA + selexipag as the background PAH therapy.

A description of the baseline characteristics of STELLAR study participants included in this within-trial analysis is shown in Table 61 (33). While there are differences in the baseline characteristics between the two groups of patients being compared, the results of this comparison are informative in the absence of alternative data (33). Without the results of these analyses, the relative efficacy of sotatercept vs. selexipag for the risk status change outcome remains a data gap (33).

Results are presented in Table 62 for the outcome of change from baseline in risk status at Week 24, a key clinical effectiveness outcome used in the cost-effectiveness analysis (detailed in section 3) (33). These show that patients treated with PDE5i + ERA + sotatercept are more likely to have an improvement in risk status from baseline to Week 24 than patients treated with PDE5i + ERA + selexipag [REDACTED] as well as less likely to have a worsening in risk status [REDACTED] (33). As mentioned earlier, the benefit of sotatercept has been shown to be independent of baseline risk status, and this is assumed to be maintained in this analysis (33).

Participant baseline characteristics

The baseline characteristics of the population included in the within-trial analysis are summarised in Table 61 (33). In assessing the comparability of the populations in this analysis, clinical expert opinion was sought on the most important baseline prognostic factors that may impact change in risk status (see Appendix O) (33). The following items which were available from the STELLAR trial were identified as important (33):

- PVR
- BNP/NT-proBNP
- 6MWD
- WHO FC

Despite randomisation being broken in this comparison, noteworthy imbalances in these patient baseline characteristics between the two arms were not observed (33). Patients in the selexipag arm had a slightly higher baseline PVR, slightly higher baseline NT pro-BNP, and slightly higher baseline proportion of FC III patients (potentially reflecting current reimbursement criteria for selexipag) (33). Conversely, the sotatercept arm had a slightly lower mean 6MWD and a greater proportion of patients in low risk at baseline (which restricts the capacity of sotatercept to demonstrate risk status improvements as one-third of patients were already in the best possible category) (33). As would be expected with patients receiving background triple therapy, the selexipag arm was associated with a slightly longer mean time since diagnosis (33).

Table 61 STELLAR study – within-trial analysis of the comparison of PDE5i + ERA + sotatercept versus PDE5i + ERA + selexipag – participant baseline characteristics

	PDE5i + ERA + sotatercept		PDE5i + ERA + selexipag + placebo	
	n	(%)	n	(%)
Participants in population	■		■	

	PDE5i + ERA + sotatercept		PDE5i + ERA + selexipag + placebo	
	n	(%)	n	(%)
Age (years)				
Participants with data	■		■	
Mean	■		■	
SD	■		■	
Median	■		■	
Range	■		■	
Age group, n (%)				
18 to <45	■	■	■	■
45 to <65	■	■	■	■
65 to <75	■	■	■	■
>=75	■	■	■	■
Sex, n (%)				
Male	■	■	■	■
Female	■	■	■	■
Race, n (%)				
White	■	■	■	■
Black or African American	■	■	■	■
Asian	■	■	■	■
Other	■	■	■	■
Missing	■	■	■	■
Ethnicity, n (%)				
Hispanic or Latino	■	■	■	■
Not Hispanic or Latino	■	■	■	■
Not reported	■	■	■	■
Height (cm)				
Participants with data	■		■	
Mean	■		■	
SD	■		■	
Median	■		■	
Range	■		■	
Weight (kg)				
Participants with data	■		■	
Mean	■		■	
SD	■		■	
Median	■		■	
Range	■		■	
BMI (kg/m²)				
Participants with data	■		■	
Mean	■		■	
SD	■		■	

	PDE5i + ERA + sotatercept		PDE5i + ERA + selexipag + placebo	
	n	(%)	n	(%)
Median	■		■	
Range	■		■	
Length of time since PAH diagnosis (years)				
Participants with data	■		■	■
Mean	■		■	■
SD	■		■	■
Median	■		■	■
Range	■		■	■
WHO diagnostic pulmonary hypertension Group I: PAH, n (%)				
Idiopathic PAH	■	■	■	■
Heritable PAH	■	■	■	■
Drug/toxin-induced PAH	■	■	■	■
PAH associated with connective tissue disease	■	■	■	■
PAH associated with simple, congenital systemic to pulmonary shunts ≥1 year following repair	■	■	■	■
WHO functional classification for symptomatic pulmonary hypertension^a, n (%)				
Class II	■	■	■	■
Class III	■	■	■	■
Type of background PAH therapy^b, n (%)				
Double	■	■	■	■
Triple	■	■	■	■
Exercise programme for cardiopulmonary rehabilitation^c, n (%)				
Yes	■	■	■	■
No	■	■	■	■
Risk status at baseline, n (%)				
Low risk	■	■	■	■
Intermediate-low risk	■	■	■	■
Intermediate-high risk	■	■	■	■
6-Minute walk distance (metres)				
n	■		■	
Mean (SD)	■		■	
Median	■		■	
Min, Max	■		■	
Missing	■		■	
NT-proBNP (pg/mL)				
n	■		■	
Mean (SD)	■		■	

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension
[ID6163]

	PDE5i + ERA + sotatercept		PDE5i + ERA + selexipag + placebo	
	n	(%)	n	(%)
Median	■		■	
Min, Max	■		■	
Missing	■		■	
PVR (dynes*sec/cm⁵)				
n	■		■	
Mean (SD)	■		■	
Median	■		■	
Min, Max	■		■	
Missing	■		■	

BMI: body mass index; CSR: clinical study report; ERA: endothelin receptor antagonist; FC: functional class; NT-proBNP: N-terminal pro-brain natriuretic peptide; PAH: pulmonary arterial hypertension; PDE5i: phosphodiesterase 5 inhibitor; PVR: pulmonary vascular resistance; SD: standard deviation; WHO: World Health Organization

Study: STELLAR (MK-7962-003)

Database Cutoff Date: 06DEC2022

Number of participants: full analysis set, participants who received PDE5i + ERA + sotatercept or PDE5i + ERA + selexipag.

^a One participant, in the sotatercept arm, is counted in the WHO FC II category in this table (in alignment with the CSR Shift from baseline in WHO functional class by Visit entries). To note is that the same participant, in the CSR Disease History Associated with PAH entry is counted in the WHO FC III category. This discrepancy was caused by an incorrect entry in the WHO FC intake form in the IRT and was recorded to match the IRT entry by the site in the study database. For purposes of efficacy and safety analyses, this participant is counted as WHO FC II.

^b One participant, in the placebo arm, is counted in the double therapy category in this table (in alignment with the CSR Disease History Associated with PAH entry). To note is that the same participant, in the CSR Concomitant Standard of Care Medications is counted in the triple therapy category. The discrepancy was caused by an incomplete end date record in the study database for a concomitant PAH therapy. For purposes of efficacy and safety analyses, this participant is counted as being on double therapy.

^c Risk stratum assessment is based on COMPERA version 2.0.

Source: MSD 2024 [Confidential Data on File] (33)

Analysis of change from baseline in risk status at Week 24

Table 62 Analysis of change from baseline in risk status at Week 24 – STELLAR study – Within-trial analysis of the comparison of PDE5i + ERA + sotatercept versus PDE5i + ERA + selexipag – Full Analysis Set

	PDE5i + ERA + sotatercept (N=36)	PDE5i + ERA + selexipag + placebo (N=26)
Improvement of risk status, m	■	■
Yes, n (%)	■	■
No, n (%)	■	■
p-value ^a	■	■
Missing due to COVID-19	■	■
Relative risk of risk stratum improvement, sotatercept versus selexipag (95% CI) ^b	■	
Maintenance of risk status, m	■	■
Yes, n (%)	■	■
No, n (%)	■	■

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

	PDE5i + ERA + sotatercept (N=36)	PDE5i + ERA + selexipag + placebo (N=26)
p-value ^a		
Missing due to COVID-19		
Worsening of risk status, m		
Yes, n (%)		
No, n (%)		
p-value ^a		
Missing due to COVID-19		
Relative risk of risk stratum worsening, sotatercept vs. selexipag (95% CI) ^c		

CI: confidence interval; CMH: Cochran-Mantel-Haenszel; COVID-19: coronavirus disease 2019; ERA: endothelin receptor antagonist; N/A: not applicable; NE: not evaluable; PAH: pulmonary arterial hypertension; PDE5i: phosphodiesterase 5 inhibitor; WHO: World Health Organization
Study: STELLAR (MK-7962-003)
Database Cutoff Date: 06DEC2022
Number of participants: full analysis set, participants who received double PAH background therapy (sotatercept: ERA+PDE5 inhibitor) or triple PAH background therapy (placebo: ERA+PDE5 inhibitor + prostacyclin (selexipag)).
Percentages are calculated as (n/m)*100. A missing result at week 24 not due to COVID-19 was considered a non-responder. Participants who missed assessments due to COVID-19 were taken out of the denominator.
^a Comparison with placebo uses CMH method stratified by WHO functional class (class II vs. III) and background PAH therapy (mono/double vs. triple therapy).
^b The relative risks were calculated by dividing the risk rate of being on-treatment with sotatercept and having a response of interest (worsening or improving) with the risk rate of being off-treatment (in this case meaning being treated with selexipag) and having a response of interest. For example, the RR for improvement of sotatercept versus selexipag is calculated as follows:

$$RR = \frac{P(R + | T+)}{P(R + | T-)}$$

Where P is the risk rate, R the response of interest, and T the treatment, with + and – indicating of having (+) or not having (-) the response of interest and being on (+) or off-treatment (-). Using the values in the table would give:

$$RR = \frac{1}{1}$$

^c Therefore, the RR for worsening of sotatercept versus selexipag is calculated as follows:

$$RR = \frac{1}{1}$$

Source: MSD 2024 [Confidential Data on File] (33)

2.9 Meta-analysis

No meta-analyses have been conducted, as no relevant meta-analyses are feasible. At the current time, data on the effect of treatment with sotatercept in relevant patients are available only from the STELLAR and PULSAR studies (as described in section 2.1), however the doses of sotatercept administered in these two studies are not mutually compatible (Table 63), and so meta-analysis of data from these two studies is not appropriate.

Table 63 Sotatercept doses used in the PULSAR and STELLAR studies

	PULSAR study	STELLAR study
Sotatercept dose(s) used	- Sotatercept 0.3 mg/kg subcutaneous injection once every 21 days plus SoC - Sotatercept 0.7 mg/kg subcutaneous injection once every 21 days plus SoC	Sotatercept starting dose 0.3 mg/kg with a target dose of 0.7 mg/kg subcutaneous injection once every 21 days plus SoC

SoC: standard of care

Additionally, the PULSAR and STELLAR trials were designed for different purposes. The PULSAR Phase 2 trial was designed to demonstrate safety and determine appropriate dosing, it was not powered to demonstrate efficacy for sotatercept as STELLAR was. This means that PULSAR not only had a different endpoint, it also had different statistical methods, and a different way of reporting results. Thus, pooling data from PULSAR and STELLAR would have been statistically inappropriate.

2.10 Indirect and mixed treatment comparisons

ITC of sotatercept versus selexipag (using trial data)

- The ITC was found to be feasible using the STELLAR trial to provide data on sotatercept and the GRIPHON and TRACE trials to provide data on selexipag.
- ITC conducted using the Bucher method found statistically significant results in favour of sotatercept vs. selexipag for the outcomes of WHO FC worsening, change in 6MWD and change in NT-proBNP.

ITC of sotatercept vs. selexipag (using trial and real-world evidence)

- A feasibility assessment was conducted for ITC of clinical effectiveness outcomes of treatment with sotatercept vs. selexipag, using individual patient data from STELLAR (for sotatercept) and real-world evidence (for selexipag). Specifically, the outcome of change in four-stratum risk status is of interest in order to inform the cost-effectiveness analysis.
- The Translational Research in Pulmonary Hypertension at Imperial College (TRIPHIC) research database was identified as a source of relevant individual data on patients with PAH treated with selexipag in the UK.
- Relevant data were available for too few patients from the TRIPHIC database to allow for a credible and meaningful indirect comparison.

ITC of sotatercept versus infused prostacyclin

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

- For patients with ESC/ERS four-strata intermediate-high risk or high risk PAH who would receive sotatercept in combination with PDE5i + ERA, the current SoC (and therefore relevant comparator) is infused prostacyclin (IV PCA) in combination with PDE5i + ERA, a comparison for which direct clinical evidence from the STELLAR study is unavailable.
- Accordingly, a feasibility assessment for ITC of clinical effectiveness outcomes of treatment with sotatercept versus infused prostacyclin, both as an add-on therapy to dual background therapy for PAH, has been conducted.
- Appropriate ITCs were not found to be feasible for this comparison using available data.

2.10.1 ITC of sotatercept vs. selexipag (using trial data)

The proposed positioning for initiating sotatercept (in combination with ERA + PDE5i) is in ESC/ERS intermediate-low risk patients as an alternative to selexipag (in combination with ERA + PDE5i; as described in Table 5 in section 1.3.3).

Therefore, a feasibility assessment for ITC of clinical effectiveness outcomes of treatment with sotatercept versus selexipag, both as an add-on therapy to dual background therapy for PAH, has been conducted.

ITC was found to be feasible using the STELLAR trial as the source of data for sotatercept and the GRIPHON and TRACE trials as the source of data for selexipag. The details of the feasibility assessment and the methodology of the analyses are provided in Appendix P, a brief summary of key details is provided below:

- Evidence base and feasibility assessment
 - The SLR of clinical effectiveness evidence (described in Appendix B) was used to identify potentially relevant studies for this ITC. The clinical effectiveness SLR identified four studies that reported on the clinical effectiveness of selexipag-based treatment regimens: the Phase 3 GRIPHON study (NCT01106014), Phase 4 TRACE study (NCT03078907), Phase 3 TRITON study (NCT02558231) and Phase 2 NCT00993408 study.
 - The TRITON study was conducted specifically for treatment-naïve patients, in contrast to the GRIPHON population, which included a mix of treatment-naïve (about 20%) and -experienced (about 80%) patients and the STELLAR population that included only treatment-experienced (heavily

pre-treated) patients, hence TRITON was excluded from further consideration due to patient heterogeneity

- The NCT00993408 study reported outcomes from baseline to week 17 only, whereas for outcomes like WHO FC, 6MWD and NT-proBNP, both GRIPHON and TRACE reported outcomes from baseline to weeks 24 or 26, which is more comparable to STELLAR (baseline to weeks 12, 24, 36, etc.). Given the comprehensive reporting of outcomes of interest in both GRIPHON and TRACE at or near the same time scales as in STELLAR, the inclusion of NCT00993408 in addition to GRIPHON and TRACE was less likely to lead to additional insights for this comparison. Thus, the NCT00993408 study was also excluded from further consideration.
- Consequently, only data from the STELLAR, GRIPHON and TRACE studies were included in the ITCs.
- Statistical methods
 - The Bucher method of ITC is simple and well-established. Given the similarities in study design between GRIPHON and TRACE, outcomes of interest reported by both trials using the same units and methodology (WHO FC improvement and worsening) were pooled. Pooling to inform the comparison of selexipag + background therapy vs. background therapy alone was performed via the inverse variance meta-analysis approach prior to the execution of Bucher ITC. For outcomes only published and available from GRIPHON or reported in incomparable units between GRIPHON and TRACE, data pooling or meta-analysis was not applicable.
- Outcomes of interest and comparative analyses
 - Table 64 presents the key outcomes of interest for the ITC. Because of limited data availability for time to event outcomes (available from GRIPHON only) and different methods used for calculating continuous outcomes (change in 6MWD and change in NT-proBNP), pooled analyses can only feasibly be conducted for the binary outcomes of WHO FC improvement and worsening because these are the only two outcomes for which data is available from both GRIPHON and TRACE that can be pooled.
 - In particular, GRIPHON reports change in 6MWD using a Hodges-Lehmann median, while TRACE reports change in 6MWD using a least squares mean. While it is possible to compare a median and mean

estimate, this requires an assumption that the distribution of the variable is normally distributed so that the median serves as a reasonable approximation of the mean. This assumption is not valid for 6MWD, as previous analyses have demonstrated 6MWD to be non-normally distributed.

- Additionally, GRIPHON reports change in NT-proBNP as a Hodges-Lehmann median difference, while TRACE reports change in NT-proBNP as a geometric least squares mean ratio.
- However, it is possible to compare the results for each selexipag trial individually with results from STELLAR for sotatercept using the Bucher ITC approach.

Table 64 Overview of outcomes of interest

No.	Outcome ^a	Outcome type	Relative treatment effect measure	Time point ^g
1.1	WHO FC improvement ^b	Binary	Relative risk; odds ratio ^e	- GRIPHON: from baseline to week 26 - TRACE: from baseline to week 24 - STELLAR: from baseline to week 24
1.2	WHO FC worsening ^b			
2	Change in 6MWD ^c	Continuous	Difference ^f	- GRIPHON: from baseline to week 26 - TRACE: from baseline to week 24 - STELLAR: from baseline to week 24
3	Change in NT-proBNP ^c			
4	Time to death or non-fatal clinical worsening ^d	Time to event	Hazard ratio ^e	End of trial follow-up with a median follow-up of: - GRIPHON: 98.1 weeks - STELLAR: 51.3 weeks
5.1	Time to death			
5.2	Time to PAH-related death			
6	Time to first PAH-related hospitalisation			

6MWD: 6-minute walking distance; FC: functional class; ITC: indirect treatment comparison; NT-proBNP: N-terminal pro-brain natriuretic peptide; PAH, pulmonary arterial hypertension; WHO: World Health Organization

^a Of note, the GRIPHON population included a very small proportion (1.8%) of patients with WHO FC I or IV at baseline, whereas the STELLAR population included only patients with FC II or III. Therefore, ideally, only the GRIPHON subpopulation with FC II or III (i.e., excluding patients with FC I or IV) would be used in the ITC whenever feasible to enable a comparative analysis.

^b For WHO FC outcomes, given the public reporting of FC improvement/worsening *stratified by baseline WHO FC* in the GRIPHON trial publication, the comparison of FC outcomes was conducted between the overall STELLAR population (with FC II or FC III) and the GRIPHON subpopulation with FC II or III only.

^c Both 6MWD and NT-proBNP outcomes were publicly reported for the overall population only in both GRIPHON and TRACE. (The GRIPHON population includes a very small proportion [1.8%] of patients with FC I or IV.) In the absence of publicly available data for the GRIPHON or TRACE subpopulations with FC II or III, the comparison of both outcomes was conducted between the STELLAR overall population (FC II or FC III) and the overall GRIPHON or TRACE population, acknowledging the small difference in the patient populations. However, this difference is expected to have minimal impact on the ITC results due to two reasons. First, patients with FC I or IV accounted for only 1.8% of the overall GRIPHON population and were excluded from TRACE, which is expected to have minimal impact on the outcomes. Second, even the minimal impact on the comparison between STELLAR and GRIPHON would likely be balanced, given the similar proportion of patients with FC I (approximately 0.8% of overall GRIPHON population) vs. FC IV (approximately 1.0% of overall GRIPHON population). Given the two reasons above, the small difference between the STELLAR and GRIPHON populations as to baseline WHO FC is expected to have very minimal impact on both the absolute treatment outcomes (for each treatment arm separately) and relative treatment effect (between selexipag + background therapy and background therapy alone).

^d In GRIPHON, the clinical worsening outcome is defined as a composite endpoint of death or a complication related to PAH, whichever occurred first, up to the end of the treatment period. Complications related to PAH were disease progression or worsening of PAH that resulted in hospitalisation, initiation of parenteral prostanoid therapy or long-term oxygen therapy, or the need for lung transplantation or balloon atrial septostomy as judged by the physician. (Placement on a transplant waiting list represented an acute measure, as confirmed by the critical-event committee, and an actual lung transplantation would also meet this criterion.) Disease progression was defined as a decrease from baseline of at least 15% in the 6MWD (confirmed by means of a second test on a different day) accompanied by a worsening in WHO FC (for the patients with WHO FC II or III at baseline) or the need for additional treatment for PAH (for the patients with WHO FC III or IV at baseline).

In STELLAR, the clinical worsening outcome is defined as a composite endpoint as a composite of death from any cause, worsening-related listing for lung or heart-lung transplantation, initiation of rescue therapy with an approved medication for PAH or increase in the prostacyclin dose by at least 10%, atrial septostomy, hospitalisation for worsening PAH (≥ 24 hours), or worsening of PAH relative to baseline as defined by both a worsened WHO FC and a decrease in 6MWD by $\geq 15\%$ [confirmed by two tests ≥ 4 hours but ≤ 1 week apart].

^e Under the Bucher framework, the *point estimate* of relative risk for the comparison of sotatercept + background therapy vs. selexipag + background therapy was calculated as: relative risk of sotatercept + background therapy vs. background therapy alone from STELLAR, divided by the relative risk of selexipag + background therapy vs. background therapy alone from GRIPHON.

The *variance* of relative risk for the comparison of sotatercept + background therapy vs. selexipag + background therapy was calculated as: the variance for the relative risk (on the log scale) of sotatercept + background therapy vs. background therapy alone from STELLAR, plus the variance for the relative risk (on the log scale) of selexipag + background therapy vs. background therapy from GRIPHON. These also apply to odds ratios and hazard ratios.

^f In GRIPHON, the relative treatment effect on continuous outcomes (e.g., changes in 6MWD and NT-proBNP) was measured through median difference from the Hodges-Lehmann estimator of location shift as per both trial protocols. In TRACE, the relative treatment effect on 6MWD was calculated as the least squares mean difference, and the relative treatment effect on NT-proBNP was calculated as the geometric least squares mean ratio. For the ITC purpose under the Bucher framework, the *point estimates* of mean difference for the comparison of sotatercept + background therapy vs. selexipag + background therapy was calculated as: median difference of sotatercept + background therapy vs. background therapy from STELLAR, minus the median difference of selexipag + background therapy vs. background therapy alone from GRIPHON.

The *variance* of mean difference for the comparison of sotatercept + background therapy vs. selexipag + background therapy was calculated as: the variance for the mean difference (on the original scale) of sotatercept + background therapy vs. background therapy alone from STELLAR, plus the variance for the mean difference (on the original scale) of selexipag + background therapy vs. background therapy alone from GRIPHON.

^g The small differences in the time points (week 24 for STELLAR and TRACE and week 26 from GRIPHON) were driven by the differences in study design, which was expected to have limited impact on the ITC results.

ITC was conducted using the Bucher method for the outcomes of WHO FC improvement, WHO FC worsening, change in 6MWD and change in NT-proBNP. The results of these comparisons are shown in Table 65. Results in favour of sotatercept where the point estimate and both ends of the 95% confidence intervals favour sotatercept were found in terms of each of these outcomes with the exception of WHO FC improvement where a significant difference between these two treatments was not found (see Table 65); however all comparisons numerically (i.e., in the point estimates) favoured sotatercept.

Table 65 Summary of key Bucher ITC results

Outcome	Source of selexipag data used in the comparison	ITC results (sotatercept vs. selexipag)	
Binary outcomes		Risk ratio (95% CI)	Odds ratio (95% CI)
WHO FC improvement	GRIPHON		
	TRACE		
	Pooled ^a (GRIPHON + TRACE)		
WHO FC worsening	GRIPHON		
	TRACE		
	Pooled ^a (GRIPHON + TRACE)		
Continuous outcomes		Difference ^b (95% CI)	
6MWD change in metres	GRIPHON		
	TRACE		
NT-proBNP change in pg/mL	GRIPHON		
	TRACE		
Time to event outcomes		Hazard ratio (95% CI)	
Time to death or non-fatal clinical worsening	GRIPHON		
Time to death	GRIPHON		
Time to PAH-related death	GRIPHON		
Time to first PAH hospitalisation	GRIPHON		

6MWD: 6-minute walking distance; CI: confidence interval; FC: functional class; ITC: indirect treatment comparison; NT-proBNP: N-terminal pro-brain natriuretic peptide; PAH: pulmonary arterial hypertension; WHO: World Health Organization

^a For Bucher ITCs, pooling of GRIPHON+TRACE was conducted via the inverse variance meta-analysis approach before execution of the ITC. Fixed effects results are reported for ITC analyses using pooled (GRIPHON+TRACE) data.

^b ITC results of continuous outcomes are presented in terms of difference in Hodges-Lehmann estimates of location shift.

2.10.2 ITC of sotatercept vs. selexipag (using trial and real-world evidence)

In order to corroborate the findings of the non-randomised within-trial analysis discussed earlier in section 2.8, a feasibility assessment was conducted for ITC of clinical effectiveness outcomes of treatment with sotatercept vs. selexipag, using individual patient data from STELLAR (on sotatercept) and real-world evidence (on selexipag). Specifically, the outcome of change in four-stratum risk status is of interest in order to inform the cost-effectiveness analysis, as well as it being a clinically relevant endpoint which cannot be informed by the trial-based ITC (as described in the earlier subsection and in Appendix P).

The TRIPHIC research database was identified as a source of relevant individual data on patients with PAH treated with selexipag in the UK. In order to perform a meaningful indirect comparison, data need to be available from the TRIPHIC database in order to:

- Calculate the COMPERA 2.0 four-stratum risk status (i.e., 6MWD, WHO FC and BNP) at relevant timepoints during selexipag treatment (i.e., start of treatment and approximately 24 weeks from start of treatment).
- Balance potentially confounding variables (e.g., PVR).

Relevant data were available for too few patients from the TRIPHIC research database to allow for a credible and meaningful indirect comparison to be feasible. Full details of this feasibility assessment is provided in Appendix O.

2.10.3 ITC of sotatercept vs. infused prostacyclin

As described in section 1.3.3, for patients with ESC/ERS four-strata intermediate-high risk or high risk PAH who would receive sotatercept in combination with PDE5i + ERA, the current SoC and therefore relevant comparator is IV PCA in combination with PDE5i + ERA. However, direct clinical evidence to support such a comparison in this population is unavailable as the STELLAR study did not make this comparison (as described in Table 5 in section 1.3.3).

Accordingly, a feasibility assessment for ITC of clinical effectiveness outcomes of treatment with sotatercept versus IV PCA, both as an add-on therapy to dual therapy for PAH, has been conducted to understand the viability of evidence to support this comparison from published clinical evidence external to the STELLAR study.

It was found that anchored ITCs would not be feasible or appropriate due to the lack of appropriate common comparator from identified studies; appropriate unanchored indirect comparisons and matching-adjusted indirect comparison are also not feasible (details provided in Appendix Q).

2.10.4 Uncertainties in the indirect and mixed treatment comparisons

The ITCs performed for the comparison of sotatercept versus selexipag (described in greater detail in Appendix P) are associated with limitations that should be considered when interpreting the results. The key limitations can be summarised in the following points:

- Data on selexipag + background therapy are available only from trial publications which do not consistently report detailed subgroup results for relevant outcomes. Therefore the ITC necessarily assumed that relative treatment effect is consistent across subgroups (though this assumption is supported by the findings of consistent relative treatment effect of selexipag + background therapy vs. background therapy alone on outcomes for which results across pre-specified subgroups are available, including the time to death/clinical worsening outcome and 6MWD).
- Sample sizes were limited for the ITCs in some outcomes, leading to wide CIs for the analyses results.
- As is common for ITCs, certain differences in study design (e.g., year of study, geographic location) could not be adjusted for. Therefore, the ITC necessarily assumed that relative treatment effect is not affected by these.

However, given (1) that STELLAR, GRIPHON and TRACE represent the best available clinical evidence for sotatercept and selexipag, respectively, and (2) an anchored ITC is feasible (placebo in combination with background therapy is a common comparator in all trials included in the analysis) and has been conducted, the evidence used to inform this ITC is robust, despite the lack of a randomised trial directly comparing sotatercept with selexipag.

2.11 Adverse reactions

Summary of adverse events information

STELLAR

- Sotatercept was demonstrated to be generally well tolerated in the STELLAR trial (29).
- The rate of serious adverse events experienced by participants taking sotatercept was not significantly greater compared to that experienced by participants taking placebo (29).
- High treatment adherence (>90%) was observed in the sotatercept arm of the STELLAR trial due to favourable tolerability (29).

SOTERIA

- SOTERIA interim results support the manageable safety of sotatercept for the treatment of PAH (6).

2.11.1 STELLAR study - summary of adverse events

The overall incidences of treatment-emergent adverse events (TEAEs) were similar in both groups (29). Incidences of TEAEs were higher in the sotatercept group than in the placebo group in the summary categories of TEAEs considered related to treatment by the investigator, TEAEs of interest, and the TEAE of special interest (29). Severe TEAEs, serious TEAEs, serious TEAEs leading to treatment discontinuation, serious TEAEs leading to death, TEAEs leading to treatment discontinuation and TEAEs leading to death occurred less commonly in the sotatercept group than in the placebo group (Table 66) (29).

Detailed AEs data from the STELLAR study are provided in Appendix D (29):

- The incidences of specific TEAEs were generally similar in both groups.
 - Frequently reported TEAEs (incidence $\geq 5\%$) with a higher incidence in the sotatercept group than in the placebo group were epistaxis, telangiectasia, dizziness, nasal congestion, thrombocytopenia and increased haemoglobin.
 - Less frequently reported TEAEs (incidence $< 5\%$) with a higher incidence in the sotatercept group than in the placebo group were abdominal pain upper, chills, gingival bleeding, injection site erythema, bursitis, polycythemia, cataract and blurred vision.
- The incidences of TEAEs leading to dose modification were generally similar in both groups.

- TEAEs leading to dose modification that occurred at a higher incidence in the sotatercept group than in the placebo group were increased haemoglobin, epistaxis and telangiectasia; these events occurred only in the sotatercept group.

Table 66 Overall summary of treatment-emergent adverse events - Safety Set - Cumulative results (DBPC+LTDB)

Number of participants with any	Sotatercept (N=163) n (%)	Placebo (N=160) n (%)	All patients (N=323) n (%)	Point estimate	
				(%) ^a	95% CI ^a
TEAEs	151 (92.6)	149 (93.1)	300 (92.9)	-0.5	(-6.4, 5.4)
TEAEs related to treatment	83 (50.9)	45 (28.1)	128 (39.6)	22.8	(12.2, 32.9)
TEAE of interest	97 (59.5)	71 (44.4)	168 (52.0)	15.1	(4.2, 25.7)
TEAE of special interest	27 (16.6)	7 (4.4)	34 (10.5)	12.2	(5.8, 19.2)
TEAEs leading to treatment discontinuation	6 (3.7)	11 (6.9)	17 (5.3)	-3.2	(-8.7, 1.8)
TEAEs leading to study discontinuation	6 (3.7)	9 (5.6)	15 (4.6)	-1.9	(-7.1, 2.9)
TEAEs leading to death	2 (1.2)	7 (4.4)	9 (2.8)	-3.1	(-7.7, 0.5)
Severe TEAEs	24 (14.7)	33 (20.6)	57 (17.6)	-5.9	(-14.3, 2.5)
Serious TEAEs	40 (24.5)	47 (29.4)	87 (26.9)	-4.8	(-14.5, 4.9)
Serious TEAEs related to treatment	3 (1.8)	2 (1.3)	5 (1.5)	0.6	NR
Serious TEAEs leading to treatment discontinuation	4 (2.5)	9 (5.6)	13 (4.0)	-3.2	(-8.2, 1.3)
Serious TEAEs leading to study discontinuation	4 (2.5)	8 (5.0)	12 (3.7)	-2.5	(-7.4, 1.8)
Serious TEAEs leading to death	2 (1.2)	7 (4.4)	9 (2.8)	-3.1	(-7.7, 0.5)

AE: adverse event; AEOL: adverse events of interest; AESI: adverse event of special interest; CI: confidence interval; DBPC: double-blind placebo-controlled period; FSH: follicle-stimulating hormone; LTDB: long-term double-blind period; NR: not reported (incidence <4 in both treatment groups); TEAE: treatment-emergent adverse event
Database Cutoff Date: 06-DEC-2022

For the sotatercept vs. placebo comparison, the 95% CI was not provided when the incidence is <4 in both treatment groups.

A treatment-emergent adverse event is an event that has a start date on or after the first dose of treatment and up to 8 weeks after the last dose of treatment.

AESIs include telangiectasia.

AEOLs include the following: Increased haemoglobin, immunogenicity, thrombocytopenia, neutropenia (including febrile neutropenia), embryo-foetal toxicity, hepatic toxicity, cardiac events, thrombotic events, renal toxicity, bleeding events, increased blood pressure and suppression of FSH.

^a Point estimates and confidence intervals for between treatment group differences are based on Miettinen and Nurminen method.

Adverse events were assessed by the investigator to be 'suspected related' or not 'suspected related'. A 'suspected related' AE is considered to be related to treatment

Source: MSD 2024 STELLAR CSR (29).

2.11.2 SOTERIA study - summary of adverse events

Safety data on long-term treatment with sotatercept is available from the 08-NOV-2023 data cutoff date of the study (922.2 patient-years of exposure) (6). The data show that sotatercept continued to be generally well tolerated, with less than 5% of participants with TEAEs leading to treatment discontinuation or death (Table 67) (6).

With regard to specific AEs of special interest (Table 68) (6):

- Telangiectasia events: No participants with serious events; only 1 event led to discontinuation.
- Thrombocytopenia events: <1% of participants with serious events.
- Bleeding events: Most were non-serious epistaxis events. 5.2% of participants had serious events and <1% led to treatment discontinuation.

Table 67 SOTERIA study overall safety and tolerability results

	Placebo-crossed (N=143)	Continued sotatercept (N=259)	Blinded (N=24)	All participants (N=426)
Median (range) duration of cumulative exposure to sotatercept, days	389.0 (21-725)	798.0 (297-1972)	129.5 (21-399)	652.5 (21-1972)
Total exposure in patient-years	159.7	751.7	10.9	922.2
Any TEAE, n (%)	131 (91.6)	240 (92.7)	16 (66.7)	387 (90.8)
TEAE related to treatment	79 (55.2)	126 (48.6)	5 (20.8)	210 (49.3)
TEAE leading to treatment discontinuation	2 (1.4)	12 (4.6)	1 (4.2)	15 (3.5)
Serious TEAE, n (%)	41 (28.7)	80 (30.9)	8 (33.3)	129 (30.3)
Serious TEAE related to treatment	5 (3.5)	6 (2.3)	0 (0)	11 (2.6)
Serious TEAE leading to treatment discontinuation	1 (0.7)	9 (3.5)	1 (4.2)	11 (2.6)
Serious TEAE leading to death	4 (2.8)	6 (2.3)	2 (8.3)	12 (2.8)

TEAE: treatment-emergent adverse event

Data cutoff date: 08-NOV-2023

Source: Preston et al. 2025 (6)

Table 68 SOTERIA study summary of adverse events of special interest

No. (%) of participants		Placebo-crossed (N=143)	Continued sotatercept (N=259)	Blinded (N=24)	All participants (N=426)
Bleeding					
Bleeding events		59 (41.3)	89 (34.4)	5 (20.8)	153 (35.9)
- Related to treatment		32 (22.4)	38 (14.7)	1 (4.2)	71 (16.7)
- Leading to treatment discontinuation		0 (0.0)	2 (0.8)	0 (0.0)	2 (0.5)
- Epistaxis (most common bleeding event)		47 (32.9)	45 (17.4)	2 (8.3)	94 (22.1)
Serious bleeding events		5 (3.5)	15 (5.8)	2 (8.3)	22 (5.2)
- Related to treatment		1 (0.7)	3 (1.2)	0 (0.0)	4 (0.9)
- Leading to treatment discontinuation		0 (0.0)	2 (0.8)	0 (0.0)	2 (0.5)
- Serious bleeding events by preferred terms	Gastrointestinal haemorrhage	1 (0.7)	5 (1.9)	2 (8.3)	8 (1.9)

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

No. (%) of participants		Placebo-crossed (N=143)	Continued sotatercept (N=259)	Blinded (N=24)	All participants (N=426)
	Haemoptysis	1 (0.7)	4 (1.5)	0 (0.0)	5 (1.2)
	Heavy menstrual bleeding	0 (0.0)	2 (0.8)	0 (0.0)	2 (0.5)
Increased haemoglobin and thrombocytopenia					
Increased haemoglobin		23 (16.1)	38 (14.7)	0 (0.0)	61 (14.3)
- Related to treatment		21 (14.7)	34 (13.1)	0 (0.0)	55 (12.9)
- Increased haemoglobin by preferred terms	Haemoglobin increased	18 (12.6)	23 (8.9)	0 (0.0)	41 (9.6)
	Polycythaemia	5 (3.5)	14 (5.4)	0 (0.0)	19 (4.5)
	Haematocrit increased	2 (1.4)	4 (1.5)	0 (0.0)	6 (1.4)
Serious increased haemoglobin		0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Thrombocytopenia		13 (9.1)	13 (5.0)	0 (0.0)	26 (6.1)
- Related to treatment		12 (8.4)	10 (3.9)	0 (0.0)	22 (5.2)
- Thrombocytopenia by preferred terms	Thrombocytopenia	13 (9.1)	10 (3.9)	0 (0.0)	23 (5.4)
	Platelet count decreased	0 (0.0)	3 (1.2)	0 (0.0)	3 (0.7)
Serious thrombocytopenia		2 (1.4)	1 (0.4)	0 (0.0)	3 (0.7)
- Related to treatment		2 (1.4)	1 (0.4)	0 (0.0)	3 (0.7)
- Serious thrombocytopenia by preferred term	Thrombocytopenia	2 (1.4)	1 (0.4)	0 (0.0)	3 (0.7)
Telangiectasia					
Telangiectasia (on or after rollover)		37 (25.9)	34 (13.1)	3 (12.5)	74 (17.4)
- Related to treatment		37 (25.9)	32 (12.4)	2 (8.3)	71 (16.7)
- Leading to discontinuation		0 (0.0)	1 (0.4)	0 (0.0)	1 (0.4)
Serious telangiectasia		0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)

Data cutoff date: 08-NOV-2023

Source: Preston et al. 2025 (6)

2.12 Ongoing studies

The ongoing SOTERIA long-term follow-up study and ZENITH study (described earlier in section 2.3.2) may provide additional evidence relevant to this appraisal in the next 12 months.

The ZENITH trial was stopped early based on the strength of its results as observed at the interim analysis of its primary endpoint. An independent data monitoring committee recommended ZENITH be stopped early and all participants be offered the opportunity to receive sotatercept through the SOTERIA open-label extension study. Data from the ZENITH study before it was stopped is expected to become available during the course of this appraisal.

The HYPERION trial (also described in section 2.3.2) lost clinical equipoise and was stopped early. The decision to stop the HYPERION study prior to its scheduled completion was based on the positive results from the interim analysis of the ZENITH trial and a review of the totality of data from the sotatercept clinical programme to date. Participants in the HYPERION trial will have the opportunity to receive sotatercept as part of the SOTERIA study. Data from the HYPERION study before it was stopped may become available in the next 12 months.

2.13 Interpretation of clinical effectiveness and safety evidence

2.13.1 Clinical evidence from the STELLAR study

STELLAR provided a robust body of evidence to support the efficacy and impact of sotatercept on overall health of participants presenting with PAH. Sotatercept has shown statistically significant and clinically meaningful improvements across well-established and clinically validated cardiovascular endpoints (39), including:

- 6MWD: patients treated with sotatercept achieved clinically meaningful and durable improvements in 6MWD (40.8m from baseline, $p < 0.001$) vs. patients treated with placebo in combination with background therapy (39).
- Time to death or clinical worsening: sotatercept reduced their risk of death or clinical worsening by 84% versus patients treated with placebo in combination with background therapy.
- Risk status: more than twice as many patients receiving sotatercept achieved or maintained low risk status vs. patients receiving placebo in combination with background therapy (39.5% vs. 18.2%, $p < 0.001$) with long-term studies demonstrating the durability of risk status maintenance and improvements.

- Functional class: more than twice as many patients receiving sotatercept improved their FC versus patients receiving placebo in combination with background therapy (29.4% vs. 13.8%, $p < 0.001$) with long-term studies demonstrating the durability of FC maintenance and improvements.
- Haemodynamics: patients receiving sotatercept demonstrated improved haemodynamics (e.g., PVR and mPAP) compared to patients receiving placebo in combination with background therapy, consistent with improved prognosis and decreased disease severity.
- PAH-SYMPACT[®]: sotatercept is the first and only therapy for PAH to demonstrate statistically significant improvements in the Physical Impacts and Cardiopulmonary Symptoms components of the PAH-SYMPACT[®] in patients already on background therapy.
- Subgroup analysis: sotatercept demonstrated consistent improvement in key endpoints regardless of patient characteristics at baseline (e.g., risk status, background therapy).

2.13.2 Clinical evidence versus selexipag from the within-trial analysis of the STELLAR study

In the absence of randomised evidence, the relative efficacy of sotatercept in intermediate-low risk patients versus selexipag for the critical outcome of risk status change can be informed via a post-hoc within-trial analysis of the STELLAR study, including only the patients treated with these regimens in each arm.

The findings of such an analysis are informative despite this analysis breaking randomisation and not being fully statistically powered. The results of the analysis indicated that treatment with the sotatercept-based regimen is more likely to lead to improvement in risk status and less likely to lead to worsening in risk status.

2.13.3 Point of initiation of treatment with sotatercept in the STELLAR study and expected in UK clinical practice

The STELLAR trial recruited patients on stable background therapy (to which sotatercept was then added), rather than patients who had failed to achieve an adequate response to background therapy (to which sotatercept was then added, which would be more aligned with how treatment escalation is recommended in the published clinical commissioning policies and ESC/ERS guidelines). Hence the point of initiation of treatment with sotatercept in the STELLAR trial may not align with how sotatercept will be used in UK clinical practice. While this misalignment cannot be reliably adjusted for and is a source of uncertainty, it is likely that

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

the findings of the STELLAR study are representative of what would be observed in patients initiating treatment with sotatercept after failing to achieve an adequate response (i.e., low risk status) with dual combination therapy. UK clinical expert opinion has been sought and aligns with this. The generally consistent clinical improvements observed with treatment using sotatercept throughout assessed subgroups (as described in section 2.8.1) indicate this assumption to be valid. Furthermore, the STELLAR results may comprise a conservative estimate of sotatercept's benefit in practice, given that patients who are clinically worse may have more potential to improve with a new treatment.

2.13.4 Indirect treatment comparisons vs. selexipag

Indirect evidence and post-hoc analyses were required to enable a comparison of sotatercept vs. selexipag. Such analyses are inevitably associated with limitations that are common with the use of such methods (as described in section 2.10.4). However, the Bucher ITC results have a usable degree of robustness in cases such as this where fully powered and randomised direct clinical trial evidence is not available. The analyses show a benefit of sotatercept versus selexipag as additions on top of background therapy. These analyses indicate that treatment with the sotatercept-based regimen is more likely to lead to improvement in WHO FC and less likely to lead to worsening in WHO FC.

3 Cost effectiveness

Summary of key cost-effectiveness information

Objective:

- The cost-effectiveness analysis reported here details the comparison of initiating sotatercept in addition to dual background therapy (PDE5i+ERA) vs. selexipag in addition to PDE5i+ERA for the treatment of patients with intermediate-low risk PAH, according to the ESC/ERS treatment guidelines.

Model structure:

- An SLR of economic evaluations identified previous cost-effectiveness analyses with model structures that predominantly classified health states according to WHO FC. However, given extensive clinical expert feedback, a de novo risk status-based model structure was developed to more appropriately reflect UK clinical practice.
- A Markov model was developed, with mutually exclusive health states: low risk, intermediate-low risk, intermediate-high risk and high risk as defined by guidelines, in addition to post-transplant and death. Off-treatment health states are used to reflect treatment continuation criteria.

Model inputs:

Patient population inputs:

- In line with the decision problem outlined in section 1.1, the expected population who will initiate sotatercept, when combined with PDE5i + ERA, consists of intermediate-low risk patients who would otherwise be treated with selexipag in combination with PDE5i + ERA.

Clinical efficacy inputs:

- Sotatercept efficacy is informed directly by observed STELLAR trial data; selexipag efficacy is informed using a within-trial analysis of STELLAR data based on a subset of the STELLAR trial population.
- Continuation criteria are defined based on clinical advice. For sotatercept, a proportion of patients are expected to discontinue after 24 weeks of treatment due to inadequate response, and thereafter upon entry to the high risk health state and switching to the current SoC (parenteral PCA). For selexipag, patients discontinue upon entry to a worse health state than intermediate-low risk, where they are escalated to parenteral PCA treatment.

- A conservative assumption that risk status improvement with sotatercept is not possible beyond 24 weeks is made. Based on clinical advice, risk status improvement with selexipag is also not considered possible beyond this timepoint.
- Risk status-specific mortality curves are used, based on published real-world evidence, as life expectancy is differentiated according to a patient's risk status.

Utility inputs:

- The base case analysis uses EQ-5D-5L utility data reported in the STELLAR clinical trial to calculate utility by health state (i.e., risk status).
- An SLR identified several studies that reported utility values according to WHO FC, however no appropriate studies reporting risk status-based values were identified. The STELLAR trial is the most appropriate source of health state utility values.
- A disutility for IV administration (reported in the literature) was applied in the health states where patients receive IV PCA, to reflect the poor QoL associated with receiving those treatments, according to patient and clinician opinion.
- The negative impact on carer QoL is also accounted for, using a value from the literature.

Costs and resource use inputs:

- Drug costs were sourced from appropriate UK sources, and the considerable resource use associated with current PAH treatments was informed by published data and discussions with UK clinicians and nurses.
- Sotatercept monitoring costs are in line with the SmPC requirements.

Base case results and sensitivity analyses:

- Sotatercept is a cost-effective use of NHS resources, indicated by a base case incremental cost-effectiveness ratio (ICER) which falls below the range NICE typically considers good value for money, compared to current SoC. This incorporates qualification for a severity modifier of 1.2.
- Deterministic and probabilistic sensitivity analyses, varying the key model parameters over plausible ranges, indicate the base case findings to be robust.

Scenario analyses:

- A range of scenario analyses were conducted, to investigate alternative assumptions for the key model inputs. The results of these analyses demonstrate that sotatercept remains cost-effective under a range of alternative assumptions.

Cost-effectiveness conclusions:

- Using a model structure which reflects current clinical practice, and leveraging inputs from the pivotal clinical trial, real-world evidence and UK clinician input where appropriate, sotatercept is shown to be a cost-effective treatment in addition to dual background therapy when initiated in intermediate-low risk PAH. The inevitable uncertainty associated with a rare disease is mitigated where possible, however greater acceptance of this uncertainty should be considered given the substantial unmet need in this disease.

3.1 Published cost-effectiveness studies

An SLR was conducted to identify cost-effectiveness studies published since 2012 that are applicable to the decision problem in patients with PAH. The date limit of 2012 was chosen to ensure capture of the most recent relevant evidence, and studies published prior to this date are deemed to no longer reflect current treatment practices. In addition, relevant conference proceedings (2020 onward) and health technology assessment (HTA) websites (no date limit) were searched. The specific methods used to identify and select studies are described in Appendix E.

The SLR identified 33 publications on cost-effectiveness analyses, of which nine were reported in full-text publications (Table 69) and 25 were related to previous HTA submissions (Table 70; AWMSG [n=4], CADTH [n=8], NCPE [n=2], PBS [n=4], SMC [n=7]). There has been no prior appraisal of a technology for PAH by NICE.

None of the studies reported in full-text format had a UK setting, but most were potentially applicable to England and Wales based on the NICE quality assessment tool for economic evaluations (40). Therapies evaluated in the studies included PCA (n=1) (41), prostacyclin receptor agonists (PRAs) (n=1) (42), PDE5is (n=3) (43-45); ERAs (n=2) (46, 47) and combination therapies (n=2) (48, 49). Seven studies utilised a Markov model with varying numbers of health states, ranging from four to seven states, with the most common being five health states (WHO FC I, II, III, IV and death) (41, 43-47, 49). One study combined a decision tree and Markov model design (47). The remaining two studies used a discrete event simulation and an individual patient simulation design, respectively (42, 48). The majority of the economic evaluations adopted a payer perspective and included a modelled lifetime horizon of 30 years, with a common cycle length of three months.

In addition to the studies included in the SLR, an economic analysis conducted by the Institute for Clinical and Economic Review (ICER) in the United States of America (USA) was identified. In this analysis, ICER developed and published a decision analytic model for assessing the

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

cost-effectiveness of sotatercept in PAH patients from a healthcare sector perspective using a lifetime horizon (Table 69) (50). The model focuses on a hypothetical cohort of patients receiving either sotatercept plus background therapy or background therapy alone and employs an intention-to-treat approach, starting with adults in WHO FC II or III who are already on background therapy (50). The model cohort's health could either improve, remain stable, or deteriorate over time. Improvement was possible in the first 24 weeks, but beyond this, health could only remain the same or worsen. Reaching WHO FC IV triggered the addition of an infused prostacyclin treatment, which provided a chance for functional class improvement in the initial cycle. However, if a patient did not show improvement within one-cycle or returned to WHO FC IV after initial progress, sotatercept treatment was discontinued (50).

A total of 25 HTA submissions for PAH therapies were identified in the SLR: two HTA submissions for PCAs (51, 52); five for PRAs (53-56); seven for PDE5is (57-63); three for guanylate cyclase stimulators (64-66); seven for ERAs (67-73) and one for a combination therapy (74). Seven HTA submissions employed a Markov model for their analyses, four of which utilised a micro-simulation Markov model (all selexipag submissions). However, it is worth noting that patient-level micro-simulation models have received criticism for their complexity, which can potentially limit the range of feasible analyses, especially in terms of probabilistic analyses (55).

Findings from these studies were utilised to ensure that previous criticisms of modelling approaches and input utilisation were addressed to the extent possible. However, no economic assessment of sotatercept in the specific target population (i.e., patients with intermediate-low risk of mortality) was identified. Furthermore, none of the literature identified included an ESC/ERS risk-based definition of health states, which as discussed previously, aligns more closely with current UK clinical practice in terms of patient categorisation, treatment pathway and clinical goal of treatment (i.e., achieving a low risk status). This further limits the utility of these studies for the analysis. Consequently, a *de novo* cost-effectiveness analysis was conducted.

Table 69 Summary list of cost-effectiveness studies and model designs

Study	Summary of model	Patient population	QALYs	Total costs	ICER (base case only)
Studies identified in the SLR					
Coyle, 2016(43)	Country: Canada Perspective: Payer Treatment: Sildenafil Comparator: Tadalafil, ambrisentan, riociguat, bosentan and standard care Model type: Markov model Health states: 7: NYHA I, II, III and IV with oral therapy alone, and NYHA III and IV with oral therapy in combination with epoprostenol, and death Time horizon: Lifetime (30 years) Cycle length: 3 months ½ cycle correction: NR Discounting: 5%	PAH WHO FC II	Sildenafil: 4.63 Tadalafil: 4.002 Supportive care: 3.218 Ambrisentan 5mg: 4.636 Ambrisentan: 10mg: 4.217 Riociguat: 4.244 Bosentan: 3.904	Sildenafil: CA\$146,254 Tadalafil: CA\$153,245 Supportive care: CA\$155,156 Ambrisentan 5mg: CA\$377,187 Ambrisentan: 10mg CA\$377,523 Riociguat: CA\$388,491 Bosentan: CA\$406,282	Sildenafil dominated all comparators
		PAH WHO FC III	Sildenafil: 3.284 Tadalafil: 3.013 Supportive care: 2.687 Ambrisentan 5mg: 3.180 Ambrisentan: 10mg: 3.043 Riociguat: 3.045 Bosentan: 2.960	Sildenafil: CA\$181,119 Tadalafil: CA\$200,584 Supportive care: CA\$204,285 Ambrisentan 5mg: CA\$351,573 Ambrisentan: 10mg CA\$376,884 Riociguat: CA\$383,582 Bosentan: CA\$412,979	Sildenafil dominated all comparators
Dong, 2023(49)	Country: China Perspective: Payer Treatment: Macitentan and tadalafil Comparator: Macitentan, tadalafil and selexipag Model type: Markov model Health states: 5: WHO FC I, II, III, IV and death state	PAH	Macitentan + tadalafil: 6.666 Macitentan + tadalafil + selexipag: 7.234	Macitentan + tadalafil: \$116,534.543 Macitentan + tadalafil + selexipag: \$357,807.588	Selexipag + macitentan + tadalafil / Macitentan + tadalafil: 424,746 USD/QALY

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Study	Summary of model	Patient population	QALYs	Total costs	ICER (base case only)
	Time horizon: Lifetime (30 years) Cycle length: 26 weeks ½ cycle correction: NR Discounting: 5%				
Ekhlesi, 2023(46)	Country: Iran Perspective: Payer Treatment: Macitentan Comparator: Bosentan Model type: Markov model Health states: 5: WHO FC I, II, III, IV and death state Time horizon: Lifetime Cycle length: 3 months ½ cycle correction: NR Discounting: 7.5%	PAH	Macitentan: 4.02 Bosentan: 3.04	Macitentan: \$19,429.36 Bosentan: \$17,245.99	Macitentan / Bosentan: 2,233.46 USD/QALY
Lilyasari, 2019(44)	Country: Indonesia Perspective: Societal Treatment: Sildenafil (for WHO FC II and WHO FC III) Comparator: Beraprost (for WHO FC II and WHO FC III) Model type: Markov model Health states: 5: WHO FC I, II, III, IV and death state Time horizon: NR Cycle length: NR ½ Cycle correction: NR Discounting: 3%	PAH WHO FC II	Beraprost: 11.90 Sildenafil: 12.47	Beraprost: \$35,863 Sildenafil: \$37,632	Sildenafil / Beraprost: 3,098 USD/QALY
		PAH WHO FC III	Beraprost: 10.08 Sildenafil: 11.58	Beraprost: \$30,807 Sildenafil: \$35,059	Sildenafil / Beraprost: 2,827 USD/QALY

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Study	Summary of model	Patient population	QALYs	Total costs	ICER (base case only)
Nosrati, 2020(47)	Country: Iran Perspective: Payer Treatment: Macitentan Comparator: Bosentan Model type: Hybrid model- decision tree and Markov model Health states: 3 (decision tree): Macitentan vs. bosentan, further splitting into AEs vs. no AEs 5 (Markov): WHO FC I, II, III, IV and death state Time horizon: Lifetime Cycle length: 3 months ½ cycle correction: NR Discounting: 3% for utility and 5% for costs	PAH	Macitentan: 7.29 Bosentan: 6.48	Macitentan: \$13,876 Bosentan: \$14,163	Macitentan / Bosentan: - 357 USD/QALY
Roman, 2012(41)	Country: Spain Perspective: Payer Treatment: Iloprost Comparator: Epoprostenol and treprostinil Model type: Markov model Health states: 5: WHO FC I, II, III, IV and death state Time horizon: 3-year Cycle length: 3 months ½ Cycle correction: NR	PAH	Iloprost: 1.737 Treprostinil: 1.728 Epoprostenol: 1.780	Iloprost: €132,840 Treprostinil: €359,869 Epoprostenol: €429,775	Iloprost: Reference Treprostinil / Iloprost: treprostinil dominated Epoprostenol / Iloprost: 6,847,284 EUR/ QALY

Study	Summary of model	Patient population	QALYs	Total costs	ICER (base case only)
	Discounting: 3%				
Thongsri, 2016(45)	Country: Thailand Perspective: Societal Treatment: First-line standard treatment followed by second-line standard treatment Comparator: various first-line/second-line combinations of beraprost, sildenafil with iloprost, sildenafil with bosentan, and standard treatment Model type: Markov model Health states: 5: WHO FC I, II, III, IV and death state Time horizon: Lifetime Cycle length: 3-month ½ cycle correction: NR Discounting: 3%	PAH patients, age ≤30, FC II	Standard treatment / Standard treatment: 4.98 Beraprost / Standard treatment: 12.17 Sildenafil / Standard treatment: 12.28 Sildenafil / Sildenafil plus iloprost: 14.15 Sildenafil / Sildenafil plus bosentan: 14.27	Standard treatment / Standard treatment: ฿408,584 Beraprost / Standard treatment: ฿1,794,440 Sildenafil / Standard treatment: ฿2,231,411 Sildenafil / Sildenafil plus iloprost: ฿4,925,951 Sildenafil / Sildenafil plus bosentan: ฿3,836,462	Standard treatment / Standard treatment: Reference Beraprost / Standard treatment: 192,752 THB /QALY Sildenafil / Standard treatment: 249,770 THB/QALY Sildenafil / Sildenafil plus iloprost: 1,440,409 THB/QALY Sildenafil / Sildenafil plus bosentan: 805,528 THB/QALY
		PAH patients, age ≤30, FC III	Standard treatment / Standard treatment: 3.35 Beraprost / Standard treatment: 10.65 Sildenafil / Standard treatment: 13.33 Sildenafil / Sildenafil plus iloprost: 13.47 Sildenafil / Sildenafil plus bosentan: 13.44	Standard treatment / Standard treatment: ฿336,305 Beraprost / Standard treatment: ฿1,806,999 Sildenafil / Standard treatment: ฿2,600,134 Sildenafil / Sildenafil plus iloprost: ฿3,076,153 Sildenafil / Sildenafil plus bosentan: ฿2,844,343	Standard treatment / Standard treatment: Reference Beraprost / Standard treatment: 201,308 THB/QALY Sildenafil / Standard

Study	Summary of model	Patient population	QALYs	Total costs	ICER (base case only)
					treatment: 226,802 THB/QALY Sildenafil / Sildenafil plus iloprost: 3,298,720 THB/QALY Sildenafil / Sildenafil plus bosentan: 2,147,137 THB/QALY
		PAH patients, age >30, FC II	Standard treatment / Standard treatment: 4.63 Beraprost / Standard treatment: 9.16 Sildenafil / Standard treatment: 9.35 Sildenafil / Sildenafil plus iloprost: 10.39 Sildenafil / Sildenafil plus bosentan: 10.47	Standard treatment / Standard treatment: ₪375,850 Beraprost / Standard treatment: ₪1,207,880 Sildenafil / Standard treatment: ₪1,456,502 Sildenafil / Sildenafil plus iloprost: ₪3,081,920 Sildenafil / Sildenafil plus bosentan: ₪2,374,242	Standard treatment / Standard treatment: Reference Beraprost / Standard treatment: 183,813 THB/QALY Sildenafil / Standard treatment: 228,531 THB/QALY Sildenafil / Sildenafil plus iloprost: 1,567,582 THB/QALY Sildenafil / Sildenafil plus bosentan: 827,283 THB/QALY

Study	Summary of model	Patient population	QALYs	Total costs	ICER (base case only)
		PAH patients age >30, FC III	Standard treatment / Standard treatment: 3.15 Beraprost / Standard treatment: 8.00 Sildenafil / Standard treatment: 9.74 Sildenafil / Sildenafil plus iloprost: 9.83 Sildenafil / Sildenafil plus bosentan: 9.81	Standard treatment / Standard treatment: \$316,114 Beraprost / Standard treatment: \$1,257,685 Sildenafil / Standard treatment: \$1,597,286 Sildenafil / Sildenafil plus iloprost: \$1,996,592 Sildenafil / Sildenafil plus bosentan: \$1,796,971	Standard treatment / Standard treatment: Reference Beraprost / Standard treatment: 194,244 THB/QALY Sildenafil / Standard treatment: 194,384 THB/QALY Sildenafil / Sildenafil plus iloprost: 4,292,518 THB/QALY Sildenafil / Sildenafil plus bosentan: 2,699,191 THB/QALY
Tran-Duy, 2021(48)	Country: Australia Perspective: Payer Treatment: Combination therapy Comparator: Monotherapy Model type: Probabilistic discrete-time simulation model Health states: NR Time horizon: NR Cycle length: 1 year	SSc-PAH	Monotherapy: 3.02 Combination therapy: 3.9	Monotherapy: AU\$163,113 Combination therapy: AU\$262,539	Combination therapy / Monotherapy – 113,823 AUD/QALY

Study	Summary of model	Patient population	QALYs	Total costs	ICER (base case only)
	½ cycle correction: No Discounting: 5%				
Włodarczyk, 2016(42)	Country: Sweden Perspective: Societal Treatment: Selexipag Comparator: Iloprost Model type: Individual patient simulation model Health states: NR Time horizon: Lifetime Cycle length: NR ½ Cycle correction: NR Discounting: 3%	PAH patients already receiving an ERA	Selexipag: 0.88 Iloprost: NR	Selexipag: kr33,032 Iloprost: NR	Selexipag / Iloprost: 37,350 SEK/QALY
Additional economic analyses identified outside the SLR					
ICER, 2023(50)	Country: US Perspective: Health care sector Treatment: Sotatercept Comparator: SoC Model type: Decision analytic model Health states: WHO FC I, II, III, IV and death state Time horizon: Lifetime Cycle length: 12 weeks ½ Cycle correction: NR Discounting: 3%	PAH patients on sotatercept + background therapy or background therapy alone, FC II or III	Sotatercept / background therapy: 3.41 Background therapy alone: 2.51	Sotatercept: US\$400,000 ^a	Sotatercept / background therapy: 2,380,000 USD/QALY

AE: adverse event; AUD: Australian dollars; CAD: Canadian dollars; ERA: endothelin receptor antagonist; EUR: Euro; FC: functional class; ICER: incremental cost-effectiveness ratio; NR: not reported; NYHA: New York Heart Association classification; PAH: pulmonary arterial hypertension; QALY: quality-adjusted life-year; SEK: Swedish krona; SSc: systemic sclerosis; THB: Thai baht; USD: United States dollars; WHO: World Health Organization

^a Analysis assumes a placeholder price for sotatercept of \$400,000 per year

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Table 70 Summary list of cost-effectiveness studies for PAH therapies identified in HTA submissions

HTA submission	Therapy	Model type	Cost-effectiveness	Recommendation
Prostacyclin analogues				
PBS (2006)(52)	Epoprostenol Comparator: Bosentan	Cost-minimisation analysis	NR ICER: NR	Recommended: Yes PBAC recommendation: Epoprostenol is accepted on the basis that it is therapeutically no worse than bosentan.
CADTH (2004)(51)	Treprostinil Comparator: NR	NR	NR ICER: NR	Recommended: Do not list CEDAC recommendation: Treprostinil is not accepted as a treatment option for NYHA III or IV PAH.
Prostacyclin receptor agonist				
AWMSG (2018)(53)	Selexipag Comparator: Iloprost	Cost-utility analysis in a patient-level micro-simulation Markov model	Selexipag is less costly than Iloprost in the base case. This results in selexipag being the dominant treatment option. ICER: Redacted	Recommended: Yes AWMSG recommendation: Selexipag is accepted as an option for restricted use in adults for the long-term treatment of WHO FC II or III PAH. Is recommended either as combination therapy in patients insufficiently controlled with an ERA and/or a PDE5i, or as monotherapy.
CADTH (2016)(54)	Selexipag Comparator: Standard care	Cost-utility analysis in a patient-level micro-simulation Markov model	A price reduction of at least 42% is required for selexipag in addition to current therapy to be considered cost-effective when considering an ICUR of 50,000 CAD/QALY compared with current therapy alone. ICER: 187,418 CAD/QALY	Recommended: Yes CEDAC recommendation: Selexipag is accepted as long-term treatment of adults with WHO FC II or III PAH (idiopathic, heritable, PAH associated with connective tissue disorders, and PAH associated with congenital heart disease) to delay disease progression. Recommended for those demonstrating inadequate control with a first- and second-line PAH therapy. It is restricted to initiation by specialist clinicians with experience in the diagnosis and treatment of PAH.
NCPE (2018)(55)	Selexipag Comparator: NR	Cost-utility analysis in a patient-level micro-simulation	Selexipag not be considered for reimbursement unless cost-effectiveness can be improved relative to existing treatments. ICER: 402,530 EUR/QALY	Recommended: Yes ERG recommendation: Selexipag is accepted for long-term treatment of WHO FC II or III PAH, either as combination therapy in patients insufficiently controlled with an ERA and/or

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

HTA submission	Therapy	Model type	Cost-effectiveness	Recommendation
		Markov model		PDE5i, or as monotherapy in patients who are not candidates for these therapies.
SMC (2018)(56)	Selexipag Comparator: Inhaled iloprost	Cost-utility analysis in a patient-level micro-simulation Markov model	For both the base case analysis and the subgroup analysis with the PAS, selexipag was estimated to be a cost-effective treatment option versus inhaled iloprost. ICER: 46,538 GBP/QALY	Recommended: Yes SMC recommendation: Selexipag is accepted for restricted use for the long-term treatment of adults with WHO FC II or III PAH either as combination therapy in patients insufficiently controlled with ERA and/or PDE5i, or as monotherapy in patients who are not candidates for these therapies.
PBS 2020 (75)	Selexipag Comparator: Placebo	Cost-utility analysis in a Markov state transition model	Base case ICER: 45,000 to 75,000 AUD/QALY Reducing selexipag price per the pre-PBAC response, ICER: 15,000 to 45,000 AUD/QALY	Recommend: Yes PBAC recommendation: Selexipag is accepted, under Section 100 – Highly Specialised Drugs, for the treatment of patients with WHO FC III or IV PAH as a triple therapy in combination with an ERA and a PDE-5i.
Phosphodiesterase type-5 inhibitor				
AWMSG (2010)(57)	Sildenafil Comparator: Ambrisentan	Cost-minimisation analysis)	Sildenafil is less expensive than with ambrisentan. ICER: NR	Recommended: Yes AWMSG recommendation: Sildenafil is accepted as a treatment option for PAH (WHO FC II or III) to improve exercise capacity. Use to be restricted to PAH specialist physicians. It is not suitable for shared care between primary and secondary care within NHS Wales.
CADTH (2007)(58)	Sildenafil Comparator: Bosentan	Cost-effectiveness model	Sildenafil provides an alternative treatment option to bosentan and is less expensive. ICER: 46,000 CAD/QALY	Recommended: Yes CEDAC recommendation: Sildenafil accepted for use as a treatment for WHO FC III PAH patients (primary or secondary to connective tissue disease with inadequate response to conventional therapy). Sildenafil is not recommended for WHO FC II PAH patients.

HTA submission	Therapy	Model type	Cost-effectiveness	Recommendation
PBS (2006)(59)	Sildenafil Comparator: Bosentan	Cost-minimisation analysis	In Year 5 of listing the financial savings to the PBS would be < USD 10 million. ICER: NR	Recommended: Yes PBAC recommendation: Sildenafil is accepted for use as a treatment for PAH and connective tissue disease–associated PH in the WHO FC III group. Sildenafil demonstrated similar efficacy and toxicity as bosentan, therefore PBAC recommended listing with the same restrictions – no increased maximum quantities or repeats will be authorised consistent with the approved maximum dosage.
SMC (2006)(60)	Sildenafil Comparator: Bosentan	Markov model	Sildenafil provides an alternative treatment option to bosentan and is less expensive. ICER: NR	Recommended: Yes SMC recommendation: Sildenafil is accepted for restricted use as a treatment for WHO FC III PAH patients, to improve exercise capacity. It is restricted to initiation by specialists working in the Scottish Pulmonary Vascular Unit and by physicians experienced in the management of pulmonary vascular disease.
SMC (2010)(61)	Sildenafil Comparator: Ambrisentan	Cost-minimisation analysis	Sildenafil provides an alternative treatment option to ambrisentan and is less expensive. ICER: NR	Recommended: Yes SMC recommendation: Sildenafil is accepted for restricted use as a treatment for WHO FC III PAH patients, to improve exercise capacity. It is restricted to initiation by specialists working in the Scottish Pulmonary Vascular Unit and by physicians experienced in the management of pulmonary vascular disease.
CADTH (2010)(62)	Tadalafil Comparator: Sildenafil	Cost-utility analysis	The daily cost of tadalafil is less than the daily cost of sildenafil and considerably less than other PAH therapies. ICER: 70,753 CAD/QALY	Recommended: Yes CEDAC recommendation: Tadalafil is accepted for use as a treatment for PAH in a similar manner to sildenafil if the cost of tadalafil does not exceed the cost of sildenafil.
SMC (2012)(63)	Tadalafil Comparator: Sildenafil	Cost-minimisation analysis	The potential cost-saving with tadalafil may be higher than sildenafil. ICER: NR	Recommended: Yes SMC recommendation: Tadalafil is accepted for restricted use as treatment of adults with WHO FC II and III PAH, to improve exercise capacity. Restricted to initiation by specialists working in the Scottish Pulmonary Vascular Unit or similar specialists.
Guanylate cyclase stimulator				

HTA submission	Therapy	Model type	Cost-effectiveness	Recommendation
AWMSG (2015)(64)	Riociguat Comparator: Bosentan, ambrisentan and macitentan	Cost-minimisation analysis	Riociguat has more associated costs than other available PAH therapies. ICER: NR	Recommended: Yes AWMSG recommendation: Riociguat is accepted for restricted use as an alternative option to ERA monotherapy for the treatment of adult patients with WHO FC II and III to improve exercise capacity. Restricted to initiation by specialists working in the Pulmonary Vascular Unit or similar specialists.
CADTH (2014)(65)	Riociguat Comparator: PDE5is, ERAs, PCs and riociguat	Cost-minimisation analysis	Riociguat as first-line therapy has more associated costs than other available PAH therapies (including ERAs and PDE5is). ICER: NR	Recommended: Yes CEDAC recommendation: Riociguat is accepted as a treatment for WHO FC II and III PAH as monotherapy or in combination with ERAs if inadequate control with a first-line PAH therapy.
SMC (2015)(66)	Riociguat Comparator: Bosentan, ambrisentan and macitentan	Cost-minimisation analysis	With the Patient Access Scheme, riociguat became a cost-effective treatment option compared to bosentan, ambrisentan and macitentan. ICER: NR	Recommended: Yes SMC recommendation: Riociguat is accepted for restricted use as an alternative option to ERA monotherapy for the treatment of adult patients with WHO FC II and III to improve exercise capacity. Restricted to initiation by specialists working in the Scottish Pulmonary Vascular Unit or similar specialists.
Endothelin receptor agonist				
CADTH (2008)(67)	Ambrisentan Comparator: Sildenafil, treprostinil and bosentan	Cost-minimisation analysis	The total annual costs per patient for ambrisentan were substantially more compared to sildenafil but less compared to bosentan and treprostinil. ICER: NR	Recommended: Yes CEDAC recommendation: Ambrisentan (up to 10 mg daily dose) is accepted for the treatment of patients with at least WHO FC III PAH (either idiopathic or associate with connective tissue disease). The use is restricted to patients that have failed to respond to sildenafil or who have contraindications to sildenafil. Funding of ambrisentan given concomitantly with other ERAs, epoprostenol, treprostinil, or sildenafil is not recommended.
NCPE (2009)(68)	Ambrisentan Comparator: Bosentan and sitaxentan	Discrete event simulation	Ambrisentan is cost-effective and was demonstrated to be dominant compared with bosentan and sitaxentan. ICER: <45,000 EUR/QALY	Recommended: Yes ERG recommendation: Ambrisentan is accepted for the treatment of patients with at least WHO FC II or III PAH as a cost-effective alternative to bosentan and sitaxentan. The cost-effectiveness of ambrisentan was very sensitive to the

HTA submission	Therapy	Model type	Cost-effectiveness	Recommendation
				assumption of cost equality with the comparator ERAs. Therefore, pricing of ambrisentan at or below the current price of comparator ERAs would maximise cost-effectiveness.
SMC (2008)(69)	Ambrisentan Comparator: Bosentan and sitaxentan	Discrete event simulation	Ambrisentan is cost-effective and was demonstrated to be dominant compared with bosentan and sitaxentan. ICER: NR	Recommended: Yes SMC recommendation: Ambrisentan is accepted for restricted use as an alternative ERA option for the treatment of adult patients with WHO FC II and III (idiopathic PAH and in PAH associated with connective tissue disease) to improve exercise capacity. It is restricted to initiation and prescribing by specialists in the Scottish Pulmonary Vascular Unit or similar specialists.
PBS (2008)(70)	Bosentan Comparator: Placebo	NR	Bosentan has an acceptable cost-effectiveness compared with standard care. ICER: Between 45,000 and 75,000 AUD/QALY at 15 years	Recommended: Yes PBAC recommendation: Bosentan is accepted for the treatment of WHO FC III or IV PAH associated with congenital systemic to pulmonary shunts including Eisenmenger's physiology (APAH-CHD), PAH associated with scleroderma, idiopathic and heritable PAH.
AWMSG (2016)(71)	Macitentan Comparator: Bosentan and ambrisentan	Cost-minimisation analysis	The discounted price of macitentan is estimated to be cost-saving compared with bosentan and ambrisentan. ICER: NR	Recommended: Yes AWMSG recommendation: Macitentan is accepted as an option for use as monotherapy or in combination for the long-term treatment of adult patients exhibiting WHO FC II and III PAH.
CADTH (2015)(72)	Macitentan Comparator: Bosentan	Transition state model	Reduction in price to ensure that the drug plan cost for macitentan does not exceed the drug plan cost for bosentan. ICER: NR	Recommended: Yes CEDAC recommendation: Macitentan is accepted for the long-term treatment of WHO FC II or III PAH (idiopathic or heritable, or associated with connective tissue disease or congenital heart disease) to reduce morbidity. Recommended in cases where contraindication or inadequate response to sildenafil or tadalafil have been demonstrated.
SMC (2014)(73)	Macitentan Comparator: Bosentan and ambrisentan	Cost-minimisation analysis using a	When the Patient Access Scheme was taken into account, macitentan became a cost-effective treatment option	Recommended: Yes SMC recommendation: Macitentan is accepted for use as long-term treatment of adult patients with WHO FC II and III. It is

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

HTA submission	Therapy	Model type	Cost-effectiveness	Recommendation
		semi-Markov model	when compared with bosentan and ambrisentan. ICER: NR	restricted to initiation and prescribing by specialists in the Scottish Pulmonary Vascular Unit or similar specialists.
Combination therapies				
CADTH (2021)(74)	Macitentan and Tadalafil Comparator: Macitentan and tadalafil taken individually	Cost-comparison analysis	The annual cost savings associated with macitentan-tadalafil compared with macitentan and tadalafil taken as individual products at the same dose ranged from CAD 7,388 to 9,140 per patient. ICER: NR	Recommended: Yes CEDAC recommendation: Macitentan-tadalafil combination therapy is accepted for the long-term treatment of WHO FC II or III PAH (idiopathic, heritable, or associated with connective tissue disease or congenital heart disease) to reduce morbidity. Macitentan-tadalafil FDC must only be used in patients who are currently treated concomitantly with stable doses of macitentan 10 mg and tadalafil 40 mg as separate tablets. Must be prescribed by a specialist with expertise in managing and treating patients with PAH.

AUD: Australian dollars; AWMSG: All-Wales Medicines Strategy Group; CAD: Canadian dollars; CADTH: Canadian Agency for Drugs and Technologies in Health; CEDAC: Canadian Expert Drug Advisory Committee; ERA: endothelin receptor antagonists; ERG: Evidence Review Group; EUR: Euro; FDC: fixed-dose combination; GBP: British pound sterling; HTA: health technology assessment; ICER: incremental cost-effectiveness ratio; ICUR: incremental cost-utility ratio; NCPE: National Centre for Pharmacoeconomics, Ireland; NHS: National Health Service; NR: not reported; NYHA: New York Heart Association classification; PAH: pulmonary arterial hypertension; PAS = patient access scheme; PBAC: Pharmaceutical Benefits Advisory Committee; PBS: Pharmaceutical Benefits Scheme; PC: prostacyclin; PDE5i: phosphodiesterase 5 inhibitors; QALY: quality-adjusted life-years; SMC: Scottish Medicines Consortium; USD: United States dollars; WHO FC: World Health Organization Functional Class.

3.2 Economic analysis

Historically, Markov model structures in PAH have been informed by the WHO FC categorisation (see section 3.1). NHSE currently commissions treatments based on WHO FC, however, as discussed in section 1.3.2, European guidelines have transitioned patient categorisation and treatment pathway to ESC/ERS risk status.

Feedback from UK clinical experts confirmed that their patient management aligns with these guidelines (see Appendix R), suggesting that risk status is a more clinically appropriate method for constructing the model for this decision problem. Specifically, the WHO FC system has been described as subjective in nature and insensitive to changes in mortality risk. Although there is some overlap between WHO FC and risk status patient profiles, the other parameters in the risk status estimation method are also important predictors of patient survival (i.e., carry an equivalent weighting in the determination; see Table 4). A model which exclusively categorises patients according to WHO FC risks:

- diverging from the current treatment pathway (i.e., treating patients in the model with the wrong treatment regimen)
- failing to characterise the clinical objective of treatment (i.e., achieving low risk status)
- producing clinically implausible outcomes.

The appropriateness of categorising patients according to ESC/ERS risk status rather than WHO FC was also endorsed by stakeholders (NHS England) in the Scope document of this reimbursement submission ID6163 (76).

Therefore, this Markov model utilises risk status health states to assign study treatments, healthcare resource use (HCRU) and costs, HRQoL, treatment escalation and continuation criteria, and mortality risk. As mentioned previously, previous presentation of patient-level micro-simulation models in PAH has been criticised.

In the absence of established precedent in the literature, the model inputs have been sourced from the STELLAR trial, real-world evidence generation, clinical expert input, and assumptions have been made where it was deemed appropriate.

3.2.1 Patient population

As detailed in section 1.3.3, the expected population who will initiate sotatercept, when combined with PDE5i + ERA, consists of intermediate-low risk patients who would otherwise be treated with selexipag in combination with PDE5i + ERA. Low risk patients receiving PDE5i + ERA are considered to be at target without the need for additional therapy; intermediate-high risk patients are currently initiated with IV PCA and comparative evidence vs. initiation

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

with sotatercept is unavailable; and data reporting the benefits of initiating sotatercept in high risk patients [REDACTED]

At the start of the model, all patients are positioned in the intermediate-low risk health state, representing the initial health condition of the target population before initiating treatment.

The intermediate-low risk group were the most prevalent at baseline in STELLAR (44.2%), however where appropriate and available, data representing not only the target population, but also the entirety of the STELLAR trial population were used, reflecting the absence of an identified subgroup effect in the trial (see section 2.8.1).

3.2.2 Model structure

A Markov cohort model with an initial decision tree structure for the first 24 weeks of treatment was developed to simulate the disease pathway of PAH, estimating the associated clinical events and costs depending on the risk status membership, determined by the occurrence of improvement, maintenance, or worsening of risk status; lung/heart transplant; and death.

More specifically, the core model framework consists of four mutually exclusive health states: low risk, intermediate-low risk, intermediate-high risk and high risk. These health states were defined in line with the refined four-strata risk assessment for PAH, that has been validated by multiple PAH registries and recommended in the 2022 ESC/ERS treatment guideline for PAH given its established prognostic relevance for one-year survival (4, 5, 77).

Furthermore, all risk-based health states for the sotatercept arm are stratified based on whether patients are on- or off sotatercept treatment based on-treatment continuation criteria. A description of all health states is presented in Table 71 followed by a graphical depiction of the possible movements in Figure 14 and Figure 15.

Table 71 Health state definitions (base case)

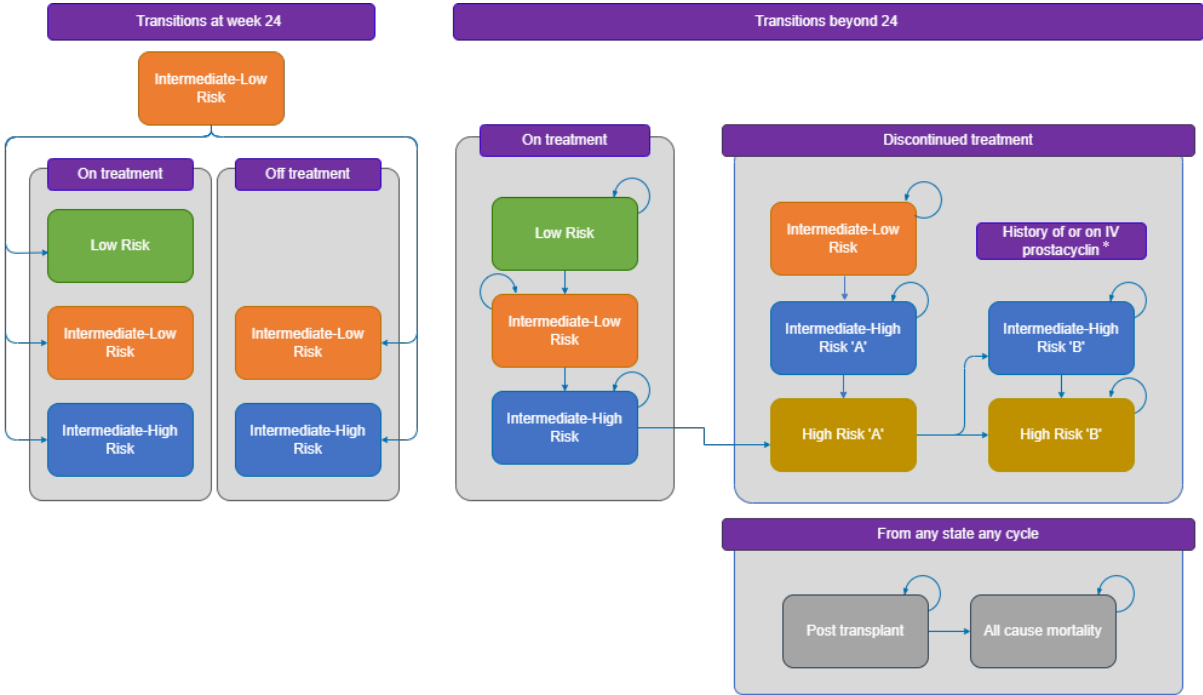
Health State	Definition	Sotatercept Arm	Selexipag Arm
Low risk	Patients with PAH without limitation of physical activity. State contains patients who have improved at week 24 following initiation of sotatercept or selexipag in intermediate-low risk.	PDE5i+ ERA+ sotatercept	PDE5i+ ERA+ selexipag
Intermediate-low risk	PAH results in slight limitation of physical activity. Ordinary physical activity causes dyspnoea, fatigue, chest pain or near syncope. Comfortable at rest. All patients start in this health state at baseline and are redistributed at week 24 following improvement, maintenance or worsening of risk status. After week 24, patients can enter intermediate-low risk only through worsening of disease from low risk.	PDE5i+ ERA+ sotatercept	PDE5i+ ERA+ selexipag
Intermediate-high risk	PAH results in marked limitation of physical activity. Less than ordinary activity causes dyspnoea, fatigue, chest pain or near syncope. Comfortable at rest. State contains patients who have worsened from a lower risk health state at any time point.	PDE5i+ ERA+ sotatercept	PDE5i+ ERA+ IV PCA (i.e., discontinue selexipag)
High risk	PAH causes inability to perform physical activity without symptoms. Dyspnoea or fatigue may be present at rest. State contains patients who have worsened from a lower risk health state at any time point.	Not applicable (patients who progress to high risk discontinue sotatercept and transition to 'High Risk 'A')	PDE5i+ ERA+ IV PCA
Low risk (off sotatercept)	Patients with PAH without limitation of physical activity achieving low risk with SoC following sotatercept discontinuation (not assumed to be possible).	PDE5i+ ERA	Not applicable (relates to sotatercept arm only)

Health State	Definition	Sotatercept Arm	Selexipag Arm
Intermediate-low risk (off sotatercept)	PAH results in slight limitation of physical activity. Ordinary physical activity causes dyspnoea, fatigue, chest pain or near syncope. Comfortable at rest. State contains patients who have discontinued sotatercept at 24 weeks without progression to a worse risk state.	PDE5i+ ERA	Not applicable (relates to sotatercept arm only)
Intermediate-high risk 'A' (off sotatercept)	PAH results in marked limitation of physical activity. Less than ordinary activity causes dyspnoea, fatigue, chest pain or near syncope. Comfortable at rest. State contains patients who have discontinued sotatercept at 24 weeks and worsened to intermediate-high risk. Patients initiate and continue to receive IV PCA until they progress to High Risk 'A'.	PDE5i+ ERA+ IV PCA	Not applicable (relates to sotatercept arm only)
High risk 'A' (off sotatercept)	PAH causes inability to perform physical activity without symptoms. Dyspnoea or fatigue may be present at rest. State contains patients who have discontinued sotatercept upon disease progression to high risk. Patients remain in this health state for one model cycle only with a chance for improvement upon initiation of parenteral PCA (even though the possibility for improvement is capped at 24 weeks), whereafter patients either move to the high risk 'B' state permanently or improve to intermediate-high risk 'B'.	PDE5i+ ERA+ IV PCA	Not applicable (relates to sotatercept arm only)
Intermediate-high risk 'B' (off sotatercept)	PAH results in marked limitation of physical activity. Less than ordinary activity causes dyspnoea, fatigue, chest pain or near syncope. Comfortable at rest.	PDE5i+ ERA+ IV PCA	Not applicable (relates to sotatercept arm only)

Health State	Definition	Sotatercept Arm	Selexipag Arm
	State contains patients who have benefitted from the one-cycle opportunity to improve from High Risk 'A' upon initiating IV PCA and remain here until progression to High Risk 'B', receive transplantation or die.		
High risk 'B' (off sotatercept)	PAH causes inability to perform physical activity without symptoms. Dyspnoea or fatigue may be present at rest. Patients who deteriorate from Intermediate-high risk 'B' to High Risk 'B' will remain in this state until they receive transplantation or die.	PDE5i+ ERA+ IV PCA	Not applicable (relates to sotatercept arm only)
Post-transplant	Semi-absorbing state. Patients who have undergone a transplant and can no longer develop PAH. Patients remain in this state until they die.	No active PAH treatment	No active PAH treatment
Death	Absorbing state: patients who have died from any risk status or post-transplant health state.	Not applicable	Not applicable

PDE5i = Phosphodiesterase 5 inhibitor; ERA = Endothelin receptor antagonists; PCA = Prostacyclin analogue; IV = intravenous; PAH = pulmonary arterial hypertension.

Figure 14 Model diagram – Possible Movements (sotatercept)



* Refers to the history of PCA use in the high risk 'A' health state

Figure 15 Model diagram – Possible Movements (selexipag)

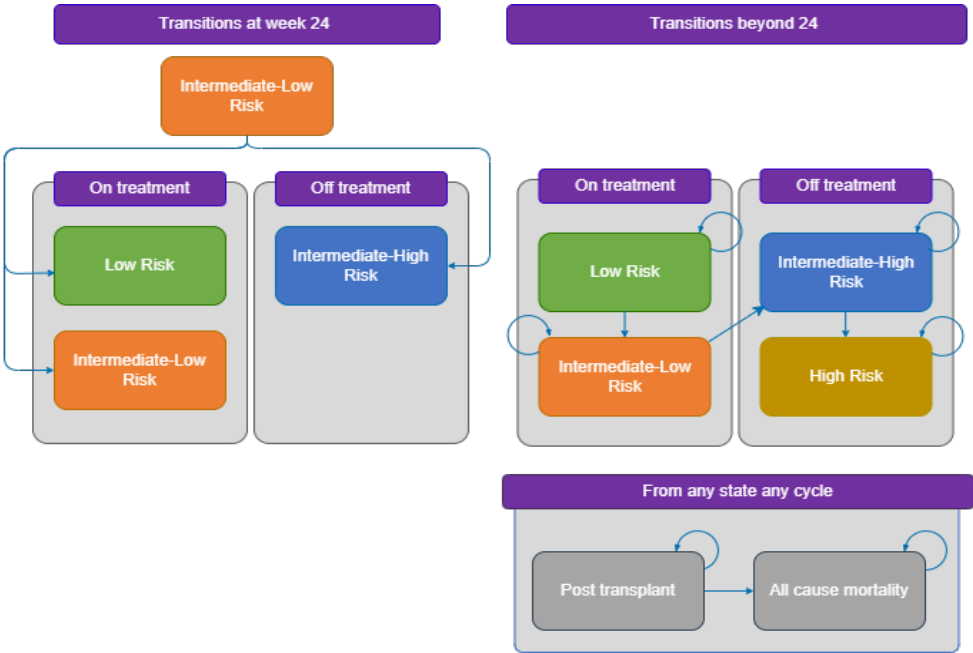


Table 72 Features of the economic analysis

Factor	Current evaluation	
	Chosen values	Justification
Summary of analytic methods	Markov model (with decision tree structure covering first 24 weeks)	Allows transparent incorporation of movement between mutually exclusive health states.
Patient population	Intermediate-low risk patients at baseline	Expected patient population for sotatercept-based on local clinical practice, feedback from UK clinical experts and availability of relevant evidence.
Time horizon	Lifetime (30 years)	Given the median starting age of 48 years for the STELLAR population, and the progressive nature of PAH, only 0.06% of patients are projected to be alive 30 years following treatment initiation, which therefore can be considered a lifetime horizon.
Perspective	NHS & PSS	Aligns with NICE guide to the methods of technology appraisal.
Discount rate	3.5% for costs and benefits	Aligns with NICE guide to the methods of technology appraisal.
Cycle length	1 st cycle (decision tree): 24 weeks 2 ^{nd+} cycles: 12 weeks	1 st cycle: transparent incorporation of treatment discontinuation. 2 ^{nd+} cycles: it is sufficiently short to identify incremental changes in the risk categorisation of patients.
Half-cycle correction	Applied	Half-cycle correction is implemented to account for the possibility of patients transitioning between different health states at any given point within a cycle. This correction helps to accurately reflect the costs and effectiveness associated with these transitions.
Treatment waning effect?	No	Maintenance of response achieved was examined and shown based on SOTERIA. Week 12-24 transitions are assumed to continue over a lifetime. To address uncertainty over the long-term benefit, a conservative assumption was adopted that all improvement beyond 24 weeks converts to maintenance, whereas in the SOTERIA study, continual improvement or maintenance from baseline has been observed.
Source of utilities	Utilities derived based on STELLAR	Incorporated using direct evidence from the relevant clinical trial while also adhering to the health state definitions according to the risk categorisation. Due to lack of population specific data, it was assumed that post-transplant patients'

Factor	Current evaluation	
	Chosen values	Justification
		quality of life is equal to those in the low risk health state.
Source of costs	NHS reference costs, BNF and eMIT, literature.	

BNF: British National Formulary; eMIT: Electronic Market Information Tool; NHS: National Health Service; NICE: National Institute for Health and Care Excellence; PSS: personal social services

Continuation criteria

Based on clinical expert opinion, discontinuation of sotatercept should be considered in the case of inadequate responders, therefore in order to reflect this and more accurately estimate treatment costs for the NHS, a rule in the model has been implemented whereby patients who do not adequately respond to sotatercept therapy in the first 24 weeks will discontinue treatment.

Based on clinical expert feedback received from UK experts, inadequate response and therefore discontinuation of sotatercept treatment should be considered based on concurrent evaluation of change from baseline for the following three endpoints: 6MWD, WHO FC and NT-proBNP.

According to the expert feedback, inadequate response is defined as a worsening in at least one of the three endpoints (6MWD, WHO FC and NT-proBNP) used in risk status assessment. Worsening of each endpoint is defined as:

- WHO FC – a higher WHO FC Score at Week 24 compared to baseline.
- 6MWD/NT-proBNP – a higher number of points assigned at Week 24 compared to baseline using COMPERA risk categories while accounting for a $\geq 10\%$ deterioration in endpoint raw scores. Thus, only a decrease/increase of $\geq 10\%$ in 6MWD/NT-proBNP value respectively at Week 24 compared to baseline is considered a deterioration. The analysis used a 10% cutoff to make sure only clinically meaningful deteriorations are captured.

The clinical expert feedback was that these criteria would be unique to sotatercept. A post-hoc analysis of STELLAR was used to estimate the proportion of sotatercept patients meeting these criteria. Results of the analysis (i.e., four-stratum risk assessment transition probabilities) from baseline to week 24 with risk status assessment at both baseline and week 24 and risk status deterioration at week 24 while accounting for a $\geq 10\%$ deterioration in continuous endpoints (6MWD and NT-proBNP) are presented in the following table. In total 5

patients met the criteria, 3 did not change risk status and 2 patients moved to intermediate-low risk health state.

Table 73 Risk Status Deterioration Between Baseline and Week 24

N		To Week 24			
From baseline	n ^b	Low Risk	Intermediate-Low Risk	Intermediate-High Risk	High Risk
Intermediate-Low Risk	1	3	3	2	2

^a Number of participants: full analysis set, participants with risk status assessment at both baseline and Week 24

^b Number of participants: full analysis set population, participants with non-missing risk assessment stratum at both analysis visits and risk status deterioration at Week 24 while accounting for a ≥ 10% deterioration in continuous endpoints (6MWD and NT-proBNP). Row percentages are based on this number. Risk status deterioration is determined by a worsening in WHO FC: a higher WHO FC Score at Week 24 compared to baseline or a worsening in 6MWD/NT-proBNP: a higher number of points assigned at Week 24 compared to baseline using COMPERA risk categories while accounting for a ≥ 10% deterioration in endpoint raw scores. Risk stratum assessment is based on COMPERA version 2.0

Study: STELLAR (MK-7962-003)
Database Cutoff Date: 06DEC2022

In addition, in the base case, patients in any of the pre-transplant states whose disease worsens at any timepoint to become high risk will discontinue treatment due to ██████████ ██████████ for sotatercept’s efficacy in these patients. STELLAR only recruited ██████████ with a baseline high risk status and ██████████ received sotatercept. Additional evidence from future trials and more clinical experience with sotatercept may establish it as a treatment for initiation in high risk patients, so the impact of this assumption is explored in a scenario analysis.

Based on clinical feedback, patients are assumed to discontinue selexipag treatment when they move to a worse health state than intermediate-low risk, where IV PCA treatment is initiated. This reflects the ESC/ERS treatment guidelines, the WSPH algorithm (summarised in Figure 2) (28), and clinical advice that IV PCA and selexipag are never administered concomitantly.

3.2.3 Intervention technology and comparators

The intervention, sotatercept, is implemented within the model as per its expected pathway position (Section 1.3.3) and is given according to the recommended dosing regimen. The comparative treatments are also implemented as per current clinical practice (i.e., ESC/ERS guidelines) and are given according to their licensed dosing regimens.

In order to align with current clinical practice for the intermediate-low risk population, the following treatment regimens were included in the base case comparison:

- PDE5i + ERA + sotatercept

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

- PDE5i + ERA + selexipag

In current clinical practice clinicians initiate treatment with selexipag in this subpopulation, however it should be noted that according to current NHS England reimbursement guidelines selexipag + dual therapy is restricted to patients who are WHO FC III (78). Clinical advice received is that reimbursement guidelines are expected to soon reflect risk status. As the risk categories are composite measures of survival risk, they can include patients from multiple WHO FC states at the same time. In STELLAR 42.3% of the baseline intermediate-low group patients belonged to WHO FC II and 57.7% to WHO FC III.

The most recent ESC/ERS guideline states that “in patients who are at intermediate-low risk despite receiving ERA/PDE5i therapy, adding selexipag should be considered to reduce the risk of clinical worsening” (4).

UK clinicians confirmed that selexipag is used in intermediate-low risk patients in the UK according to the recommendation of the ESC/ERS guideline therefore the comparison of PDE5i + ERA + sotatercept vs. PDE5i + ERA + selexipag is appropriate.

It should be noted that the comparator in the STELLAR trial was placebo added to background therapy, therefore the transitions observed in the trial comparator arm are not appropriate for use in this cost-effectiveness analysis. However, this comparator arm includes background triple therapy with a subset of patients who received selexipag in combination with PDE5i+ERA. It is possible to utilise this subset as a proxy for the comparator relevant to this appraisal, and the relevant outcome of risk status change was recorded. This analysis has been described previously in section 2.8.1 and its application in the model is described in the next section.

3.3 Clinical parameters and variables

3.3.1 Baseline characteristics

STELLAR is the pivotal trial for this indication hence the baseline patient characteristics applied in the model reflect the average PAH patient included in the STELLAR trial. It is not expected that patient characteristics (i.e., weight) would significantly differ based on the health state of the patient therefore it was assumed that weight remains constant over time. To better reflect the likely UK population in terms of weight distribution, mean weight of patients outside of the USA was derived from STELLAR, as clinical opinion questioned the generalisability of the USA trial participants to UK practice for this parameter. A scenario analysis using the mean weight of the full STELLAR population is also presented.

Table 74 Baseline characteristics (ITT Population)

Demographic parameter	STELLAR ITT population (N=323)	Reference
Female, n (%)	256 (79.3)	STELLAR(5)
Mean age, years	47.9	
Weight (SD)		STELLAR(5) excluding the USA trial participants

ITT: intent-to-treat; SD: standard deviation; USA: United States of America

3.3.2 Clinical parameters and variables

Overview of data sources

STELLAR trial data, publicly available data and real-world evidence have been used to estimate the clinical parameters used in the economic model. A summary is provided in Table 75.

Table 75 Summary of clinical data sources

Parameter	Source of parameter		Relevant section(s)
	Sotatercept	Selexipag	
Transition probability parameters			
Short-term	STELLAR based on transition from baseline to 24 weeks	RR of improvement or worsening based on within-trial analysis vs. sotatercept from STELLAR	Short-term transition probabilities (up to week 24)
Long-term	STELLAR based on transition from week 12 to 24 (assumed to be representative of subsequent 12-week cycles, except for risk status improvements, which are not permitted beyond week 24, see Table 72)		Long-term transition probabilities (post week 24)
Continuation rule	STELLAR, based on transition from baseline to 24 weeks and clinical expert opinion	Based on transition to worse health states than intermediate-low risk (28)	Continuation criteria
Mortality			
Reference + HR model	Low risk group fitted independently to KM presented by Rosenkranz et al. (2023) followed by application of HRs also derived from Rosenkranz et al. (2023) (79). Scenario analysis using low risk curves and HRs based on Boucly et al. (2022) (77).		Mortality
General population mortality	Office for National Statistics. National Life Tables: UK. 2017-2019.		Mortality
Lung/heart transplant			
Probability of lung/heart transplant	COMPERA. MSD data on file (2023). (80)		Lung/heart transplant
Post-transplant mortality	Bernstein et al. (2018) (81)		Lung/heart transplant
PAH hospitalisation			

Probability of PAH hospitalisation	COMPERA. MSD data on file (2023) (82)	PAH hospitalisation
HR for PAH hospitalisation	ITC of overall STELLAR vs. GRIPHON (WHO FC II or III only)	PAH hospitalisation

FC: functional class; HR: hazard ratio; ITC: indirect treatment comparison; PAH: pulmonary arterial hypertension; RR: relative risk; UK: United Kingdom; WHO: World Health Organization

The base case comparison of sotatercept versus selexipag (both as an add-on therapy to dual background therapy) relied on the within-STELLAR analysis previously described in section 2.8.1 due to the absence of a selexipag randomised trial that reports risk status change, or sufficient granular patient-level data to enable retrospective calculation of risk status over time. Given the limitations associated with this analysis and to provide a range of relative efficacy, a scenario analysis explores the impact of using the WHO FC change results from the trial-based Bucher ITC reported in section 2.10 as a proxy for risk status change.

Detailed description of the sources and inputs used are provided in the sections below.

Comparison of sotatercept vs. selexipag

As discussed in section 2.10, it was not feasible to conduct an ITC to estimate the relative treatment effect of sotatercept versus selexipag for risk status change as results for selexipag therapy were only published according to the WHO FC categorisation (the indirect comparison for this endpoint has been previously discussed [section 2.10]). At the time of submission, in order to appropriately parameterise the model, the aforementioned within-trial analysis of STELLAR was conducted to provide comparative data for risk status change for sotatercept versus selexipag, both added to background dual therapy. In this subgroup analysis patients in the sotatercept arm who received it in combination with ERA + PDE5i were compared with patients in the placebo arm who received it in combination with ERA + PDE5i + selexipag. This analysis involves some limitations.

- Potential bias may be introduced by breaking the randomisation between treatment arms, leading to potential imbalance in effect modifiers as STELLAR was not designed to test comparative efficacy between sotatercept and selexipag. However, as previously discussed, a review of the characteristics identified as important prognostic factors (see Appendix O) did not identify any noteworthy imbalances (see Table 61).
- Being a non-randomised subgroup, the sample size is reduced and uneven between arms, as illustrated in the table below. However, in a rare disease such as PAH, this sample size challenge is difficult to mitigate.
- It represents a comparison between sotatercept-naïve patients and selexipag-experienced patients, which may be a source of confounding.

However, more robust comparative evidence for this endpoint is not available. Acknowledging the limitations of this analysis, a scenario analysis using the ITC for WHO FC change as a proxy is also explored.

Table 76 Demographics and baseline characteristics – Participants who received double therapy (sotatercept: ERA+PDE5i) or triple therapy (placebo: ERA+PDE5i+prostacyclin [selexipag])

	PDE5i + ERA + sotatercept		PDE5i + ERA + selexipag + placebo		Total	
	n	(%)	n	(%)	n	(%)
Participants in population	■		■		■	
Age (years)						
Participants with data	■		■		■	
Mean	■		■		■	
SD	■		■		■	
Median	■		■		■	
Range	■		■		■	
Age group, n (%)						
18 to <45	■	■	■	■	■	■
45 to <65	■	■	■	■	■	■
65 to <75	■	■	■	■	■	■
≥75	■	■	■	■	■	■
Sex, n (%)						
Male	■	■	■	■	■	■
Female	■	■	■	■	■	■
Race, n (%)						
White	■	■	■	■	■	■
Black or African American	■	■	■	■		■
Asian	■	■	■	■	■	■
Other	■	■	■	■	■	■
Missing	■	■	■	■	■	■
Ethnicity, n (%)						
Hispanic or Latino	■	■	■	■	■	■
Not Hispanic or Latino	■	■	■	■	■	■
Not reported	■	■	■	■	■	■
Height (cm)						
Participants with data	■		■		■	
Mean	■		■		■	
SD	■		■		■	
Median	■		■		■	
Range	■		■		■	
Weight (kg)						
Participants with data	■		■		■	

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Mean						
SD						
Median						
Range						

ERA: endothelin receptor antagonists; PAH: pulmonary arterial hypertension; PDE5i: phosphodiesterase type 5 inhibitor; SD: standard deviation

The results of the subgroup analysis show a large increase in the proportion of patients achieving improvement with sotatercept in combination with dual therapy (████ vs. █████) while █████ sotatercept patients showed worsening of the disease compared to █████ of the patients in the placebo group (see Table 62 in section 2.8.1). █████ was applied. When estimating relative risks a common problem arises if one of the cells in a 2x2 contingency table █████ To address this, █████

This adjustment helps stabilise the estimates and prevents [REDACTED] without invalidation of the results_(see section 2.8.1 for detailed calculations). Since the control arm patients received no active additional treatment in STELLAR (i.e., placebo) the observed efficacy is thus attributed to their background therapy (i.e., PDE5i+ERA+selexipag).

Based on the above result, the derived relative risks of risk stratum worsening/improving are presented below. The clinical plausibility of these relative effects were discussed with UK clinical experts and were deemed to be credible, albeit with acknowledgement of the methodological limitations and the wide confidence intervals.

Table 77 Relative risks of risk stratum worsening/improving

Parameter	Mean	95% CI, LL	95% CI, UL	Source
RR, Risk stratum worsening (sotatercept vs. selexipag)	■	■	■	Within- STELLAR analysis
RR, Risk stratum improvement (sotatercept vs. selexipag)	■	■	■	

CI: confidence interval; LL: lower limit; RR: relative risk; UL: upper limit

Transition probabilities for the sotatercept arm were calculated using directly observed data from the STELLAR trial while transition probabilities for the selexipag arm were calculated by applying the inverse of the RRs of sotatercept vs. selexipag to the sotatercept arm (presented in Table 77).

In the process of applying the inverse of the RRs to the sotatercept arm transition probabilities were first transformed to rates. After applying the RRs, the new rates were transformed back into probabilities. To convert a probability into a rate, the following formula was used:

$$r = \frac{-\ln(1 - p)}{t}$$

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension
[ID6163]

where 'r' is the rate, 'p' the corresponding probability for improving/ worsening and 't' time. After multiplying the rate with the inverse of the RR, it is converted back into a probability, using the following formula:

$$p = 1 - e^{(-rt)}$$

Risk stratum transition probabilities

STELLAR data were used to inform the health state transition probabilities across the four risk strata (low risk, intermediate-low risk, intermediate-high risk and high risk) in the model. Clinical improvement was represented by improvement in risk stratum (e.g., from high risk to intermediate-high risk, from intermediate-high risk to intermediate-low risk, from intermediate-low risk to low risk). Clinical maintenance was represented by maintenance in the same risk stratum (e.g., from intermediate-low risk to intermediate-low risk). Clinical worsening was represented by worsening in risk stratum (e.g., from low risk to intermediate-low risk, from intermediate-low risk to intermediate-high risk, from intermediate-high risk to high risk).

Clinical effectiveness was modelled over two distinct time periods:

- The short-term period followed the STELLAR trial duration (24 weeks).
- The long-term period followed the time starting from the end of STELLAR week 24 and continued until the end of the model time horizon.

Given the absence of an identified subgroup effect and the limited trial sample size, the transition probabilities for the overall population are deemed to be representative of the intermediate-low risk population.

Short-term transition probabilities (up to week 24)

Given the availability of patient-level STELLAR data, observed STELLAR data by visit (i.e., the number of patients transitioning across the risk strata) from baseline to 24 weeks in the overall population were directly used to inform the short-term transitions for the sotatercept arm.

All patients start treatment in the intermediate-low population and are redistributed at the end of the first model cycle (decision tree) at week 24. According to the continuation rules outlined in section 3.2.2, patients could discontinue either combination regimen which is reflected in the distribution of patients at 24 weeks presented below. Overall, there were 5 sotatercept patients who were intermediate-low at baseline and who also showed ≥10% deterioration in continuous endpoints (6MWD and NT-proBNP) and/or worsening of the disease measured by Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

having a higher FC score at Week 24 and therefore they were assumed to discontinue sotatercept therapy. Of these 5 patients, 3 patients remained in intermediate-low risk at week 24 and 2 patients worsened to intermediate-high risk. Hence, the number of on-treatment intermediate-low risk patients in the sotatercept arm was calculated as 32 (total number of intermediate-low risk patients at week 24) minus 3 = 29 patients, as presented in the table below. In line with clinical guidelines (28) and expert advice, discontinuation of selexipag was defined as transition to a worse health state than intermediate-low risk.

Table 78 Distribution of patients at baseline and at 24 weeks (Sotatercept)

Health state	Treatment status	Distribution at baseline	Distribution at week 24
		Sotatercept, n (%)	Sotatercept, n (%)
Low risk	On sotatercept treatment	0.0%	■
Intermediate-low risk		■ (100.0%)	■
Intermediate-high risk		0.0%	■
Intermediate-low risk	Off sotatercept treatment	0.0%	■
Intermediate-high risk 'A'		0.0%	■

The RR for improvement/worsening was applied to these values to derive the selexipag transitions. Patients discontinued selexipag treatment when moving to intermediate-high and high risk health state.

Table 79 Distribution of patients at baseline and at 24 weeks (Selexipag)

Health state	Treatment status	Distribution at baseline	Distribution at week 24
		Selexipag (%)	Selexipag (%)
Low risk	On selexipag treatment	0.0%	■
Intermediate-low risk		100.0%	■
Intermediate-high risk	Off selexipag treatment	0.0%	■
High risk		0.0%	■

Long-term transition probabilities (post week 24)

In the base case, long-term (i.e., post week 24) transition probabilities relied on the observed STELLAR data from week 12 to 24, with the exception of any transitions to an improved risk status – which were assumed to be maintenance transitions instead. This conservative approach was used due to a lack of longer-term RCT evidence (i.e. STELLAR) and is supported by the absence of a treatment waning effect in the first interim analysis (IA) of the SOTERIA open-label extension study. Given the titration period required by the sotatercept SmPC, it is assumed that week 12 to 24 is a more reflective interval of the subsequent model cycles than the earlier transitions.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

SOTERIA is a phase 3, optional open-label long-term follow-up study with primary objective to assess the long-term safety and tolerability of sotatercept in PAH when added to background PAH therapy. Details of the SOTERIA study were provided in section 2.3.2.

Eligibility for enrolment in the study required prior participation and completion of relevant study requirements from the parent PAH sotatercept clinical trials. A notable parent PAH study is the STELLAR trial, and only SOTERIA participants from STELLAR are considered in this assessment of maintenance in order to understand if the STELLAR effect is maintained.

For the combined endpoint of improvement or maintenance, the number of participants who showed no worsening are presented for each time point, relative to Week 24 of the STELLAR trial. As outlined in the table below, [REDACTED]

[REDACTED] A low sample size is observed at the later timepoint as enrolment in SOTERIA is continuing i.e., with subsequent interim analyses, the number of patients reaching week 108 post-baseline is expected to grow. These results indicate a high degree of stability in the results initially achieved in the STELLAR trial.

Table 80 Analysis of the proportion of risk status improvement or maintenance relative to STELLAR Week 24

	Sotatercept (N ^a =153)		
Endpoint	n/N ^b	%	(95% CI) ^c
Risk status improvement or maintenance relative to STELLAR Week 24			
Week 36	[REDACTED]	[REDACTED]	[REDACTED]
Week 52 (Year 1)	[REDACTED]	[REDACTED]	[REDACTED]
Week 108 (Year 2)	[REDACTED]	[REDACTED]	[REDACTED]

Database Cutoff Date: STELLAR (MK-7962-003) - 06DEC2022 / SOTERIA (MK-7962-004) - 08NOV2023

CI: confidence interval; n: number of participants with improvement or maintenance of the endpoint assessment at a given timepoint relative to Week 24

^a Participants who enrolled to SOTERIA from STELLAR sotatercept arm and continued receiving sotatercept. 154 participants from STELLAR sotatercept arm enrolled to SOTERIA, but one participant discontinued before receiving study dose in SOTERIA.

^b Participants with non-missing endpoint assessment at both given timepoint and STELLAR Week 24. Row percentages are based on this number.

^c Derived based on the Clopper-Pearson method (exact binomial distribution).

SOTERIA provides robust evidence that the efficacy observed in STELLAR does not wane beyond week 24, but clinical expert opinion advised that confirmation of continued improvement beyond this timepoint will depend on additional data. Therefore, an appropriate assumption to make is to cap risk status improvement in the model beyond week 24. A similar assumption was made in the ICER assessment (50).

Following this timepoint patients do not experience clinical improvement, only clinical maintenance and worsening are possible. As a result, patients in the model cannot transition to health states with a more beneficial health profile than their current health state.

It was assumed that the long-term transition probabilities would remain constant over time for the remainder of the model time horizon. The existing evidence lacks a sufficiently long follow-up period to enable the investigation of the relationship between the duration spent in a particular health state and the likelihood of disease progression. The same approach was also utilised by ICER and was also the general assumption in the identified publications (50).

Given the relatively short follow-up period of the STELLAR trial and the marked clinical benefit observed with sotatercept [REDACTED] (see Table 81 below). However, considering the progressive nature of PAH, [REDACTED]

To inform this transition, a conservative assumption was made: [REDACTED]

[REDACTED] This means the transitions from FCIII are used in place of transitions from intermediate-high risk, in order to maintain a sum to 100%. Given the comparative results show greater efficacy for sotatercept on the risk status change endpoint, using FC as a proxy is likely to underestimate the treatment benefit in delaying progression to high risk; however, in the absence of alternative data, the conservative assumption is made.

Table 81 Risk stratum transition probabilities for sotatercept observed between week 12-24 in STELLAR

Transition (from: rows, to: columns)	Low risk	Intermediate-low risk	Intermediate-high risk	High risk
Patient numbers (n)				
Low risk [REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
Intermediate-low risk [REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
Intermediate-high risk [REDACTED] ^a	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
Intermediate-high risk (FC III proxy) [REDACTED] ^b	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
High risk [REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
Transition probabilities				
Low risk	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
Intermediate-low risk	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
Intermediate-high risk ^a	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
Intermediate-high risk (FC III proxy) ^b	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
High risk	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

FC: functional class

^a Not used in the model as 0 transitions from intermediate-high to high risk in the long-term is not considered clinically plausible.

^b These represent transitions from FCIII and are used as a proxy for transitions from intermediate-high risk.

As discussed previously, no improvement beyond 24 weeks was allowed, therefore the transitions to better states and maintenance of the health state were summed (i.e., Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163])

improvement transitions are converted to maintenance ones). Adjusted probabilities which are used in the model base case are presented below.

Table 82 Risk stratum transition probability inputs used in model for sotatercept (long-term): no improvement allowed beyond week 24

Transition (from: rows, to: columns)	Low risk	Intermediate-low risk	Intermediate-high risk	High risk
Patient numbers				
Low risk ■	■	■	■	■
Intermediate-low risk ■	■	■	■	■
Intermediate-high risk ■ ^a	■	■	■	■
Intermediate-high risk (FC III proxy) ^b	■	■	■	■
High risk ■	■	■	■	■
Transition probabilities				
Low risk	■	■	■	■
Intermediate-low risk	■	■	■	■
Intermediate-high risk ^a	■	■	■	■
Intermediate-high risk (FC III proxy) ^b	■	■	■	■
High risk	■	■	■	■

FC: functional class; NA: not applicable

^a Not used in the model as 0 transitions from intermediate-high to high risk in the long-term is not considered clinically plausible

^b These represent transitions from FCIII and are used as a proxy for transitions from intermediate-high risk

Long-term transitions for the selexipag arm were calculated by applying the RR of selexipag vs. sotatercept to long-term transitions for the sotatercept arm as discussed in the section 'Comparison of sotatercept vs. selexipag

'. This assumption was validated with UK clinical experts, who confirmed that selexipag has a limited value in improving a patient's risk status.

Table 83 Risk stratum transition probability inputs for selexipag (long-term)

Transition (from: rows, to: columns)	Low risk	Intermediate-low risk	Intermediate-high risk	High risk
Low risk	■	■	■	■
Intermediate-low risk	■	■	■	■

*sum of proportions improving and maintaining risk status

Based on the ESC/ERS guidelines and UK clinician feedback, selexipag is not considered for patients whose disease worsens to intermediate-high risk, which results in clinicians discontinuing selexipag in patients who progress beyond the intermediate-low risk group. Based on Table 83 above, this rule resulted in 41.6% of patients discontinuing treatment due to moving from intermediate-low to intermediate-high risk every 12 weeks.

Patients who discontinue sotatercept at 24 weeks receive background dual therapy without escalation to PCA if they remained in the intermediate-low risk health state and initiate infused PCA if they progressed to intermediate-high risk. Similarly, with selexipag, if patients progressed to intermediate-high risk they discontinue selexipag and initiate infused PCA. As previously stated, patients were assumed to only discontinue sotatercept by experiencing inadequate response at 24 weeks, otherwise patients remain on-treatment until they progress to high risk (when they initiate infused PCA without sotatercept). Transition probabilities according to the description above are presented below.

Table 84 Risk stratum transition probability inputs (off sotatercept and selexipag)

Transition (from: rows, to: columns)	Low risk	Intermediate-low risk	Intermediate-high risk	High risk	Source
Transition probabilities					
Low risk	■	■	■	■	STELLAR ^a (5, 83)
Intermediate-low risk	■	■	■	■	
Intermediate-high risk	N/A	N/A	91.2% (0%+16.9% +74.3%)	8.8%	Roman, 2012 (41)
High risk (selexipag only)	■	■	■	■	STELLAR (5, 83)

NA: not applicable

^a Week 12-24 transition probabilities based on placebo arm from STELLAR

Transition probabilities only utilised by the sotatercept arm (following discontinuation of treatment) of the model are in Table 85 presented below. Of note, the High Risk 'A' transition to Intermediate-high risk 'B' is the only allowed improvement in risk status beyond week 24. This allows patients in the model who have newly discontinued sotatercept upon progression to high risk and subsequently initiated PCA to have a one-cycle opportunity for risk status improvement.

Table 85 Risk stratum transition probability inputs (off sotatercept only)

Transition (from: rows, to: columns)	Intermediate-high risk 'A'	High risk 'A'	Intermediate-high risk 'B'	High risk 'B'	Source
Transition probabilities					
Intermediate-high risk 'A'	91.2%	8.8%	NA	NA	Roman, 2012 (41)
High risk 'A'	NA	NA	38.5%	61.5%	
Intermediate-high risk 'B'	NA	NA	91.2%	8.8%	
High risk 'B'	■	■	■	■	STELLAR (5, 83)

NA: not applicable

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

In the presence of competing events (i.e., to determine which event occurs first), the risk stratum transitions were further adjusted by applying first (1) probabilities of all-cause mortality (see 'Mortality' in section 3.3.2) and then (2) probabilities of lung/heart transplant (see 'Lung/heart transplant' in section 3.3.2) such that the sum of the transition probabilities (e.g., from low risk to low risk, intermediate-low risk, intermediate-high risk, high risk, lung/heart transplant and death) would equal 100%.

Lung/heart transplant

In the event of clinical worsening despite maximal pharmacotherapy, patients would be eligible to receive a lung/heart transplant as an end-stage therapeutic option (4). It was assumed in the model that only patients with intermediate-high and high risk PAH would be eligible for a lung/heart transplant, consistent with the 2022 treatment guideline, and would remain in this state until death (4).

Given the extremely small number of transplant events in STELLAR (and limited data in the literature in general, given the rare nature of transplant in patients with PAH), external data were used to inform transplant-related model inputs. Specifically, the probabilities of having a lung/heart transplant were obtained from unpublished real-world COMPERA registry data, (80) meanwhile post-transplant mortality was obtained from the published literature through a targeted literature review (Table 86). MSD acknowledge the scarcity of lung/heart transplants in current UK clinical practice.

Table 86 Probability of lung/heart transplant

Parameter	Value	Distribution (SE)	Source
Probability of lung/heart transplant, per 12-week cycle			
Low risk	0.00%	<u>Not varied</u>	Assumption ^a
Intermediate-low risk	0.00%	<u>Not varied</u>	Assumption ^a
Intermediate-high risk	0.05%	<u>Beta (0.01%)</u>	MSD data on file (2022). COMPERA (80)
High risk	0.07%	<u>Beta (0.01%)</u>	Merck data on file (2022). COMPERA (80)
Other transplant-related parameter			
Post-transplant mortality per 12-week cycle	2.8% ^b	Beta (0.6%)	Bernstein et al. (2018) (81)

^a This assumption was consistent with the 2022 treatment guideline that only patients with intermediate-high or high risk would be recommended for lung/heart transplant (4).

^b Reported as mortality rate of 12.41/100 person-years among adults with PAH, converted to 12-week probability
SE: standard error

Mortality

Patients with PAH face an increased risk of mortality, with increasingly higher risk as the disease progresses (4, 84). To capture the effect of disease progression on survival and

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

predict patient survival, all-cause mortality probabilities reported for a PAH cohort were applied to the proportion of patients in each health state to estimate the proportion of patients transitioning from each risk stratum health state to death. Data reporting PAH-specific mortality probabilities were not identified.

Data sources to inform risk status-specific survival were identified through a targeted review of the literature and decisions about their suitability were based on clinical expert opinion. Inclusion criteria were publications which reported Kaplan-Meier (KM) survival data per ESC/ERS risk status. Two eligible studies were identified (77, 79).

In the base case, KM curves for the no-comorbidity cohort in Rosenkranz et al. (2023) (79) were used, as the patient characteristics in this cohort were considered by clinical experts to be the more comparable to that of the STELLAR ITT population. Alternative KM curves from Boucly et al. (2022) (77) were used in scenario analyses.

Table 87 A Comparison of baseline patient characteristics of STELLAR and COMPERA

Parameter	STELLAR ITT (N=323)	COMPERA (Rosenkranz et al. 2023) ^a (N=208) (79)	French PAH registry (Boucly et al. 2022) (N=2879) (77)
Age, years			
Mean	47.9	NR	61.0
Median	48.0	51.0	NR
Female, n (%)	256 (79.3%)	141 (67.8%)	1737 (60.3%)
BMI, kg/m ²			
Mean	26.4	NR	27.2
Median	25.5	24	NR
WHO FC, n (%)			
I	NR	3 (1.5%)	925 (32.0%)
II	157 (48.6%)	41 (19.9%)	
III	166 (51.4)	135 (65.5%)	
IV	NR	27 (13.1%)	413 (4.0%)
6MWD, metres			
Mean	401.1	NR	NR
Median	424.0	373	300
Time from diagnosis to randomisation, years			
Mean	8.8	0	NR
Median	7.3	0	NR

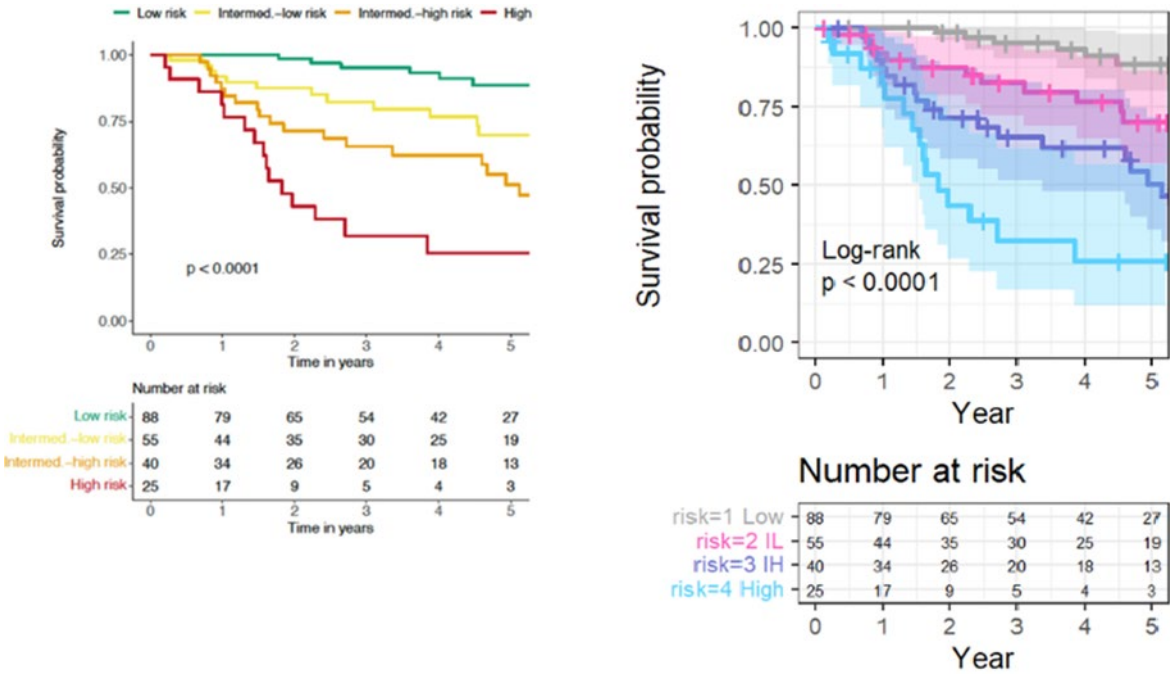
6MWD: 6-minute walking distance; BMI: body mass index; FC: functional class; ITT: intent-to-treat; NR: not reported; PAH: Pulmonary Arterial Hypertension; WHO: World Health Organization

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

^a Among all the reported subgroups in Rosenkranz et al. (2023) (79), data for the no-comorbidity cohort was used, as the baseline patient characteristics in this cohort were the most comparable to that of the STELLAR ITT population, especially in terms of relatively younger baseline age. Hence, the baseline patient characteristics of this cohort are presented above.

To inform long-term all-cause mortality probabilities by risk stratum, the Guyot (2012) (85) algorithm was employed to reconstruct individual patient-level data from the published KM survival curves from Rosenkranz et al. (2023) (79). Parametric survival functions were then fitted to re-constructed KM survival curves by risk status, assessed by the ESC/ERS 4-strata model, and used in the economic model (Figure 16).

Figure 16 Kaplan-Meier curves for survival from Rosenkranz et al. (2023) (79) (original – left; re-created patient-level data per the Guyot et al. algorithm (85) – right)



IH: intermediate-high; IL: intermediate-low

Previous PAH cost-effectiveness models identified in the literature (see Appendix E) were largely based on a WHO FC structure, and a commonly adopted approach has been to derive mortality probabilities by applying a hazard ratio relative to WHO FC I to estimate survival of patients in WHO FC II, III and IV states (43, 50, 54). Thenappan et al. (2010) (86) reported HRs of 4.51, 7.94 and 11.6 for WHO FC II, III and IV, respectively, relative to WHO FC I using a Cox proportional hazards model.

As the economic model built for this appraisal follows a risk-based approach rather than a FC-based one (in order to reflect current clinical practice), the findings of Thenappan et al. (2010) (86) cannot be directly applied, nevertheless the method utilised by the previous economic evaluations is applied in the base case approach.

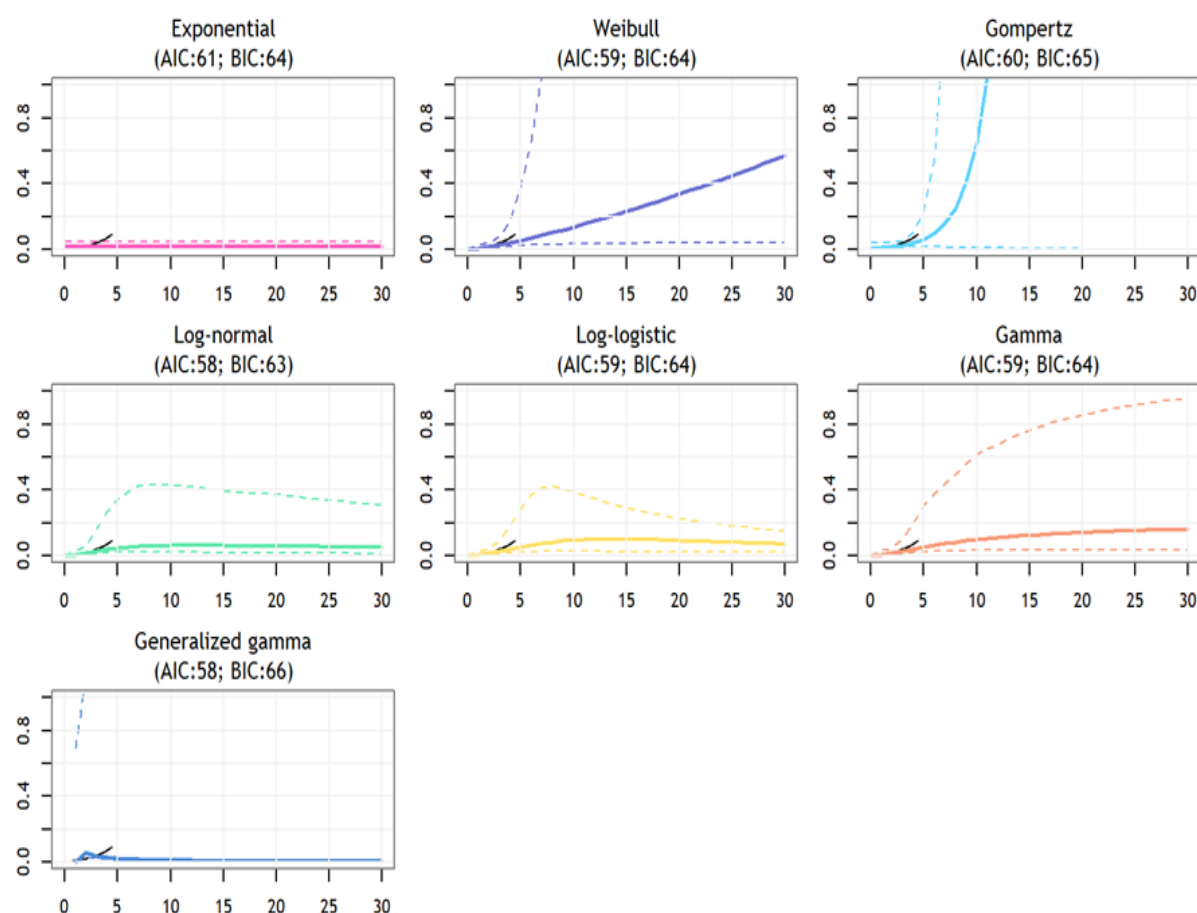
Base case: Single parametric model using low risk as reference state

The approach in this submission involved fitting parametric models to the KM curve for the low risk cohort to derive the mortality probabilities for the low risk health state and applying HRs relative to the low risk–specific mortality to derive the mortality for other risk strata health states. The low risk stratum is used as the reference risk stratum (to which HRs relative to other risk strata were applied) in this model based on precedent for WHO FC I as the reference state in the literature, but also because the low risk group had the highest number of patients at risk allowing for a higher event count (see Figure 16).

This approach allows for the adjustment and exploration of different survival assumptions as well as potential uncertainty of HRs across risk strata in a transparent manner, especially given the existence of literature that reports HRs across risk strata. For example, Boucly et al. (2022) reported HRs of 2.65, 5.65 and 11.08 for intermediate-low risk, intermediate-high and high, respectively, relative to low risk, using a Cox proportional hazards model.

In selecting an appropriate parametric model, it was proposed that an increasing hazard pattern over time would be clinically plausible and anticipated, and this was confirmed in discussions with clinical experts. Given the progressive nature of the disease, a pattern where the probability of death decreases over time is not considered valid in the case of Group 1 PAH patients treated with currently available treatments. Furthermore, the Rosenkranz study reported curves for patients according to their baseline risk status, without adjusting for subsequent risk status worsening, which would be expected to occur in PAH before death (79). This underscores the case for selecting models with an increasing hazard pattern as a patient's risk of death is expected to increase over time from the risk status with which they were categorised at baseline. Among all the fitted independent models for the low risk patients (Figure 17), only the independent Gompertz, Weibull and Gamma models for Rosenkranz et al. exhibited a clearly monotonically increasing hazard pattern and thus were deemed more clinically plausible from the hazard pattern perspective (79).

Figure 17 Fitted independent parametric regression models for low risk-specific survival from Rosenkranz et al. (2023)(79): hazard plots



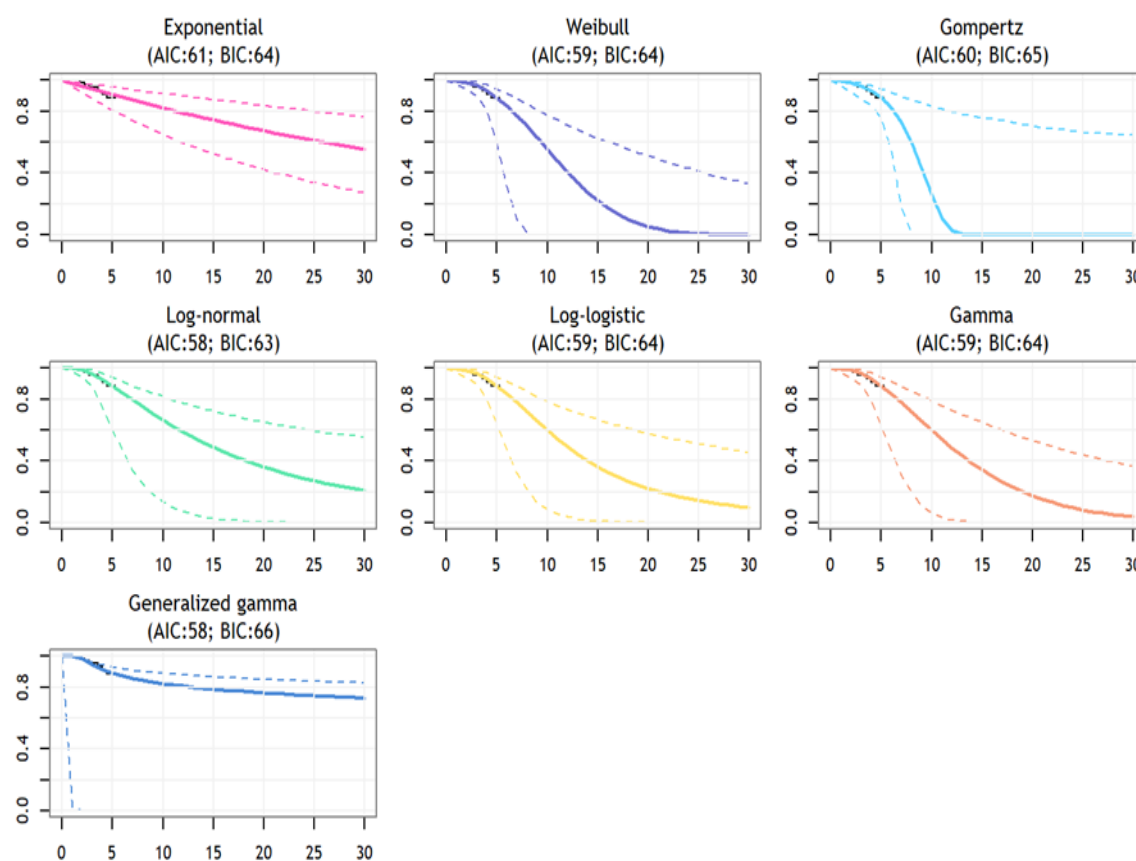
AIC: Akaike information criterion; BIC: Bayesian information criterion

(1) The x-axis refers to time in year since enrolment into COMPERA, whereas the y-axis refers to the cumulative proportion of patients alive

(2) The transition probabilities from each risk stratum health state to death at time point t was estimated as: survival function at time $t-1$ minus survival function at time t , divided by survival function at time $t-1$.

The translation of these hazard plots to survival curves are presented below. Given the small sample size (namely, small number of patients at risk over time from the KM curve), these models were associated with substantial uncertainty (as indicated by the 95% CI per the dashed lines) (Figure 18) and thus should be used and interpreted with caution.

Figure 18 Fitted independent parametric regression models for low risk-specific survival from Rosenkranz et al. (2023)(79): Survival



AIC: Akaike information criterion; BIC: Bayesian information criterion

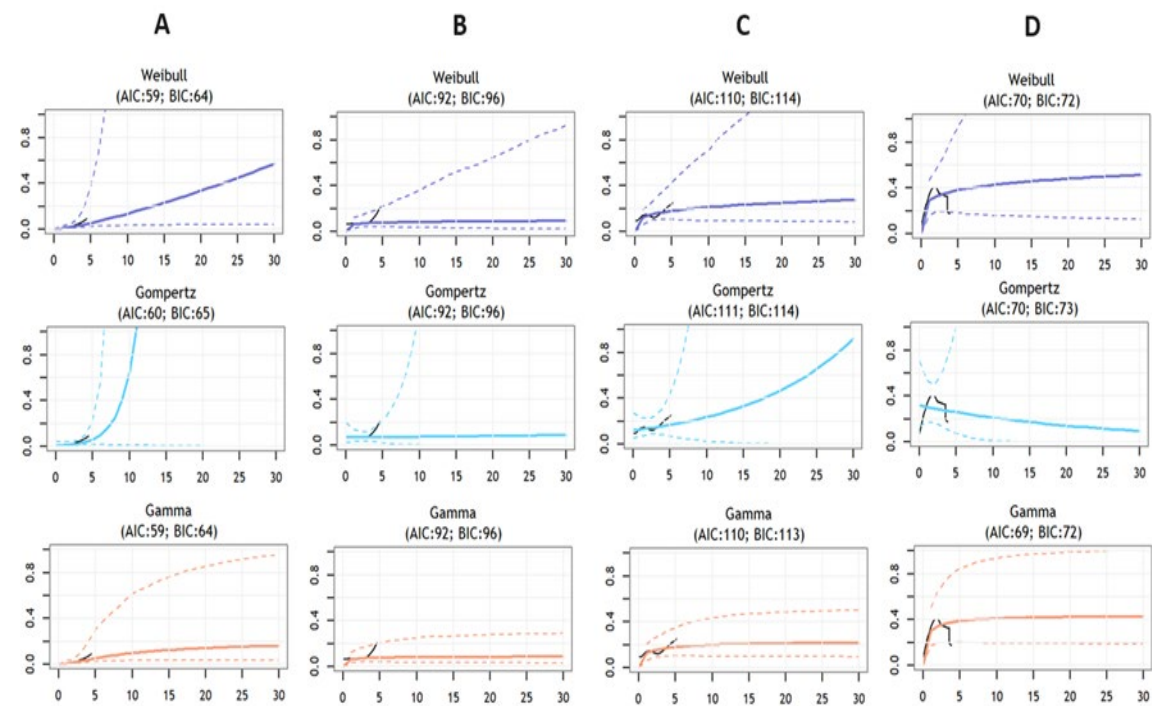
(1) The x-axis refers to time in year since enrolment into COMPERA, whereas the y-axis refers to the cumulative proportion of patients alive.

(2) The transition probabilities from each risk stratum health state to death at time point t was estimated as: survival function at time $t-1$ minus survival function at time t , divided by survival function at time $t-1$.

Upon further examination of the Gompertz, Weibull and Gamma hazard plots in comparison to the plots across all different risk categories (Figure 19), a number of conclusions were reached. The Gompertz and Weibull curves exhibit implausible outcomes from a clinical perspective. According to the definition, the probability of death is expected to be the lowest in the low risk group and the highest in the high risk group. However, in both cases, patients appear to have the highest risk of death in the low risk group, hence these curves are deemed clinically implausible. As a result, these options have been excluded from further analysis.

On the other hand, the Gamma curve demonstrates a more realistic probability of survival for the low risk population and had similar quality of fit statistics. Therefore, it has been selected as the reference curve for the low risk group.

Figure 19 Fitted independent parametric regression models for all risk-specific survival from Rosenkranz et al. (2023)(79): hazard plots of Weibull, Gompertz and Gamma



AIC: Akaike information criterion; BIC: Bayesian information criterion
A: low risk; B: intermediate-low risk; C: intermediate-high risk; D: high risk
The x-axis refers to time in year since enrolment into COMPERA, the y-axis refers to the hazards

In the next step, HRs for survival were estimated by applying a Cox proportional hazards model to the KM curves stratified by risk strata, with low risk as the reference strata, reported in Rosenkranz et al. (2023) (Table 88) (79). These values are used in the base case analysis.

Table 88 Model inputs: HRs for all-cause mortality

Parameter	Value	Lower bound of 95% CI	Upper bound of 95% CI	Distribution (SE) ^a	Source
Other risk strata relative to low risk					
Intermediate-low risk	■	■	■	■	Rosenkranz et al. (2023)(79)
Intermediate-high risk	■	■	■	■	
High risk	■	■	■	■	

CI: confidence interval; HR: hazard ratio; SE: standard error
^a The SE of each hazard ratio on the log scale was calculated by converting the lower bound and upper bound of the CI to the logarithmic scale, finding one-half of the distance between them, and dividing by the reference statistic z* used to calculate the upper bound

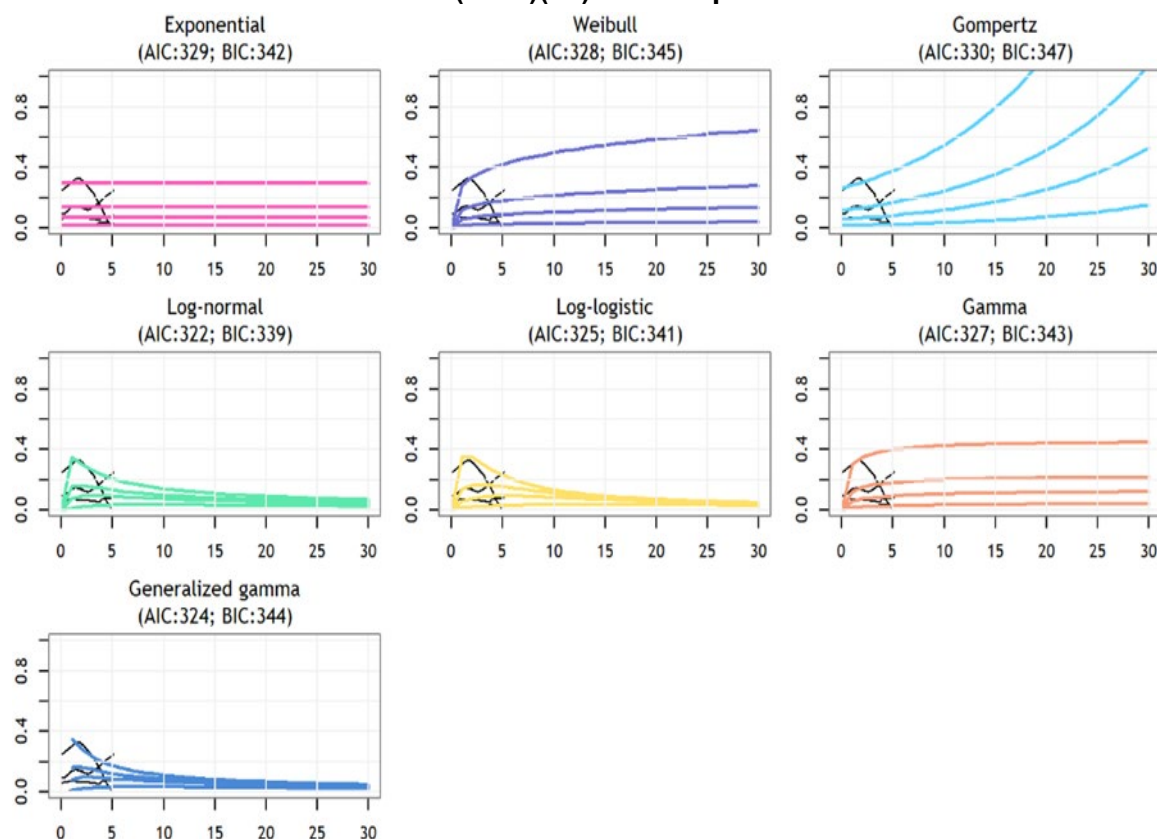
Scenario analysis: Dependent parametric models for all risk strata

Given the uncertainty around risk status-based mortality in PAH, an alternative approach was also tested in scenario analysis, where all available data were jointly fitted, using non-low risk

states as covariates. This approach maximises the utilisation of all the available data to incorporate all four KM curves (one for each risk stratum) in one parametric regression model.

As before, the selection of parametric survival models depended on clinical experts' opinions regarding long-term hazard pattern, following by assessment of goodness of fit (as assessed via fit statistics including AIC and BIC and visual inspection).

Figure 20 Fitted dependent parametric regression models for risk stratum-specific survival from Rosenkranz et al. (2023)(79): hazard plots

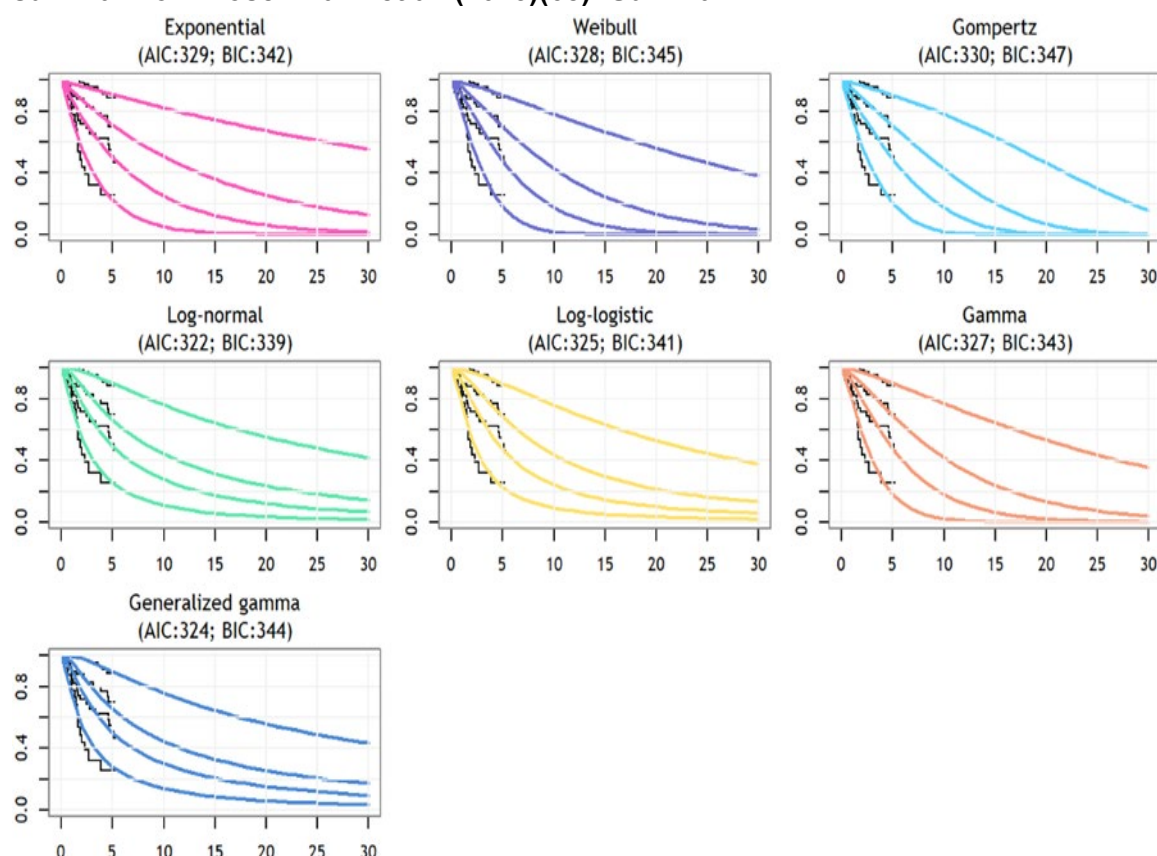


AIC: Akaike information criterion; BIC: Bayesian information criterion

- (1) Risk strata were included as covariates in each of the parametric models above.
- (2) The x-axis refers to time in year since enrolment into COMPERA, whereas the y-axis refers to the hazards.
- (3) The transition probabilities from each risk stratum health state to death at time point t was estimated as: survival function at time $t-1$ minus survival function at time t , divided by survival function at time $t-1$.

Among all the dependent models fitted to the KM curves from Rosenkranz et al. (Figure 21) (79), the dependent Gompertz and Weibull models (i.e., those with a clearly monotonically increasing hazard pattern) were considered by clinical experts to be more clinically plausible. The Gamma model, although displaying an increasing trend, was considered to behave more similarly to the exponential curve, showing hazards which are nearly constant or increasing at a slow pace for the majority of the time horizon, hence not aligning with clinical expectations. The Gompertz model was associated with higher mortality rates in the long-term compared to the Weibull model and thus led to more conservative survival outcomes for sotatercept. Thus, the Gompertz model was selected as more appropriate for inclusion in the scenario analysis. Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Figure 21 Fitted dependent parametric regression models for risk stratum-specific survival from Rosenkranz et al. (2023)(68): Survival



AIC: Akaike information criterion; BIC: Bayesian information criterion

(1) Risk strata were included as covariates in each of the parametric models above.

(2) The x-axis refers to time in year since enrolment into COMPERA, whereas the y-axis refers to the cumulative proportion of patients alive.

(3) The transition probabilities from each risk stratum health state to death at time point t was estimated as: survival function at time $t-1$ minus survival function at time t , divided by survival function at time $t-1$.

As mentioned previously, it should be noted that the KM curves (to which the aforementioned parametric models were fitted) were stratified by risk stratum at the first follow-up (i.e., at least 3 months after start of therapy) in the absence of longitudinal risk stratum data (e.g., KM curves for those who remained in specific risk strata) (79). The KM curve for the cohort of patients with intermediate-low risk at enrolment would include at least three sub-cohorts, including (1) patients remaining in intermediate-low risk, (2) patients transitioning from intermediate-low risk to low risk, and (3) patients transitioning from intermediate-low risk to intermediate-high risk. Individual KM curves for each of these sub-cohorts was not available from the study. Instead, the selection of parametric models, as discussed above, relied on clinical experts' views about long-term hazard patterns in patients with PAH, which, along with the exploration of multiple parametric models, may mitigate this limitation.

Approach not presented: Independent parametric models for all risk strata

A potential third approach to estimating mortality would involve deriving mortality probabilities by fitting independent regression models to each risk stratum separately.

The use of independent regression models generally has the advantage of accommodating different hazard patterns across patient cohorts. These models work well when the available data are considered mature and the number of events are high enough that the confidence intervals at the end of the extrapolations are reasonably narrow for decision-making. However, in this case, we observe relatively smaller sample size for one or more patient categories, uncertainty of the tail of KM curves, and clinically counter-intuitive (and even implausible) hazard patterns that are difficult to interpret. An example of these counter-intuitive results was presented in Figure 19, which ultimately led to the conclusion that the third approach is not appropriate to be used for decision-making.

Age adjustment

As reported in Table 85, the median age of the no-comorbidity cohort reported by Rosenkranz et al. was 51.7 years (mean value not reported) (79). The baseline starting mean age for the model cohort was reflective of the STELLAR population (mean 47.9 years) and given that is lower than the starting point of the mortality curve, an adjustment was made which meant the model cohort did not assume the hazard of a patient older than 51 years until the model cohort reached that age. Until the cohort reached the selected age (51.7 years) constant mortality rates were considered equivalent to the starting mortality in the source (Rosenkranz).

This adjustment was made so as not to overstate the hazard of the model cohort when the average patient is younger than the population from which mortality was informed, given the important impact of age on mortality. Hoepfer et al. highlighted the existence of age-related survival differences (87). This is further supported by the findings of the COMPERA registry, which reported progressively worsening survival rates in older versus younger patients over time, with significantly lower 1-, 2-, and 3-year survival rates in patients aged over 65 years, even when adjusted for age and sex (87, 88).

The age cap was not included in the sensitivity analysis as arbitrarily adjusting it could introduce unrealistic scenarios that undermine the intended purpose of this constraint. Therefore, we have opted to incorporate the cohort age in our sensitivity analyses, ensuring that the patient age distribution around the mean is accurately represented.

Conclusion

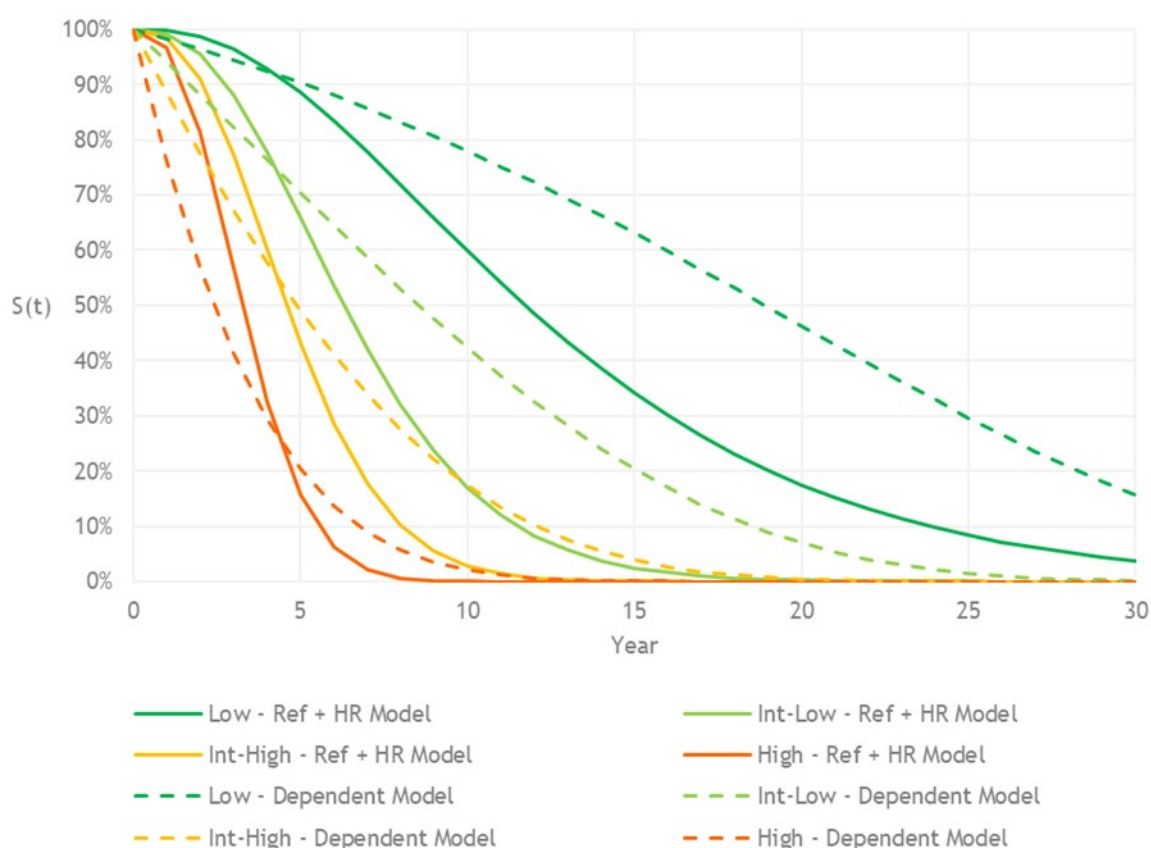
Both the HR-based, and dependent models were deemed suitable for survival extrapolation based on their characteristics presented above. The decision for selecting a base case

approach was ultimately considered based on the clinical validity of the projected curves as well as the resulting LYs the model estimated when using one or the other.

The dependent model consistently predicted longer survival compared to the HR-based model. In the case of low and intermediate-low groups, approximate median survival of 18 and 8 years versus 12 and 6 years (dependent versus HR-based models, respectively) were estimated, with a sizeable proportion of the patients still alive at the end of the time horizon according to the dependent model (general population unadjusted). 7-8% of the patients who were members of the intermediate-low risk group at baseline were predicted to be alive at 20 years into the time horizon based on the dependent model, while a similar proportion of patients were alive at 12 years with the HR-based model.

Patients in the intermediate-high and high risk groups were also predicted to live longer according to the dependent model however with a lesser absolute difference. The HR-based model estimated a higher proportion of patients to be alive within the first 5 years of the projections with a steeper decline towards the end, when compared to the dependent model.

Figure 22 Comparison of HR-based and Dependent Model Projections



HR: hazard ratio

Considering the transitions of the cohort between health states, with sotatercept predominantly associated with the low risk group and selexipag with the high risk group, the dependent model tends to favour sotatercept due to longer survival in healthier states. In contrast, the HR-based model would favour selexipag, reflecting comparable or superior survival (vs. dependent model predictions for the same risk group) in worse health states. It should be underscored that the longer the survival period, the greater the uncertainty in the predictions, thereby complicating the assessment of their accuracy. This factor particularly complicates the evaluation of survival in patients receiving sotatercept when using the dependent model.

Based on this, results for the HR-based model were deemed more clinically plausible therefore it was decided that the HR-based model is more appropriate in the base case and the dependent model is evaluated in scenario analysis.

In the subsequent phase, age adjustment was integrated into the computations using the HR-based model and Gamma distribution. This led to the following survival estimates: proportions of patients alive in each risk group from the end of the first cycle to 20 years, a point at which the majority of patients have passed away.

General mortality

The model did not directly rely on general mortality (per the UK life tables (89)) as model inputs. However, when the modelled all-cause mortality (from the fitted parametric regression models) was lower than general mortality of the general UK population of equivalent age and sex ratio, then general mortality was used in the model to inform the mortality inputs at that cycle, in order to avoid clinically implausible outcomes for this population.

PAH hospitalisation

Patients with PAH often experience hospitalisations due to the disease, with higher risk patients facing a greater risk of PAH hospitalisation. The model therefore accounts for the impact of PAH hospitalisations on costs and utility. This was achieved by applying health state-specific PAH hospitalisation probabilities to all non-death health states during each model cycle (Table 89).

As mentioned in the clinical section, sotatercept demonstrated a profound reduction in risk of hospitalisation compared to patients treated with placebo and background therapy. A similar benefit of sotatercept was observed versus selexipag in an ITC conducted between STELLAR and GRIPHON, using WHO FC based endpoints (details are presented in Appendix P). It was assumed that the benefit observed in the ITC would be observed in any patient groups regardless of how patients were categorised according to their mortality risk. The hazard ratio

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

of [REDACTED] was applied to the probabilities of hospitalisation in the low risk and intermediate-low risk health states based on the aforementioned ITC in the base case. Due to the uncertainty associated with the ITC discussed in Appendix P, a scenario analysis explores the impact of removing this HR. It should be noted that due to the overall low impact of hospitalisation events on the total costs the influence of this assumption is minor. Nevertheless, it underlines the beneficial efficacy and safety profile of sotatercept compared to selexipag.

Table 89 PAH hospitalisation

Parameter	Value	Distribution (SE)	Source
Probability of PAH hospitalisation, per 12-week cycle			
Low risk	[REDACTED]	[REDACTED]	MSD data on file (2023). COMPERA (82)
Intermediate-low risk	[REDACTED]	[REDACTED]	
Intermediate-high risk	[REDACTED]	[REDACTED]	
High risk	[REDACTED]	[REDACTED]	
Post lung/heart transplant	[REDACTED]	[REDACTED]	Assumption ^a
HR for PAH hospitalisation (sotatercept + BGT vs. selexipag + BGT)			
Low risk	[REDACTED]	[REDACTED]	MSD data on file (2023). ITC of overall STELLAR vs. GRIPHON (90)
Intermediate-low risk	[REDACTED]	[REDACTED]	
Intermediate-high risk	[REDACTED]	[REDACTED]	Assumption
High risk	[REDACTED]	[REDACTED]	Assumption

BGT = background therapy; HR: hazard ratio; PAH: pulmonary arterial hypertension; SE: standard error

* The input for the “post lung/heart transplant” state was assumed to be the same as that of the low risk state due to the lack of transplant-related data in STELLAR and the literature in general, primarily driven by the rare nature of transplant in patients with PAH.

** The standard error for log-transformed HR is presented.

3.4 Measurement and valuation of health effects

3.4.1 Health-related quality-of-life data from clinical trials

The base case cost-effectiveness analysis uses EQ-5D-5L utility data reported in the STELLAR clinical trial to calculate utility by health state (i.e., risk status). The trial endpoints WHO FC, 6MWD and NT-proBNP were used to classify the STELLAR trial participants into the four-stratum risk assessment categories used to define the model health states.

Several studies were identified (see section 3.1) that used the FC categorisation when sorting patients into health states, and only one was identified which derived health state utilities according to risk status (91). However data from this study were deemed inappropriate for the model, owing to the low sample size (n=6 for the low risk population, n=4 for high risk), higher mean age compared to the STELLAR population, and implausibly high utility values, including a higher value for intermediate risk than for high risk (91). The only relevant evidence for this model structure therefore is the quality-of-life data collected in the STELLAR study, which is

used in the base case analysis. The impact of using WHO FC health state utilities as a proxy for risk status is explored in a scenario analysis.

EuroQol 5-Dimension (EQ-5D) is a standardised measure of health status developed by the EuroQol Group to provide a simple, generic measure of health for clinical and economic appraisal. The EQ-5D measure in the STELLAR study consisted of a descriptive system of five questionnaire items (EQ-5D-5L) as well as a visual analogue scale (EQ-5D VAS). The EuroQol 5-Dimension 5-Level (EQ-5D-5L) descriptive system was comprised of the following five dimensions, each describing a different aspect of health: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension had five response levels of severity rated as, no problems (level 1), slight, moderate, severe and extreme problems (level 5).

The utility data recorded for the Full Analyses Set (FAS) population, defined as all randomised participants, was used to generate all reported EQ-5D Utility Scores summaries. The utility analyses only included patients from the FAS population with non-missing EQ-5D Score assessments at baseline and week 24. The utility data at baseline and post-baseline visits were combined in the analysis used to inform the model base case utility.

3.4.2 Mapping

The EQ-5D-5L responses collected in STELLAR are mapped to EuroQol 5-Dimension 3-Level (EQ-5D-3L) using a published UK crosswalk algorithm incorporating participant's age and sex with probabilities to compute the scores based on the EQ-5D-5L responses (92). EQ-5D-3L is specified in the NICE Methods Manual for technology appraisal submissions (93).

3.4.3 Health-related quality-of-life studies

An SLR was conducted to identify studies published since 2012 assessing utilities and HRQoL of adult patients with PAH. The date limit of 2012 was chosen to ensure capture of the most recent relevant evidence, and studies published prior to this date are deemed to no longer reflect current treatment practices. Details of the methods used to identify and select the relevant studies are described in Appendix F.

The SLR identified 36 publications. Of these, 16 publications reported utility indices, while 25 publications focused on HRQoL.

NICE has not previously evaluated a technology for PAH (although other HTA bodies have done so, including the All-Wales Medicines Strategy Group [AWMSG] and Scottish Medicines Consortium [SMC] in the UK) (section 3.1).

Utilities

All 16 utility studies reported outcomes using the EQ-5D utility index or VAS (15, 91, 94-107). Among the three studies that included PAH patients receiving treatment, all three reported improvements in the EQ-5D-3L and EQ-5D-5L utility indices following treatment (97, 98, 102). Comorbidities in patients with PAH as well as higher WHO FC categories were associated with worse EQ-5D index scores (97, 99, 104-106). Of note, patients with PAH, even those with the most favourable EQ-5D scores, had significantly lower EQ-5D utility values compared to the population norms for the UK (UK norms: mean EQ-5D: ~0.86; mean EQ-5D VAS: ~82.8) (108).

Data from RCTs in PAH were reviewed separately and are summarised in Appendix B. Only two studies included in the clinically focused SLR reported utility index scores: PATENT-1 and PATENT-2 examined the effect of riociguat treatment on EQ-5D index scores (109, 110).

Health-related quality-of-life

Of the 25 identified studies reporting HRQoL data, nine reported data from within the UK and are therefore directly relevant for the England and Wales setting. The remaining 16 studies reported data from within the EU27. The instruments reported across all 25 studies included the modified Borg Dyspnoea scale, CAMPHOR, emPHasis10, 36-item Short-Form Survey (SF-36) and Minnesota living with heart failure questionnaire (MLFHQ). Overall, there was evidence that PAH therapies improve HRQoL over time. In addition, both the CAMPHOR and emPHasis10 data indicated that patients with a higher WHO FC may experience worse HRQoL (94-96, 106, 107, 111-122).

3.4.4 Adverse reactions

This economic evaluation does not account for AEs given the comparable safety outcomes between the sotatercept + background therapy and placebo + background therapy arms in STELLAR (severe AE: 8.0% vs. 13.1%, absolute difference: -5.1% [95% CI: -12.2 to 1.6]; serious AE: 14.1% vs. 22.5%, absolute difference: -8.4% [95% CI: -16.9 to 0.1]), indicating that the addition of sotatercept to background therapy does not increase the incidence of AEs (see section 2.11 for further details).

The most common AEs of interest in the STELLAR clinical trial included epistaxis (22.1% of all participants), increased haemoglobin (14.3% of all participants), and telangiectasia (17.4% of all participants), all of which require minimum resources to address (see Section 2.11 for details on AE rates).

According to clinical expert opinion, the sotatercept AEs which may necessitate clinical intervention are bleeding events. However, none of the events observed in the clinical trial

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

were deemed to be severe and over half of the bleeding events reported were categorised as epistaxis (61%), resulting in low resource requirements for their resolution and minimal impact on a patient’s QoL. Although a higher number of AEs were observed with selexipag treatment in the GRIPHON trial (90), the severity of these events was not published, leading to their exclusion from further analysis. This constitutes a conservative model assumption, as discussion with clinical experts highlighted the many AEs that can be experienced with selexipag in real-world practice, particularly during the titration phase.

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

Health state utilities

The descriptive summaries of number of participants, mean, standard deviation, median, first quartile, third quartile, minimum and maximum) for the cross-walked UK utility scores by risk status are presented in Table 90. Due to the limited EQ-5D-5L data that could be collected according to the STELLAR protocol, descriptive analyses were deemed sufficient for informing the model’s health state values.

It is assumed patients in the post-transplant state have the same utility as low risk patients. This assumption was made due to lack on quality-of-life data in the target population and it should be added that transplantation is the last resort of patients with PAH uncontrollable by other means. Sotatercept is expected to delay time to transplantation via improved disease control and keeping patients in the lower risk health states compared to selexipag. Assuming that more patients receive transplantation with selexipag favours the comparator by assuming a higher quality-of-life for patients who have undergone transplantation. Advice from UK clinical experts is that lung/heart transplants in PAH patients occur very rarely in current clinical practice and the National PAH Audit reports that just 6 transplants took place in the most recently reported year, so the impact of this assumption for the decision problem is low (17).

Table 90 Utility by health state UK crosswalk (from EQ-5D-5L to EQ-5D-3L) mapping algorithm utility scores: pooled records (baseline and Week 24)

	N	Mean (SD)	Median (Q1; Q3)	Min; Max
Low risk	█	█	█	█
Intermediate-low risk	█	█	█	█
Intermediate-high risk	█	█	█	█
High risk	█	█	█	█

EQ-5D-3L: EuroQol 5-Dimension 3-Level; EQ-5D-5L: EuroQol 5-Dimension 5-Level; Q1: 1st quartile; Q3: 3rd quartile; UK: United Kingdom
Source: Hernandez Alava et al. 2023 (123)

As discussed in Section 3.4.1, only one study was identified from the literature reporting risk-based utilities scores in patients with PAH however the values lacked face validity and therefore MSD believe the STELLAR trial utility scores to be the most relevant for decision-making (91). To enable the impact of health state utility values on the model results to be explored, a scenario analysis is provided using the WHO FC-based utility scores from the STELLAR trial FAS population as a proxy for risk status. The utility scores based on the UK crosswalk are pooled across both treatment arms as well as baseline and week 24, presented in Table 91 below (123).

Table 91 Scenario analysis: health state utilities based on WHO Functional Class as a proxy for risk status (123)

	N	Mean (SD)	Median (Q1; Q3)	Min; Max
WHO FC I	█	██████	██████	██████
WHO FC II	█	██████	██████	██████
WHO FC III	█	██████	██████	██████
WHO FC IV	█	██████	██████	██████

FC: Functional Class; Q1: 1st quartile; Q3: 3rd quartile; SD: standard deviation; WHO: World Health Organisation

Administration disutility

Additionally, patients may accrue a disutility while receiving treatment associated with the administration method, as detailed in Table 92. This is based on the study by Davies et al, identified by the SLR (Section 3.4.3) that created four “health states” linked to administration route (oral, inhaled, subcutaneous, or intravenous administration) based on interviews with six PAH patients that had different administration methods for ERA/PDE5i treatment (102); the UK general public completed EQ-5D-5L surveys based on these health states to generate the utility associated with each state and therefore allow for the derivation of the disutility with respect to oral administration. This study has the dual benefits of using health state descriptions informed by the lived experience of patients with the disease receiving the treatment intravenously and then eliciting values directly from members of the UK general population, reflecting their preferences. Based on patient testimony and clinical expert advice, the results are deemed to be intuitive, as the poor QoL experienced by patients receiving IV PCA has been described (see section 1.3.1). The QoL of the SC injection route is assumed to be equivalent to an oral treatment, given the training and support that will be delivered alongside sotatercept initiation.

Table 92 Disutility for administration method

Administration method	Disutility	SE	Reference/Justification
SC injection	0	0	Assumption
IV infusion	-0.307	0.0614	Davies 2018(102)
Inhaled	0	0	Assumption
Oral	0	0	Assumption

IV: intravenous; SC: subcutaneous; SE: standard error

Clinical event disutility

The impact of QoL decrease associated with PAH hospitalisation was also incorporated into the analysis as a one-off decrement. Disutility data for PAH patients who have been hospitalised were not identified via the systematic review, although a targeted literature review identified a systematic review of utility values in heart failure, which provided an overview of disutility values attributable to a hospitalisation event (124). Given the data gap in PAH, this was proposed to be a suitable proxy. Amongst the studies, the value for “hospitalised within previous 30 days” reported by McMurray et. al (2018) (125) was deemed to be the most appropriate, based on the assumption that the QoL impact is felt most accurately as close to the date of the event as possible and because the study employed the UK tariff. Disutility values for subsequent model cycles following hospitalisation have also been reported for heart failure patients in the literature, however it is assumed using a 12-week cycle length in the current model sufficiently addresses that, so a one-cycle disutility is applied.

A QoL decrement equal to the impact of hospitalisation was assumed for receiving lung/heart transplant due to the lack of data in the target population. This is believed to be a conservative assumption as patients undergoing a major surgery with a long recovery period (see Table 96 for average HCRU of transplant patients) would likely experience a higher utility drop than in the average hospitalisation, therefore assuming a more moderate disutility favours the selexipag arm as a greater proportion of patients in this arm transition to the post-transplant health state.

Table 93 Disutility for clinical event

Event	Disutility	SE	Reference/ Justification
PAH hospitalisation	0.105	Normal (0.021)	MCMurray et al. (2018) (125)
Lung/heart transplant	0.105	Normal (0.021)	Assumption ^a

PAH: pulmonary arterial hypertension; SE: standard error

^a Assumed to be the same as that for PAH hospitalisation due to the lack of transplant-related data in STELLAR and the literature in general, primarily driven by the rare nature of transplant in patients with PAH.

No disutility for sotetercept AEs was included as the model assumes no treatment-related adverse events occur that would result in substantial disutility to the patient (see Section 2.11.1). Clinical expert opinion also views the safety profile of sotatercept to be mild.

3.4.6 Utility adjustment based on age

Utility values were adjusted for age using a multiplicative approach. This involved dividing the utility value for the current age by the utility value at the age of model entry (which corresponds to the average age of the sample used to inform health state utility values), and then multiplying this adjustment factor by the health state utility value. Consequently, patient utility values will always be equal to or lower than the health state utility value (assuming the current age is the same as or younger than the age at entry), with gradual decreases over time reflecting reductions in age-adjusted utility values. Inputs were derived from Hernández Alava et al 2022 (92).

An alternative approach for utility age adjustment is to cap the health state utility values at each cycle if they exceed the utility of the general population. Over the 30-year time horizon of the economic model, only patients in the low risk health state have the potential to exceed this cap. While this is a simpler methodology that ensures utility values at each cycle do not exceed that of the general population, modelled patient utility values for each health state do not reflect the increasing age of the modelled population. This approach is presented as a scenario analysis to explore the impact on model results.

3.4.7 Carer quality of life

Patients with PAH can have difficulty performing everyday tasks, especially as the disease progresses resulting in progressive heart failure and other comorbidities (4). Patients' disease is expected to impact their relationships with and QoL of their caregivers, as reported by the PHA UK survey (19) (details in Section 1.3.1). To capture the burden of progressive disease on caregivers, the model applies a caregiver utility impact based on the patient's health state (see Table 94 for details). This relationship has been reported by Pennington et al. (2024) (24), identified by a targeted literature review

The model applies a approach to capturing caregiver QALY impacts which avoids the "caregiver QALY trap," a phenomenon where a treatment that extends patient life penalises caregiver QALYs since the longer the patient remains alive, the longer the caregiver experiences a disutility (126). To overcome this, the CEM uses high risk as the reference stratum and applies a "caregiver benefit" in which caregivers of patients in lower risk health states receive a utility gain, with the high risk patient's caregiver receiving no utility benefit. This approach rewards maintaining patients in lower risk health states.

The caregiver benefit was derived from Pennington et al. (2024) (24), that reported a fixed effects model for carers' SF-6D index score. Specifically, the fixed effect model for the SF-6D score of carers that were family members utilises an interaction coefficient of 0.123 for care recipient's SF-6D score (i.e. improving the HRQoL of a patient by 0.1 would improve the HRQoL of the carer by 0.0123); This coefficient was used in combination with patient health state utility in the model to approximate a utility gain for caregivers by:

- Determining the difference between each state value versus the high risk state value to get a patient utility gain per state.
- Multiplying the patient gain by the coefficient reported in Pennington et al. (2024) (24) to calculate the caregiver utility gain per health state.

Table 94 Disutility for caregivers

Health State	Patient Utility	Patient Utility Increment (Ref: High Risk)	Caregiver Utility Increment (Ref: High Risk)	SE	Reference
Low Risk	■	0.292	0.036	-0.007	Calculation based on Pennington (2024) (24).
Intermediate-Low Risk	■	0.185	0.023	-0.005	Calculation based on Pennington (2024) (24).
Intermediate-High Risk	■	0.104	0.013	-0.003	Calculation based on Pennington (2024) (24).
High Risk	■	NA	NA	NA	NA
Post lung/heart transplant	■	0.292	0.036	-0.007	Calculation based on Pennington (2024) (24).

NA: not applicable; Ref: reference; SE: standard error

There are some limitations to the current caregiver utility method. The Pennington et al. (2024) (24) publication looked at any caregiver, and followed a disease-agnostic approach, so it is unclear how generalisable this model would be for caregivers of PAH patients. Additionally, the model uses SF-6D as a measure of patient and carer QoL, hence there are potential uncertainties when applying this model to patient utility values collected using EQ-5D (the utility measurement used in the STELLAR clinical trial). However, these limitations are necessitated by the lack of other data sources related to the utility of caregivers of PAH patients. To acknowledge the uncertainty about the quantitative impact of PAH on carer QoL, a scenario is provided whereby carer utility is not accounted for in the model.

3.5 Cost and healthcare resource use identification, measurement and valuation

An SLR was conducted to identify HCRU and cost studies reporting on PAH populations relevant to the economic model since 2012. The date limit of 2012 was chosen to ensure capture of the most recent relevant evidence, and studies published prior to this date are deemed to no longer reflect current treatment practices. The specific methods used to identify and select relevant studies are described in Appendix G.

The SLR identified 91 publications, of which four reported PAH-related costs and HCRU among the English and/or Welsh population. One study reported both cost and HCRU outcomes (21), one study reported cost data only (127) and two studies reported HCRU data only (Table 95 and Table 96) (113, 128).

Across the four identified studies, costs and HCRU are reported for the PAH population as a whole, without stratification for WHO FC or risk status. No studies were identified that reported on costs/HCRU in PAH stratified by patients' WHO FC or risk status.

Details of the costs and HCRU reported in the four studies identified in the SLR are described in Appendix G.

Table 95 Abbreviated summary of costs relevant to UK decision-making

Author (Year), Country	Currency, Cost year	Cost type	Population	Total annual cost
Exposto (2021)(21), UK	GPB, 2017	Total cost for patients with PAH over a 5-year period	PAH patients	£43.2 M
		Total cost of inpatient admissions for patients with PAH (2013-2017)	PAH patients	£33.9 M
		Total cost of outpatient admissions for patients with PAH (2013-2017)	PAH patients	£8.3 M
		Total cost of A and E visits for patients with PAH (2013-2017)	PAH patients	£0.9 M
		Cost of PAH-related hospitalisation, per patient (2013-2017)	PAH patients	£1,544
Montpart Viñas (2022)(127), Belgium, France, Germany, Italy, Poland, Portugal, Spain, the Netherlands, UK	EUR, 2021	PAH-related hospitalisation costs	PAH patients on low dose treprostinil	€1,507
			PAH patients on medium dose treprostinil	€2,110

			PAH patients on high dose treprostinil	€4,571
			PAH patients on high dose treprostinil	€4,571

A and E: Accident and Emergency; EUR: Euros; GBP: Great British Pounds; M: million; NR: not reported; PAH: pulmonary arterial hypertension, UK: United Kingdom

Table 96 Abbreviated summary of HCRU relevant to UK decision-making

Author (Year), Country	Population	HCRU type	Mean (SD) PPPA use	Mean (SD) LOS per patient/ range where reported
Exposto (2021)(21), UK	PAH patients	PAH patients with any inpatient admissions (2017)	2.9 (4.3)	3.3 (7.6)
		Patients with PAH-related inpatient admissions (2017)	2.0 (3.1)	3.8 (8.1)
		PAH patients with any outpatient visits (2017)	10.0 (7.4)	NA
		PAH patients with any A and E visits (2017)	0.8 (1.0)	NR
		Patients with PAH-related A and E visits (2017)	0.3 (0.9)	NR
Sabashnikov (2019)(128), UK	PAH patients post lung transplant	Postoperative hospital LOS (days)	65.0	26.0-125.0
		Postoperative ICU LOS (days)	28.0	12.0-40.0
		Ventilator use (hours)	97.0	36.0-84.0
		Postoperative ECMO use	NR	NR
Small (2014)(113), USA Germany, Italy, UK	Number of PAH-related hospitalisations	PAH patients receiving monotherapy	NR	NA
		PAH patients receiving combination therapy without intravenous infusion of prostacyclin	NR	NA
		PAH patients receiving combination therapy with intravenous infusion of prostacyclin	NR	NA

A and E: Accident and Emergency; ECMO: extracorporeal membrane oxygenation; HCRU: health care resource utilisation; ICU: intensive care unit; LOS: length of stay; NA: not applicable; NR: not reported; PAH: pulmonary arterial hypertension; PPPA: per patient per annum; SD: standard deviation; UK: United Kingdom; USA: United States

3.5.1 Intervention and comparators' costs and resource use

The model explicitly captures all treatment costs associated with a patient's treatment regimen. Details of each treatment, including the route of administration and assumed dosing regimen are presented in Table 97.

Table 97 Treatment options

Drug Class	Drug	Administration Method	Dosing ^a	Source
ActRIIA-IgG1	Sotatercept	SC injection	0.3 mg/kg every 21 days (loading dose) 0.7 mg/kg every 21 days	STELLAR trial protocol
ERA	Bosentan	Oral	125 mg twice daily	Tracleer [prescribing information] 2022
	Ambrisentan	Oral	10 mg once daily	Letairis [prescribing information] 2020
	Macitentan	Oral	10 mg once daily	Opsumit [prescribing information] 2021
PDE5i	Sildenafil citrate	Oral	20 mg three times daily	Revatio [prescribing information] 2020
	Tadalafil	Oral	40 mg once daily	Adcirca [prescribing information] 2021
PGI2 analogue	Epoprostenol	IV infusion	2 ng/kg/min	Flolan [prescribing information] 2021; CADTH submission for selexipag ^b
	Treprostinil	IV infusion	1.25 ng/kg/min	Remodulin [prescribing information] 2022; CADTH submission for selexipag ^b
	Iloprost	Inhaled	5 mcg 6-9 times daily	Ventavis [prescribing information] 2022
PGI2 agonist	Selexipag	Oral	1.6 mg twice daily	Upravi [prescribing information] 2021

ActRIIA-IgG1: activin receptor type IIA-Fc; CADTH: Canadian Agency for Drugs and Technologies in Health; ERA: endothelin receptor antagonist; IV: intravenous; PDE5i: phosphodiesterase type 5 inhibitor; PGI2: prostaglandin I2; SC: subcutaneous; sGC: soluble guanylate cyclase

^a Dosage information was based on target dose.

^b Model assumes a daily dose which corresponds to the midpoint of the range reported in a prior CADTH submission in PAH(54).

In the base case, the model compares sotatercept with dual (ERA+PDE5i) background therapy versus selexipag with dual (ERA+PDE5i) background therapy. The composition of dual background therapy is not expected to differ between arms, hence the distribution of individual ERA and PDE5i treatments which comprise the background therapy is assumed to be equal for the sotatercept and selexipag treatment arms.

The distribution of treatments at the beginning of the simulation are presented in Table 98.

Table 98 Treatment distribution by health state and treatment arm

Drug	Low risk	Intermediate-low risk	Intermediate-high risk	High risk
Sotatercept + dual background therapy				

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Sotatercept	100.00%	100.00%	100.00%	100.00%
Bosentan	12.76%	12.76%	12.76%	12.76%
Ambrisentan	43.29%	43.29%	43.29%	43.29%
Macitentan	43.95%	43.95%	43.95%	43.95%
Sildenafil citrate	61.87%	61.87%	61.87%	61.87%
Tadalafil	38.13%	38.13%	38.13%	38.13%
Epoprostenol	0.00%	0.00%	0.00%	61.94%
Treprostinil	0.00%	0.00%	0.00%	4.95%
Iloprost	0.00%	0.00%	0.00%	33.11%
Selexipag	0.00%	0.00%	0.00%	0.00%
Selexipag + dual background therapy				
Sotatercept	0.00%	0.00%	0.00%	0.00%
Bosentan	12.76%	12.76%	12.76%	12.76%
Ambrisentan	43.29%	43.29%	43.29%	43.29%
Macitentan	43.95%	43.95%	43.95%	43.95%
Sildenafil citrate	61.87%	61.87%	61.87%	61.87%
Tadalafil	38.13%	38.13%	38.13%	38.13%
Epoprostenol	0.00%	0.00%	61.94%	61.94%
Treprostinil	0.00%	0.00%	4.95%	4.95%
Iloprost	0.00%	0.00%	33.11%	33.11%
Selexipag	100.00%	100.00%	0.00%	0.00%
Discontinued sotatercept ('off-treatment')				
Sotatercept	0.00%	0.00%	0.00%	0.00%
Bosentan	12.76%	12.76%	12.76%	12.76%
Ambrisentan	43.29%	43.29%	43.29%	43.29%
Macitentan	43.95%	43.95%	43.95%	43.95%
Sildenafil citrate	61.87%	61.87%	61.87%	61.87%
Tadalafil	38.13%	38.13%	38.13%	38.13%
Epoprostenol	0.00%	0.00%	61.94%	61.94%
Treprostinil	0.00%	0.00%	4.95%	4.95%
Iloprost	0.00%	0.00%	33.11%	33.11%
Selexipag	0.00%	0.00%	0.00%	0.00%

IV: intravenous; SC: subcutaneous

* Data from National Audit of Pulmonary Hypertension Great Britain 2023-2024(129)

** Patients who discontinued sotatercept or selexipag.

The costs associated with each treatment include acquisition costs, calculated from the dosing information in Table 97 and the package costs detailed in Table 99, and administration costs, based on the administration method specified in Table 97 and the unit administration costs specified in Table 99.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Table 99 Treatment package costs

Drug	Strength per tablet or vial	Units / package	Cost / package	Reference
Sotatercept, SC injection*	45mg	1	£5,422.50	MSD UK
	45mg	2	£10,845.00	
	60mg	1	£7,230.00	
	60mg	2	£14,460.00	
Bosentan, Oral	125	56	£42.09	eMIT (July 2023 to June 2024)
Ambrisentan, Oral	10	30	£34.59	eMIT (July 2023 to June 2024)
Macitentan, Oral	10	30	£2,306.00	BNF (accessed 21 st February 2025) -- Opsumit 10mg tablets Janssen-Cilag Ltd
Sildenafil citrate, Oral	20	90	£40.79	eMIT (July 2023 to June 2024)
Tadalafil, Oral	20	56	£18.39	eMIT (July 2023 to June 2024)
Epoprostenol, IV infusion	1.5	1	£44.76	BNF (accessed 21 st February 2025) -- Flolan 1.5mg powder and solvent (pH12) for solution for infusion vials GlaxoSmithKline UK Ltd
Treprostinil, IV infusion	200	1	£2,811.40	eMIT (July 2023 to June 2024)
Iloprost, Inhaled	0.01	168	£2,465.19	eMIT (July 2023 to June 2024)
Selexipag, Oral	1.6	60	£3,000.00	BNF (accessed 21 st February 2025) -- Upravi 1,600microgram tablets Janssen-Cilag Ltd

BNF: British National Formulary; eMIT: Electronic Market Information Tool; IV: intravenous; SC: subcutaneous; UK: United Kingdom

*List price

The calculations assumed that in all cases when treatment dosing is dependent on patient weight, wastage would occur, therefore the number of vials are rounded up to the minimum number of vials required. As indicated in the scoping document by NHS England, the majority of patients will self-administer the treatment in their own homes (76). Consequently, vial sharing is deemed infeasible.

In the case of sotatercept, a detailed approach was taken, where the number of 45mg or 60mg vials were distributed according to number of vials needed per weight band (see Table 100

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

below). Using the mean weight and associated standard deviation from the non-USA population in the STELLAR trial [REDACTED], in line with the starting population for the economic model (Table 74), and assuming a log-normal distribution the proportion of patients falling into each weight band was calculated. The weighted average of the cost per vials needed and distribution across weight categories were calculated to get the cost per administration of sotatercept.

Table 100 Patient weight and required dosing

Patient Weight Range (kg)	Injection Volume (mL)	Kit type
30.0 – 31.7	0.4	1 x 45 mg vial
31.8 – 38.9	0.5	
39.0 – 46.0	0.6	
46.1 – 53.2	0.7	
53.3 – 60.3	0.8	
60.4 – 67.4	0.9	
67.5 – 74.6	1.0	1 x 60 mg vial
74.7 – 81.7	1.1	
81.8 – 88.9	1.2	
89.0 – 96.0	1.3	2 x 45 mg vials
96.1 – 103.2	1.4	
103.3 – 110.3	1.5	
110.4 – 117.4	1.6	
117.5 – 124.6	1.7	
124.7 – 131.7	1.8	2 x 60 mg vials
131.8 – 138.9	1.9	
139.0 – 146.0	2.0	
146.1 – 153.2	2.1	
153.3 – 160.3	2.2	
160.4 – 167.4	2.3	
167.5 and above	2.4	

The resulting cost of drug acquisition and administration per 12-week cycles and per year are presented in Table 101 below.

Table 101 Calculated drug acquisition costs per cycle and per year

Drug	Drug acquis. / 12 wks (£) ^a	Drug acquis. cost / yr (£)
Sotatercept, SC injection	[REDACTED]	[REDACTED]
Bosentan, oral	126	549
Ambrisentan, oral	97	421
Macitentan, oral	6,457	28,076
Sildenafil citrate, oral	114	497

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Tadalafil, oral	55	240
Epoprostenol, IV infusion	12,633	54,931
Treprostinil, IV infusion	3,571	15,526
Iloprost, inhaled	9,244	40,197
Selexipag, oral	8,400	36,525

IV: intravenous; SC: subcutaneous; wks: weeks; yr: year

^a Rounded to nearest GBP

3.5.2 Health state unit costs and resource use

The SLR of HCRU and costs identified no studies whereby resource use and costs in patients with PAH were stratified by risk status (see Appendix G for full details). Extensive discussions with clinical experts in the UK demonstrated that HCRU is dependent on the treatment regimen a patient receives rather than their risk status alone. Therefore, the approach in the economic model is to assign HCRU based on the treatment regimen received by a patient in that health state: which includes the resource use and costs associated with treatment initiation, administration, and the ongoing monitoring costs.

The combination of ERA + PDE5i is applicable to all risk status health states and therefore resource use is assumed to be equal in both modelled intervention and comparator arm, and consequently associated resource use is not included in the economic model.

Sotatercept + dual therapy resource use

Costs associated with the initiation of sotatercept in addition to ERA + PDE5i are applied as a one-off cost at the start of the model to all patients initiating sotatercept in the intermediate-low risk health state. Based on expert opinion from a UK centre, sotatercept patients will be invited to attend two face-to-face appointments prior to initiation: one to discuss the initiation and to gain agreement, and one for training on the administration. All patients will then have their first dose administered at a further face-to-face appointment, and for an estimated 50% of patients, the second dose will also be administered in this way, to ensure technique compliance and to answer any patient/caregiver queries. Therefore, an average of 3.5 face-to-face appointments per sotatercept patient is costed in the model base case. To account for variation in practice across UK centres, this assumption is tested in a scenario analysis. The SmPC specifies that the first 5 sotatercept doses should be accompanied by a blood test, to monitor for the onset of any AEs (130-133). Expert opinion received from a UK centre is that these tests will be conducted remotely, at a cost of £20 per test.

Table 102: Sotatercept resource use and costs

Resource use	Unit cost	Frequency	Source
Appointments associated with initiation	£95.80	3.50	Costs: National Schedule of NHS Costs Year 2023/24 Non-Consultant Led General Medicine WF01B

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

			Multiprofessional Non-Admitted face-to-Face Attendance, First Frequency: Clinical expert opinion
Blood test	£20.00	5.00	SmPC (130-133)

SmPC: summary of product characteristics

Selexipag + dual therapy resource use

Table 103: Selexipag regimen resource use and costs

Resource use	Unit cost	Frequency	Source
Titration period (4 months)			
Outpatient consultation prior to initiation	£147.30	One-off	Cost: National Schedule of NHS Costs Year 2023/24 Outpatient, Non-consultant led, Respiratory Physiology Service, WF01B, Non-admitted face-to-face attendance, First
PAH specialist nurse video/telephone call	£82.46	Every 2 weeks (Total of 8 in 4-month period)	Cost: National Schedule of NHS Costs Year 2023/24 District nursing service, N08AN, Specialist Nursing, Asthma and Respiratory Nursing/Liaison, Adult, Non-face-to-face Frequency: Clinical expert opinion
Total one-off cost			£806.98
Monitoring			
PAH specialist nurse video/telephone call	£82.46	Comfortable patients (90%): 1 call/ 3 months Patients of concern (10%): 1 call/ month	Cost: National Schedule of NHS Costs Year 2023/24 District nursing service, N08AN, Specialist Nursing, Asthma and Respiratory Nursing/Liaison, Adult, Non-face-to-face Frequency: Clinical expert opinion
Total recurring cost per cycle			£32.98

PAH: Pulmonary Arterial Hypertension

Prostacyclin + dual therapy resource use

A targeted review of available materials was conducted to understand the HCRU associated with IV PCAs. However, no source was identified that provided directly applicable costs or breakdown of resource use. Consequently, the micro-costing approach applied below relies on feedback from UK clinical experts and PAH nurse specialists who directly interact with patients during PCA use initiation and long-term usage. The HCRU associated with inhaled iloprost is assumed to be similar to that of selexipag, an oral treatment, based on the expert input.

Table 104 Intravenous prostacyclin resource use and costs

	Unit cost	Frequency	Source
Initiation			

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

	Unit cost	Frequency	Source
Training, day case	£147.30	One-off	National Schedule of NHS Costs Year 2023/24 Outpatient, Non-consultant led, Respiratory Physiology Service, WF01B, non-admitted face-to-face attendance, First
Catheterisation	Total cost: £4,830	One-off	Clinical expert opinion indicated that hospitalisation for catheterisation would cost £483 per night for a total 10-night stay
Outpatient appointment: stitch removal and decision on dose increase	£140.19	One-off	National Schedule of NHS Costs Year 2023/24 Non-Consultant Led, Respiratory Physiology Service, WF01A, Non-admitted Face-to-Face Attendance, Follow-up
Dose up-titration	Via video call appointments (90%): £82.46 per appointment Via day case re-admission (10%): £278 per appointment	One-off, 4 appointments	Video call appointment: National Schedule of NHS Costs Year 2023/24 District nursing service, N08AN, Specialist Nursing, Asthma and Respiratory Nursing/Liaison, Adult, non-face-to-face Day case re-admission: National Schedule of NHS Costs Year 2023/24 Day case, YR43A, Attention to Central Venous Catheter, 19 years and over
Total one-off cost			£ 5,525.55
Maintenance			
Consumables	£450	Per month	Clinical expert opinion
Hire of two pumps	£250	Per month	Clinical expert opinion
Routine care follow-up	£82.46	2 / month**	Costs: National Schedule of NHS Costs Year 2023/24 District nursing service, N08AN, Specialist Nursing, Asthma and Respiratory Nursing/Liaison, Adult, non-face-to-face Frequency: Roman et al. (2012) (41)
A&E admission for mild/moderate IV complication (5 days)	£4,923.30*	8% / year	Costs: National Schedule of NHS Costs Year 2023/24 Non-Elective Inpatient – Short Stay, WH07A, Infections or Other complications of procedures, with multiple interventions, with CC Score 2+ Frequency: Roman et al. (2012) (41)

	Unit cost	Frequency	Source
A&E admission for severe IV complication (10 days)	£12,711.80	8% / year	Costs: National Schedule of NHS Costs Year 2023/24 Non-Elective Inpatient – Long Stay, WH07A, Infections or other complications of procedures, with multiple interventions, with CC score 2+ Frequency: Roman et al. (2012) (41)
Re-catheterisation	£2,287.60	20% / year	Costs: National Schedule of NHS Costs Year 2023/24 Non-elective inpatient – Long stay, YR41A, Insertion on tunnelled central venous catheter, 19 years and over Frequency: Clinical expert opinion
Total recurring costs per cycle			£896.92

A&E: Accident & Emergency; IV: intravenous; NHS: National Health Service

*Assumption that Short Stay tariff of £984.66 is equivalent to one-day admission

** Clinical expert opinion indicated that routine care follow-up after catheterisation is initially once a month and extended to every 3 months once a patient is comfortable. This has been simplified to twice a month which is in line with published literature

Table 105 Inhaled prostacyclin resource use and costs

	Unit cost	Frequency	Source
Titration period (4 months)			
Outpatient consultation prior to initiation	£147.30	One-off	Cost: National Schedule of NHS Costs Year 2023/24 Outpatient, Non-consultant led, Respiratory Physiology Service, WF01B, non-admitted face-to-face attendance, First Frequency: assumed equal to selexipag + background therapy
PAH specialist nurse video/telephone call	£82.46	Every 2 weeks (Total of 8 in 4-month period)	Cost: National Schedule of NHS Costs Year 2023/24 District nursing service, N08AN, Specialist Nursing, Asthma and Respiratory Nursing/Liaison, Adult, non-face-to-face Frequency: assumed equal to selexipag + background therapy
Total one-off cost			£806.98
Monitoring			
PAH specialist nurse video/telephone call	£82.46	Comfortable patients (90%): 1 call/ 3 months	Cost: National Schedule of NHS Costs Year 2023/24 District nursing service, N08AN, Specialist Nursing, Asthma and Respiratory Nursing/Liaison, Adult, non-face-to-face

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

		Patients of concern (10%): 1 call/ month	Frequency: assumed equal to selexipag + background therapy
Total recurring costs per cycle			£32.98

A&E: Accident & Emergency; IV: intravenous; NHS: National Health Service; PAH: Pulmonary Arterial Hypertension

Costs applied in CEM

Resource use and costs identified for regimens of selexipag + background therapy and PCA + background therapy consisted of both one-off costs associated with the initiation of the drug as well recurring long-term maintenance and monitoring costs. Given the memoryless nature of the model, newly progressing cohorts are unable to be followed individually in every cycle. In order to preserve model parsimony, a simplified costing approach is adopted in the economic model whereby one-off costs are divided equally across all 12-week cycles in which a patient spends time in a health state where that cost is relevant. The one-off cost portion in each cycle is added to that cycle's recurring monitoring and maintenance costs (as presented in Table 103 to Table 105) to give a total cost per 12-week cycle that is applied to patients in that health state. A breakdown of this costing approach is provided below:

- All one-off costs were summed and were assumed to occur in the 1st month of a 12-weekly (approximately 3 monthly) period. Selexipag and inhaled PCA initiation was estimated to cost £806.98 and IV PCA to cost £ 5,525.55 in total.
- To avoid introducing unnecessary complexity, the one-off costs were summed with the recurring costs and applied in every cycle. To prevent overestimation of costs, the one-off value was distributed equally based on the undiscounted average number of cycles a patient is expected to live following the progression on sotatercept or selexipag and subsequent initiation of parenteral PCA. In the base case model, patients were estimated to live 3.80 model cycles while receiving selexipag + background therapy (ie. in either low or intermediate-low risk), while they were estimated to be receiving parenteral PCAs for 4.64 and 25.07 model cycles in the sotatercept + background therapy arm and selexipag + background therapy arm, respectively, following the applicable discontinuation criteria.
- The recurring costs for monitoring and maintenance were adjusted to reflect 3-monthly costs, which for simplicity is assumed to correspond to a per 12-week cycle cost.
- To avoid double counting (of initiation + maintenance) in the first cycle, the calculation assumed that only two-thirds of the per cycle maintenance cost would be taken into account during the first cycle. The calculation is presented below; the first cycle

receives this reduced maintenance cost, while the remaining cycles (reduced in number by the one-cycle already accounted for) receive the full per cycle maintenance cost. The total value was then divided by the total number of cycles to calculate the average recurring cost:

$$\bullet \text{ Average recurring costs} = \frac{\left(\frac{2}{3} \times \text{per cycle cost}\right) + (\text{number of remaining cycles} \times \text{per cycle cost})}{\text{Total number of cycles}}$$

Lastly, to account for the different administration method, and resource use, between inhaled and IV PCA, the parenteral PCA HCRU was weighted by the treatment distribution. The HCRU of patients receiving PCA via inhaled administration was assumed to be equivalent to the initiation and monitoring costs of patients receiving selexipag (Table 105). The recurring costs of IV PCA (i.e. epoprostenol and treprostinil) and inhaled PCA (i.e. iloprost) were weighted based on the treatment distributions of PCA in UK clinical practice (see Table 98).

After implementing the calculations above, recurring costs were applied in each cycle for patients undergoing selexipag or PCA treatment. Costs associated with sotatercept were assigned solely at the initiation of treatment, as no recurring monitoring costs were identified.

Table 106 Recurring HCRU costs used in the CEM

Treatment / Health State	Cost per 12 weeks	Cost per year
Selexipag in low and intermediate-low risk	■	■
On PCA, intermediate-high and high risk following sotatercept	■	■
On PCA, intermediate-high and high risk, following selexipag	■	■

CEM: cost-effectiveness model; HCRU: healthcare resource use; PCA: Prostacyclin analogue

3.5.3 Adverse reaction unit costs and resource use

No costs associated with treatment-related adverse event management since AEs are not considered in the model. As discussed earlier, none of the events observed in the clinical trial were deemed severe, resulting in low resource requirements for their resolution.

3.5.4 Miscellaneous unit costs and resource use

The model also includes one-time costs for clinical events, namely hospitalisation, transplantation and death. Details on these event costs and how they were derived from literature is presented in Table 107. All transplant events in the economic model are costed as lung transplants to reflect the nature of transplants in PAH patients in UK clinical practice. One-off costs for lung transplantation and mortality are applied upon progression to the post-transplant and death health states, respectively, meanwhile PAH hospitalisation costs are applied to the proportion of patients experiencing a hospitalisation event per cycle.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Table 107 Summary of clinical event costs

State	One-time event cost: Mean	SE	Source/Justification
PAH hospitalisation event	£1,825.14*	£356.27	Exposto 2021 (a HES database analysis of burden of PAH). Supplementary Figure 3. (21)
Lung/heart transplant event	£78,315.68	£15,663.14	National Schedule of NHS Costs Year 2023/24. Weighted average of Admitted Patient Care DZ01Z Lung Transplant
Mortality event	£15,743.00	£3,148.60	The Unit Costs of Health and Social Care 2023 Manual. Costs of hospital and social care services by 'Congestive heart failure' diagnostic group. Jones et al. (134)

FC: functional class; GBP: British pound sterling; NHS: National Health Service; PAH: pulmonary arterial hypertension; SE: standard error
*Reported as £1,544 in 2017 cost year, inflated to 2024 cost year using Unit Costs of Health and Social Care 2023 Manual. Table 12.1.1: The NHS Inflation Index (NHSCII). Page 98

3.6 Severity

The severity of the condition, defined as the future health lost by people living with the condition with standard care in the NHS, was calculated for the population receiving selezipag in combination with dual background therapy. The extent of unmet health need is reflected by the absolute and proportional QALY shortfall.

Inputs for the QALY shortfall calculation are informed by clinical trials and published data. The cohort characteristics in the STELLAR trial are assumed to be representative of the patient population of interest, with a mean age of 47.9 years and 20.7% being male (Table 108) (135).

Table 108 Summary features of QALY shortfall analysis

Factor	Value	Reference to section in submission
Sex distribution (male)	20.7%	See section 3.2
Starting age (mean)	47.9 years	

QALY: quality-adjusted life-year

Health state utilities inputs were informed by the EQ-5D analysis based on the STELLAR trial. For the calculation of QALYs of patients without the condition over their remaining life expectancy, UK life tables and UK age and sex adjusted utilities based on Hernandez Alava et al. 2022 have been used (92).

A summary of the QALY shortfall analysis is presented in Table 109, including the calculated and modelled discounted QALYs for the general population and for the population of interest on SoC treatment, respectively. This resulted in an absolute QALY shortfall of [REDACTED] and a proportional shortfall of [REDACTED] versus the general population. As stated in the NICE Technical Support Document 23 (TSD 23 (136)), the definition of QALY shortfall that provides the highest severity weight should be used. The proportional QALY shortfall is below 85%, however as Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

the absolute QALY shortfall is above 12 the disease severity qualities for a QALY weight of 1.2 to be applied to the modelled incremental QALYs.

Table 109 Summary of QALY shortfall analysis

Expected total QALYs for the general population	Total QALYs that people living with a condition would be expected to have with current treatment (BAT)	Absolute QALY shortfall	Proportional QALY shortfall	QALY weight
16.74				1.2

BAT: best available therapy; QALY: quality-adjusted life-year

3.7 Uncertainty

Uncertainty resulting from limitations in the evidence base is inevitable in the HTA of treatments for rare diseases. PAH is a rare and progressive disease with a paucity of evidence documenting the long-term outcomes of patients living with this condition.

A key challenge arose from the fact that the model was developed to reflect how patients are treated in clinical practice. As confirmed by UK clinicians, patients are managed according to the risk status-based approach outlined by the most recent ESC/ERS guidelines (4). While the analysis of the STELLAR trial provided strong evidence in support of this approach, it complicated the ITC. Neither the publication of the GRIPHON trial nor any relevant literature identified reported risk status-based results, or results for the constituent parameters to a sufficiently granular level, making an ITC versus selexipag for the endpoint of risk status change infeasible. To overcome this, a within-trial analysis of STELLAR was conducted. This analysis found that patients treated with PDE5i + ERA + sotatercept were more likely to experience an improvement in risk status from baseline to Week 24 compared to patients treated with PDE5i + ERA + selexipag (), and were less likely to experience a worsening in risk status (). Although the analysis is associated with limitations (please see section 2.8), it remains the only available source providing evidence for the comparative efficacy of sotatercept and selexipag in terms of risk status change.

Furthermore, UK clinicians confirmed that estimating the long-term survival of patients with PAH in the target population presents significant challenges. No patient remains in the same risk stratum over the long-term, making it very difficult to provide an accurate life expectancy at any given time.

As noted in section 3.3, a restriction was applied in the model, preventing patients from further risk status improvement beyond the follow-up duration of the STELLAR trial, a conservative approach supported by UK clinical experts. While PAH is a progressive disease and assuming indefinite improvement with sotatercept could be an overestimation of its clinical benefits, there is currently no data proving the opposite, i.e., that the improvement stops for all patients at a certain timepoint. Due to data uncertainty, a cap on the possibility for improvement was implemented.

It is currently assumed that all patients that discontinue sotatercept or selexipag due to disease progression go onto replace sotatercept or selexipag with a parenteral PCA (epoprostenol [62%], treprostinil [5%], or iloprost [33%]), per the guidelines (4). However, PCA adherence in clinical practice tends to be lower due to tolerability issues, which would imply some patients off-treatment (i.e., in intermediate-high and high risk off-treatment health states) would solely receive dual therapy. However, PCA usage has increased in recent years according to the 15th National Audit of Pulmonary Hypertension (2024), which may mitigate the discrepancy in adherence between clinical practice and treatment guidelines (7).

To mitigate these challenges, key assumptions and model inputs were thoroughly discussed and validated with UK clinical experts to ensure that the modelled outcomes are generalisable to UK clinical practice and the licensed indication. Additionally, a substantial number of sensitivity and scenario analyses were conducted to account for various uncertainties and enhance the robustness of the assessment.

Given the rarity of PAH and the evolving nature of the management of PAH (particularly in terms of treatment decisions being made based on a patient's risk status rather than WHO FC), there is inevitable uncertainty in the economic case. However, efforts have been made to optimise the use of the available data and to conduct extensive scenario analysis to inform the decision problem. Given the substantial unmet need experienced by PAH patients and the potential for sotatercept to offer a step-change in treatment, a greater acceptance of such uncertainty should be considered.

3.8 Managed access proposal

MSD believe sotatercept is a candidate for routine commissioning.

3.9 Summary of base case analysis inputs and assumptions

3.9.1 Summary of base case analysis inputs

Table 110 provides a summary of the inputs and settings used in the base case cost-effectiveness analysis.

Table 110 Summary of variables applied in the economic model

Variable	Value (reference to appropriate table or figure in submission)	Measurement of uncertainty and distribution: confidence interval (distribution)	Reference to section in submission
Model general settings			
Perspective	NHS and PSS	N/A	NICE reference case (93)
Time horizon	Lifetime (30 years)	N/A	Assumption
Discount rate: costs and outcomes	3.5%	N/A	NICE manual (78)
Model structure	Risk-based	N/A	Assumption
Patient characteristics			
Mean age, years	47.9	0.8 (Normal)	MSD data on file (2023). STELLAR (137)
Females, %	79.3%	2.3% (Beta)	MSD data on file (2023). STELLAR (137)
Body weight, kg	■	■	MSD data on file (2023). STELLAR – non-US population (137)
Age-specific mortality rate	UK national life tables rates by age and sex	N/A	ONS (138)
Markov Model			
Initial distribution	All patients in intermediate-low risk on sotatercept/ selexipag + dual therapy	N/A	Assumption per eligible population
Week-24 distribution (sotatercept + dual therapy) from intermediate-low risk	Low Risk: ■ Intermediate-low risk: ■ Intermediate-high risk: ■	(Dirichlet) Low risk: ■ Intermediate-low risk: ■	MSD data on file (2023). STELLAR (139)

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Variable	Value (reference to appropriate table or figure in submission)	Measurement of uncertainty and distribution: confidence interval (distribution)	Reference to section in submission
	High risk: ■ Low risk discon., sota: ■ Intermediate-low risk discon., sota: ■ Intermediate-high risk discon. A, sota: ■ High risk discon. A, sota: 0%	High risk: ■ Low risk siscon., sota: ■ Intermediate-low risk discon., sota: ■ Intermediate-high risk discon. A, sota: ■ High risk discon. A, sota: ■	
Week-24 distribution (selexipag + dual therapy) from intermediate-low risk	Low Risk: ■ Intermediate-low risk: ■ Intermediate-high risk: ■ High risk: ■	(Dirichlet) Low risk: ■ Intermediate-low risk: ■ Intermediate-high risk: ■ High risk: ■	MSD data on file (2023). STELLAR (139)
Long-term transition probabilities (sotatercept + dual therapy)	See Matrix in Table 82	NA	MSD data on file (2023). STELLAR (139)
RR, RS worsening, selexipag vs. sota	■	■	MSD data on file (2023). STELLAR (140)
RR, RS improvement, selexipag vs. sota	■	■	MSD data on file (2023). STELLAR (140)
Baseline age for mortality adjustment	51.7	10.3 (Normal)	Rosenkranz 2023 (79)
Mortality distribution for single parametric model fitted to low risk	Gamma	NA	Rosenkranz 2023 (79)
HR for mortality (relative to low risk)	Intermediate-low risk (including discon.): ■ Intermediate-high risk (including discon.): ■ High risk (including discon.): ■	(Log-normal) Intermediate-low risk: ■ Intermediate-high risk: ■ High risk: ■	Rosenkranz 2023 (79)
Post-transplant 12-week mortality risk	2.8%	0.6% (Beta)	Bernstein 2018 (81)

Variable	Value (reference to appropriate table or figure in submission)	Measurement of uncertainty and distribution: confidence interval (distribution)	Reference to section in submission
PAH hospitalisation risk	Low risk (including discon.): 1.8% Intermediate-low risk (including discon.): 2.7% Intermediate-high risk (including discon.): 4.3% High risk (including discon.): 5.7% Lung/heart transplant: 1.8%	(Beta) Low risk: 0.3% Intermediate-low risk: 0.2% Intermediate-high risk: 0.2% High risk: 0.6% Lung/heart transplant: 0.3%	MSD data on file (2023). COMPERA (82)
HR for PAH hospitalisation, sotatercept vs. selexipag in low and intermediate-low risk health states	■	■ (Log-normal)	MSD data on file (2023). ITC of overall STELLAR vs. GRIPHON (II or III only)
Lung/heart transplant risk	Intermediate-high risk: 0.05% High risk: 0.07%	0.01% (Beta)	MSD data on file (2023). COMPERA(80)
Utility			
Utility by health state	Low risk (including discon.): ■ Intermediate-low risk (including discon.): ■ Intermediate-high risk (including discon.): ■ High risk (including discon.): ■ Heart/Lung Transplant: ■	(Beta) Low risk: ■ Intermediate-low risk: ■ Intermediate-high risk: ■ High risk: ■ Heart/lung transplant: ■	UK Crosswalk (from EQ-5D-5L to EQ-5D-3L) Mapping Algorithm Utility Scores: Pooled Records (baseline and Week 24)
Disutility for treatment administration	IV infusion: 0.307 SC injection: 0 Inhaled: 0 Oral: 0	(Normal) IV infusion: 0.061 SC injection: 0 Inhaled: 0 Oral: 0	Infusion based on Davies 2018 (zero disutility for other administration methods is assumed)

Variable	Value (reference to appropriate table or figure in submission)	Measurement of uncertainty and distribution: confidence interval (distribution)	Reference to section in submission
Hospitalisation disutility	0.105	0.021 (Normal)	Tanna 2021 (a SLR of utilities for heart failure)
Transplant disutility	0.105	0.021 (Normal)	Assume equal to hospitalisation disutility
Carer utility gain	Low risk: -0.036 Intermediate-low risk: -0.023 Intermediate-high risk: -0.013 High risk:0	(Normal) Low risk: -0.007 Intermediate-low risk: -0.005 Intermediate-high risk: -0.003 High risk:0	Calculated using multiplier from Pennington et al. 2024 (24) and the difference in patient health state utility compared to patient high risk utility.
Costs			
Drug acquisition costs (12-weeks) <u>Branded drug costs were not included in the one-way and probabilistic sensitivity analyses, given the precision of the input value.</u>	Bosentan, oral: £126.27 Ambrisentan, oral: £96.85 Sildenafil citrate, oral: £114.21 Tadalafil, oral: £55.16 Treprostinil, IV infusion: £3,570.70 Iloprost, inhaled: £9244.46	(Gamma) Bosentan, oral: £25.25 Ambrisentan, oral: £19.37 Sildenafil citrate, oral: £22.84 Tadalafil, oral: £11.03 Treprostinil, IV infusion: £714.14 Iloprost, inhaled: £1848.89	Dosing and drug package information
Number of administrations in 12 weeks	Sotatercept, SC injection: 4 Bosentan, oral:168 Ambrisentan, oral: 84 Macitentan, oral: 84 Sildenafil citrate, oral: 252 Tadalafil, oral: 84 Epoprostenol, IV infusion: 84 Treprostinil, IV infusion: NA Iloprost, inhaled: 630 Selexipag, oral: 168	NA	Relevant labels and clinical trials

Variable	Value (reference to appropriate table or figure in submission)	Measurement of uncertainty and distribution: confidence interval (distribution)	Reference to section in submission
Treatment distribution sotatercept (low risk, intermediate-low risk and intermediate-high risk)	Sotatercept, SC injection: 100% Bosentan, oral: 12.8% Ambrisentan, oral: 43.3% Macitentan, oral: 44.0% Sildenafil citrate, oral: 61.9% Tadalafil, oral: 38.1% Epoprostenol, IV infusion: 0% Treprostinil, IV infusion: 0% Iloprost, inhaled: 0% Selexipag, oral: 0%	(Beta) Sotatercept, SC injection: 20% Bosentan, oral: 3% Ambrisentan, oral: 8% Macitentan, oral: 9% Sildenafil citrate, oral: 13% Tadalafil, oral: 7% Epoprostenol, IV infusion: 0% Treprostinil, IV infusion: 0% Iloprost, inhaled: 0% Selexipag, oral: 0%	NHS 2024 UK PAH audit (7)
Treatment distribution sotatercept (high risk)	Sotatercept, SC injection: 0% Bosentan, oral: 12.8% Ambrisentan, oral: 43.3% Macitentan, oral: 44.0% Sildenafil citrate, oral: 61.9% Tadalafil, oral: 38.1% Epoprostenol, IV infusion: 61.9% Treprostinil, IV infusion: 5.0% Iloprost, inhaled: 33.1% Selexipag, oral: 0%	(Beta) Sotatercept, SC injection: 0% Bosentan, oral: 3% Ambrisentan, oral: 8% Macitentan, oral: 9% Sildenafil citrate, oral: 13% Tadalafil, oral: 7% Epoprostenol, IV infusion: 12% Treprostinil, IV infusion: 0% Treprostinil, SC infusion: 0% Iloprost, Inhaled: 7% Selexipag, oral: 0%	NHS 2024 UK PAH audit (7)
Treatment distribution selexipag (low risk and intermediate-low risk)	Sotatercept, SC injection: 0% Bosentan, oral: 12.8% Ambrisentan, oral: 43.3%	(Beta) Sotatercept, SC injection: 0% Bosentan, oral: 3%	NHS 2024 UK PAH audit (7)

Variable	Value (reference to appropriate table or figure in submission)	Measurement of uncertainty and distribution: confidence interval (distribution)	Reference to section in submission
	Macitentan, oral: 44.0% Sildenafil citrate, Oral: 61.9% Tadalafil, oral: 38.1% Epoprostenol, IV infusion: 0% Treprostinil, IV infusion: 0% Iloprost, inhaled: 0% Selexipag, oral: 100%	Ambrisentan, oral: 8% Macitentan, oral: 9% Sildenafil citrate, oral: 13% Tadalafil, oral: 7% Epoprostenol, IV infusion: 0% Treprostinil, IV infusion: 0% Iloprost, inhaled: 0% Selexipag, oral: 20%	
Treatment distribution selexipag (intermediate-high risk and high risk)	Sotatercept, SC injection: 0% Bosentan, oral: 12.8% Ambrisentan, oral: 43.3% Macitentan, oral: 44.0% Sildenafil citrate, oral: 61.9% Tadalafil, oral: 38.1% Epoprostenol, IV infusion: 61.9% Treprostinil, IV infusion: 5.0% Iloprost, inhaled: 33.1% Selexipag, oral: 0%	(Beta) Sotatercept, SC injection: 0% Bosentan, oral: 3% Ambrisentan, oral: 8% Macitentan, oral: 9% Sildenafil citrate, oral: 13% Tadalafil, oral: 7% Epoprostenol, IV infusion: 12% Treprostinil, IV infusion: 0% Iloprost, inhaled: 7% Selexipag, oral: 0%	NHS 2024 UK PAH audit (7)
One-Off HCRU: Sotatercept admin costs	Cost per admin: £95.80 Average number of admins: 3.5	(Gamma) £19.16 0.70	
One-Off HCRU: Sotatercept Blood Tests	Cost per Test: £20 Number of Tests: 5	(Gamma) £4.00 1.00	
Recurring HCRU costs by health state per 12-week cycle (sotatercept + dual therapy)	Low risk: £0	(Gamma) £0	National Schedule of NHS Costs Year 2023/24 (141)

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Variable	Value (reference to appropriate table or figure in submission)	Measurement of uncertainty and distribution: confidence interval (distribution)	Reference to section in submission
	Intermediate-low risk: £0 Intermediate-high risk: £0 High risk: £2,555	£0 £0 £511	Clinical expert opinion
Recurring HCRU costs by health state (selexipag + dual therapy)	Low risk: £302 Intermediate-low risk: £302 Intermediate-high risk: £1,966 High risk: £1,966	£60 £60 £393 £393	National Schedule of NHS Costs Year 2023/24 (141) Clinical expert opinion
Recurring HCRU costs by health state (discontinued sotatercept + dual therapy)	Low risk: £0 Intermediate-low risk: £0 Intermediate-high risk: £1,966 High risk: £2,555	£0 £0 £393 £511	National Schedule of NHS Costs Year 2023/24 (141) Clinical expert opinion
Cost of hospitalisation event	£1,825.14	£365.03 (Gamma)	Exposto 2021 (a HES database analysis of burden of PAH). Supplementary Figure 3
Cost of transplant event	£78,315.68	£15,663.14 (Gamma)	National Schedule of NHS Costs Year 2023/24. Total HRGs DZ01Z Lung Transplant (141)
Mortality event cost	£15,743.00	£3,148.60 (Gamma)	The Unit Costs of Health and Social Care 2023 Manual. Costs of hospital and social care services by 'Congestive heart failure' diagnostic group. Jones et al.

IV: intravenous; N/A: not applicable; NHS: National Health Service; NICE: National Institute for Health and Care Excellence; ONS: Office for National Statistics; PSS: Personal Social Services; SC: subcutaneous

3.9.2 Assumptions

A list of all assumptions applied in the model base case are summarised in Table 111.

Table 111 Summary of assumptions and corresponding rationale

Description of assumption used in the base case	Justification
Categorising patients according to ESC/ERS risk status rather than WHO FC	Feedback from clinical experts in the UK confirmed that patient management aligns with these guidelines, suggesting that risk status is a more clinically appropriate method for constructing the model for this decision problem. WHO FC subjectivity and the potential for implausible results were factors in this decision.
30-year time horizon is applied	As 0.06% of patients are alive at 30 years this time horizon was considered long enough to capture the long-term clinical and economic impacts of PAH, an incurable disease requiring treatment until end of life. Alternative horizons are tested in scenario analyses.
A 12-week cycle length is assumed in the Markov model, following an initial decision tree of 24 weeks duration	A half-cycle correction is applied in the Markov model to account for the fact that events and transitions can occur at any point during the cycle, not necessarily at the start or end of each cycle. The 12-week cycle lengths implemented in the model provide an appropriate level of detail to capture changes in risk status and associated costs.
Discontinuation of sotatercept is considered in the case of inadequate responders at 24 weeks after treatment start	Based on KOL feedback, clinicians wait until week 24 to consider discontinuing a new therapy. They will also consider the components of the ESC/ERS risk categorisation when making this decision (6MWD, WHO FC and NT-proBNP). The removal of this rule is tested in a scenario analysis.
All patients discontinue sotatercept therapy who transition to the high risk state	Due to uncertain evidence, patients whose disease worsens significantly and move to the high risk state will discontinue sotatercept therapy. The removal of this rule is tested in a scenario analysis.
All patients discontinue selexipag therapy who transition to the intermediate-high risk or high risk state	Based on clinical feedback and current guidelines, the discontinuation of selexipag treatment was defined as movement to a worse health state than intermediate-low risk.
No risk status improvement is assumed beyond 24 weeks	The conservative assumption was adopted that all improvement beyond 24 weeks converts to maintenance, equivalent to the duration of follow-up available from STELLAR. This is a conservative assumption in the absence of randomised evidence that examined the improvement of patients receiving sotatercept on a longer-term.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Description of assumption used in the base case	Justification
Patients receiving sotatercept: In the model, the transition probabilities from FCIII are used as a proxy for all transitions from the intermediate-high risk state	██████████ were observed from intermediate-high to high risk in STELLAR. To address this data gap, a proxy was applied using transitions from the FCIII category, to allow transitions from intermediate-high to high risk.
Patients who discontinued sotatercept and selexipag: transition probabilities for intermediate-high and high risk is assumed to be equal to transition probability for FC III and FC IV based on Roman 2012 (41) – Epoprostenol arm.	Published literature was used to inform these transitions as the STELLAR trial lacks the specific patient population required for this input. Specifically, Roman 2012 reported 12-week WHO FC-based transition probabilities for patients receiving epoprostenol, iloprost and treprostinil. In UK practice, epoprostenol is used most often therefore the transitions reported for epoprostenol use were used in this analysis.
Health state utilities are defined by ESC/ERS risk status rather than WHO FC	The derivation of health state utilities based on risk status is appropriate given the health state definitions applied in the analysis. Data directly available from the STELLAR trial were utilised. A scenario analysis using WHO FC values from the trial is presented.
Patients receiving lung/heart transplant are assumed to have the same quality of life as low risk patients	No HRQoL information were identified for patients who have undergone lung/heart transplant in PAH; therefore a conservative approach was taken where patients are assigned the best HRQoL possible. Due to the low number of transplants in the UK, this is not a driver of model results.
Disutility applied to patients receiving IV infusion	The administration of IV PAH therapies has been reported to impact a patient's QoL negatively and a study reporting a value, Davies et al. 2018 (102), was identified.
Adverse events	This economic evaluation does not account for AEs given the comparable safety outcomes between the sotatercept + background therapy and placebo + background therapy arms in STELLAR, indicating that the addition of sotatercept to background therapy does not increase the incidence of AEs. Furthermore, all of the AEs observed were mild or moderate in terms of their severity, assumed to have minimal impact on the health care system and patient's QoL. The severity of the AEs reported in the GRIPHON trial was not reported.

6MWD: 6-minute walk distance; AE: adverse event; ERS: European Respiratory Society; ESC: European Society of Cardiology; IV intravenous; KOL: key opinion leader; NT-proBNP: N-terminal pro-brain natriuretic peptide; PAH; pulmonary arterial hypertension; PCA: prostacyclin analogue; RCT: randomised controlled trial; UK: United Kingdom; WHO FC: World Health Organization Functional Class; QoL = quality-of-life

3.10 Base case results

3.10.1 Base case incremental cost-effectiveness analysis results

The base case results, for initiating sotatercept + dual background therapy vs selexipag + dual background therapy in intermediate-low risk, incorporating the model settings specified in Table 110 are presented in Table 112.

Sotatercept in combination with dual therapy accumulated [REDACTED] of costs and [REDACTED] QALYs and selexipag in combination with dual therapy accumulated [REDACTED] of costs and [REDACTED] QALYs. The resulting ICER of [REDACTED] per QALY, inclusive of the 1.2 QALY severity modifier for which sotatercept qualifies (see Section 3.6), falls below the £30,000/QALY willingness-to-pay threshold and is consequently cost-effective.

Additional base case results can be found in Appendix H, including clinical outcomes (i.e., hospitalisations, transplants and survival) as well as detailed health economic outcomes (i.e., life-years [LYs] and QALYs by health state, individual QALY losses, and cost categories by health state).

Table 112 Deterministic base case results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)	ICER with severity modifier (£/QALY)
Sotatercept + dual therapy	████	██	██	████	██	2.50	████	████
Selexipag + dual therapy	████	6.39	██	-	-	-	-	-

ICER: incremental cost-effectiveness ratio; LYG: life-years gained; QALY: quality-adjusted life-years

Table 113 presents the incremental deterministic net health benefit (NHB) and net monetary benefit (NMB) versus sotatercept.

Table 113 Deterministic net health benefit and net monetary benefit

Technologies	Incremental costs (£)	Incremental QALYs	NHB at £20,000	NHB at £30,000	NMB at £20,000	NMB at £30,000
Selexipag + dual therapy	████	2.50 * 1.2 = 3.00	██	██	████	████

NHB: net health benefit; NMB: net monetary benefit; LYG: life-years gained; QALY: quality-adjusted life-years

3.11 Exploring uncertainty

3.11.1 Probabilistic sensitivity analysis

A probabilistic sensitivity analysis (PSA) was run for 1,000 iterations.

The parameters varied in the PSA and the nature of the uncertainty around their value are described in Table 110 above.

The mean summary results from the probabilistic analysis are presented in

Table 114 below. Graphical results including the cost-effectiveness plane (incremental costs vs. incremental QALYs) and the cost-effectiveness acceptability curve (CEAC) are shown below in Figure 23 and Figure 24 respectively.

Table 114 Probabilistic base case results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)	ICER with severity modifier (£/QALY)	ICER weighted according to modifier (£/QALY)
Severity modifier									
Sotatercept + dual therapy									
Selexipag + dual therapy									

ICER: incremental cost-effectiveness ratio; LYG: life-years gained; QALY: quality-adjusted life-years

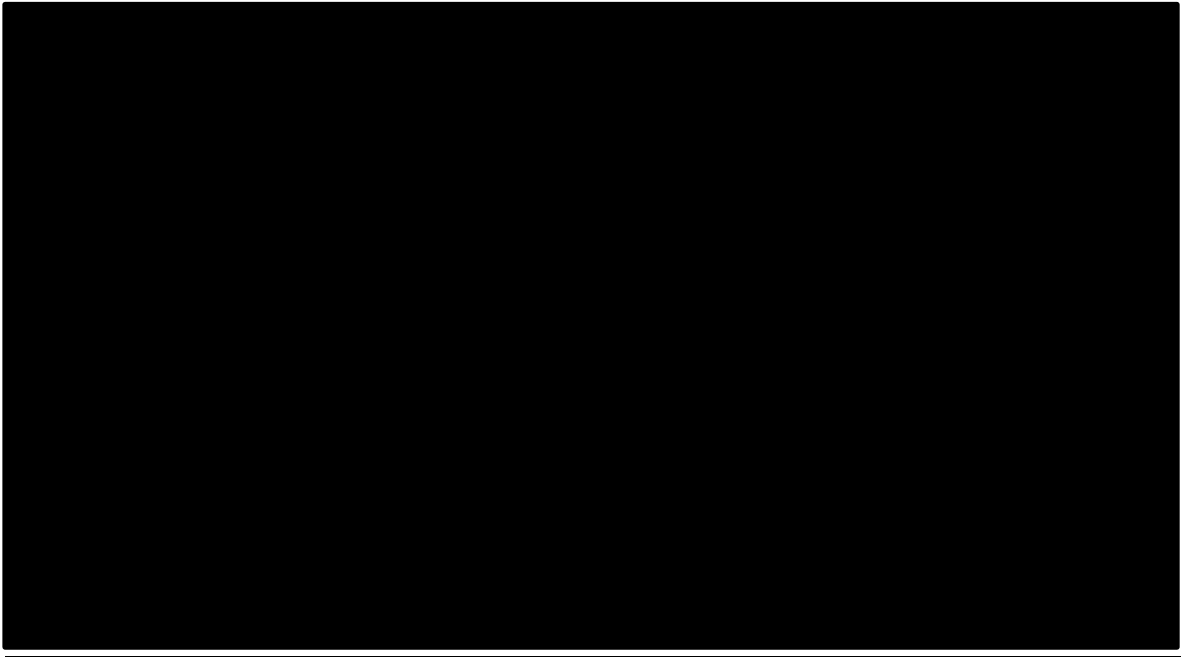
Based on the 1000 iterations, sotatercept compared to selexipag add as an add-on to dual background therapy in intermediate-low risk was cost-effective in █ of iterations at a willingness-to-pay threshold of £30,000 per QALY, with more than █ of those being dominant (i.e., sotatercept is more efficacious and less costly). In less than █ of the iterations was sotatercept projected to be dominated by selexipag or less efficacious but less costly.

Figure 23 Cost-effectiveness plane scatter plot (1,000 iterations) – Incremental Costs and QALYs (severity modifier applied)



QALY: quality-adjusted life-year; WTP: willingness-to-pay

Figure 24 Cost-effectiveness acceptability curve (1,000 iterations) (severity modifier applied)



BGT: background therapy; QALY = quality-adjusted life-year

3.11.2 Deterministic sensitivity analysis

A deterministic sensitivity analysis (DSA) was conducted to explore the impact of key model parameter values on the model outcomes, especially on the ICER (£/QALY). The lower and upper bounds for parameter values were varied in the DSA according to their 95% CI. The CI was calculated based on reported uncertainty metrics from the source or, where CIs were not available, assuming a standard error (SE) of 20% of the reported mean value. The top 10 most influential parameters on ICER per QALY outcomes are presented in the tabular results in Table 115 and the tornado diagram in Figure 25.

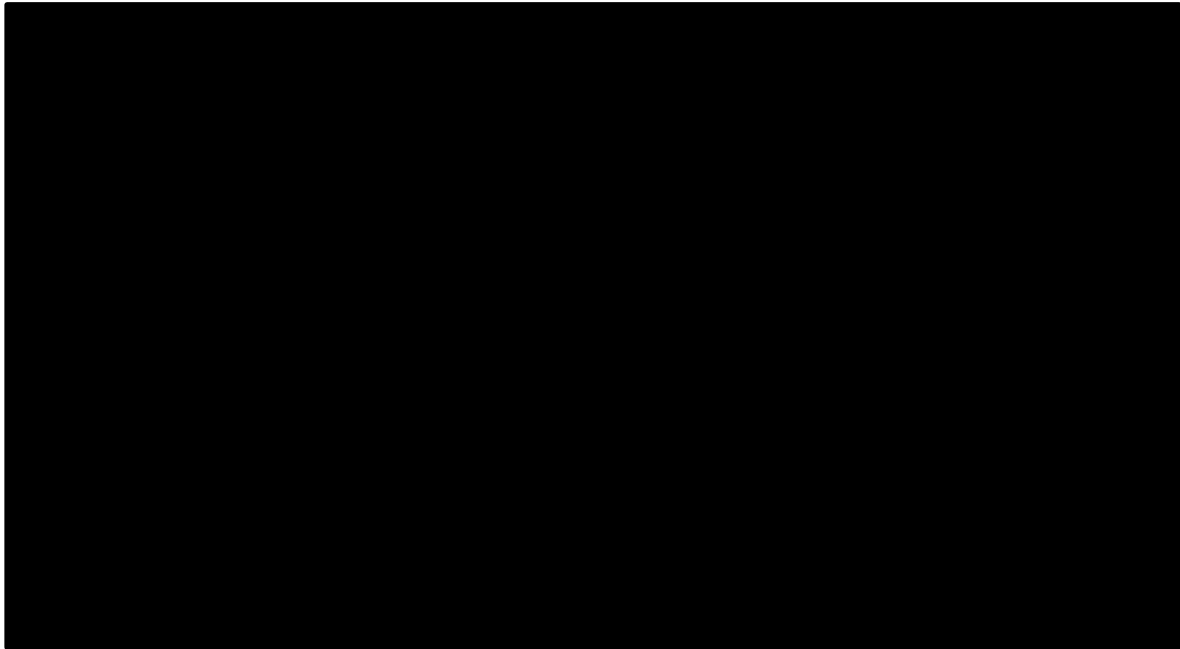
With the severity modifier applied, all DSA results remain below the £30,000/QALY threshold. The two parameters that have the greatest impact on the results are the drug acquisition cost of iloprost and the RR of risk status worsening between sotatercept and selexipag. The cost of iloprost per 12 weeks is derived using the most recent costs reported by eMIT, as per the NICE reference case, and is therefore a robust representation of the NHS iloprost drug acquisition cost, using best available data. While the RR of risk status worsening is derived from a within-trial analysis of sotatercept vs. selexipag in STELLAR, the value is varied across an arbitrary range in the DSA, rather than reflecting the true clinical uncertainty of this parameter. A more plausible assumption has been assessed in a scenario analysis. All other parameters have minimal impact and are not considered to be drivers of the model results.

Table 115 Ten most impactful parameters from one-way deterministic sensitivity analysis on ICER (severity modifier applied)

Parameter, (Base Case [Lower Bound, Upper Bound])	ICER for lower bound value	ICER for upper bound value
Base case		
Drug acquisition cost/12 wks: Iloprost, Inhaled (9244.456 [5982.52, 13204.82])		
RR, RS worsening (sotatercept vs selexipag from within-STELLAR analysis) (0.11 [0.0, 0.74])		
Macitentan, Oral - Sotatercept + dual BGT - Int-High (0.44 [0.27, 0.61])		
Age (years) (47.9 [46.33, 49.47])		
Non-sotatercept PAH drugs: HCRU Cost Selexipag + dual BGT: High Risk (1966.354 [1272.52, 2808.75])		
Macitentan, Oral - Sotatercept + dual BGT - Int-Low (0.44 [0.27, 0.61])		
Macitentan, Oral - Selexipag + dual BGT - Int-High (0.44 [0.27, 0.61])		
Macitentan, Oral - Sotatercept + dual BGT - Low (0.44 [0.27, 0.61])		
Non-sotatercept PAH drugs: HCRU Cost Selexipag + dual BGT: Intermediate-High Risk (1966.354 [1272.52, 2808.75])		
Admin. utility decrement: IV infusion (0.307 [0.19, 0.43])		

BGT = background therapy; HCRU = health-related resource use; ICER: incremental cost-effectiveness ratio; RS = risk status

Figure 25 Tornado diagram for ICER impact for one-way deterministic sensitivity analysis (severity modifier applied)



BGT: background therapy; HCRU: healthcare resource use; ICER: incremental cost-effectiveness ratio; IV: intravenous; PAH: pulmonary arterial hypertension; QALY: quality-adjusted life-year; RR: relative risk; RS: risk status; SC: subcutaneous injection; wks: weeks

3.11.3 Scenario analysis

Various scenarios were explored to understand the impact of structural and input assumptions in the model base case. The list of scenarios and the rationale behind their inclusion in the analysis is presented in Table 116.

Table 116 Scenario Analyses

Scenario	Rationale
Hospitalisation HR = 1	Given the inherent uncertainty associated with ITC, the impact of excluding the relative benefit of sotatercept in preventing PAH-related hospitalisations is tested.
Relative risk of improvement and worsening from WHO FC ITC	An alternative source was used, relying on the Bucher ITC conducted versus selexipag, based on the GRIPHON and TRACE trials (as reported in Table 65 of section 2.10). Since GRIPHON and TRACE did not report risk-based outcomes, WHO FC change was used as a proxy, although the outcome is not directly comparable. WHO FC is considered a more subjective measure and is less sensitive to treatment efficacy. This scenario is conducted to provide an alternative to the only available source of relative efficacy versus selexipag that was included in the base case.
Mortality approach: dependent model using Gompertz	As described in section 3.3.2, the dependent model to estimate long-term overall survival is tested in scenarios. The evolution of risks over time, as provided by the Gompertz and Weibull distributions, were judged to be potential alternatives to the base case choice and therefore included in the scenario analyses.
Mortality approach: dependent model using Weibull	
Mortality approach: Boucly et al 2022 + UK age	An alternative source to estimate survival is based on Boucly et al., as described in section 3.3.2. The scenario included updated HRs reported by Boucly, applied to the projected low risk survival also reported by Boucly (reported HRs of 2.65, 5.65 and 11.08 for intermediate-low risk, intermediate-high, and high, relative to low risk, using a Cox proportional hazards model) (77). As the mean age in the source dataset was 61.0 years, the starting age of the cohort is set to the average UK age of 58.0 years, as reported in the PAH audit. Age adjustment is applied up to the age of 61.0 years.
Sotatercept administered in high risk with 100% PCA use	Based on the uncertain evidence in high risk patients, it was assumed in the base case that patients would not receive sotatercept in the high risk state. In a scenario analysis, the impact of keeping patients on sotatercept while concurrently initiating PCA therapy is tested.
Sotatercept administered in high risk with 75% PCA use	In addition to the above assumption, based on clinician feedback, it is hypothesised that for patients receiving a combination of sotatercept and PCA, the use of PCA would decrease to 75% of the original uptake.
Sotatercept administered in high risk with 50% PCA use	In addition to the above assumption, based on clinician feedback, it is hypothesised that for patients receiving a combination of sotatercept and PCA, the use of PCA would decrease to 50% of the original uptake.

Scenario	Rationale
No treatment discontinuation at 24 weeks	Clinical experts proposed base case discontinuation criteria at 24 weeks for patients who exhibit an insufficient response, which is not stipulated in the product SmPC. The impact of removing this discontinuation assumption is explored in scenario analysis.
Quality of life capped at general population values	The base case age adjustment of utility values is removed and health state utilities are capped at general population normal values if the model estimates them to be higher. This is only observed in the low risk health state.
WHO FC-based health state utility values	Given the sparsity of reported state values for risk status categories, an alternative approach using values according to WHO FC is explored, which reflects the more common approach in the literature. The scenario assumes that WHO FC-based QoL is a reasonable approximation of risk status, despite their known differences in predicting survival. Utilities based on WHO FC from the STELLAR study were derived in a similar manner to how they were derived based on risk groups.
Removal of carer quality of life	The impact of excluding the impact of carer utility is tested.
Removal of transplant disutility	The impact of excluding the impact of disutility due to transplantation is tested.
Mean weight of STELLAR ITT population	Sotatercept is a weight-based therapy, and USA trial participants were excluded from the base case mean, for reasons of generalisability. The impact of this exclusion is tested in a scenario, where the mean weight of the STELLAR ITT population is used. The mean (SD) weight was 72.3 (18.3) kg.
Sotatercept HCRU	The sotatercept SmPC specifies blood tests are required alongside the first 5 doses. The base case assumption is that these tests will be conducted remotely, however to account for difference in monitoring approaches across PAH centres, the scenario tests the impacts of this monitoring taking place face-to-face. The 5 appointments would follow an initial decision-making meeting and training session, leading to a total of 7 face-to-face appointments for all patients.
Time horizon of 10 years	An alternative time horizon of 10 years is tested.
Time horizon of 20 years	An alternative time horizon of 20 years is tested.
Discount rate of 1.5%	A discount rate of 1.5% for both costs and health benefits is used, reflecting the scenario discussed in the NICE Methods Manual. The extent and timing of the longer-term benefits of sotatercept are uncertain but the innovative mechanism of action merits this exploration.

FC: functional class; HCRU: healthcare resource utilisation; HR: hazard ratio; ICER: incremental cost-effectiveness ratio; ITC: indirect treatment comparison; ITT: intention-to-treat; LY: life-years; QALY: quality-adjusted life-years; PAH: pulmonary arterial hypertension; PCA: prostacyclin analogue; UK: United Kingdom

The results of the scenario analyses are presented in Table 117. In almost all scenarios, the ICER remains below the threshold of £30,000/QALY indicating the base case conclusion to

be robust. The limited number of scenarios in which the scenario analysis ICER exceeds the threshold are those which employ alternative approaches to modelling mortality.

As discussed in section 3.3.2 Mortality, the alternative mortality approach, using the dependent model whereby parametric survival functions are jointly fitted to each risk status KM curve, is likely to overestimate the long-term survival to give less clinically plausible results. This approach is also linked to greater uncertainty due to low patient numbers at risk in the KM curves for the higher risk states. Furthermore, the source of risk-based mortality is the COMPERA cohort with no comorbidities reported by Rosenkranz, which is considered to be a more appropriate reflection of the STELLAR baseline characteristics, noting the impact of comorbidities on mortality, as confirmed by UK clinical experts.

The ICER decreases relative to the base case with the use of alternative time horizons and discount rates, and of health state utility values reported per WHO FC.

Table 117 Scenario analyses deterministic results

Scenario	Incremental costs	Incremental LYs	Incremental QALYs	ICER (£/QALY)	ICER with severity modifier (£/QALY)
Base case	████	██	2.50	████	████
Hospitalisation HR = 1	████	██	2.46	████	████
Relative risk of improvement and worsening from WHO FC ITC	████	██	2.06	████	████
Mortality approach: dependent model using Gompertz	████	██	2.76	████	████
Mortality approach: dependent model using Weibull	████	██	2.82	████	████
Mortality approach: Boucly et al 2022 + UK age	████	██	2.32	████	████
Sotatercept administered in high risk with 100% PCA use	████	██	2.47	████	████
Sotatercept administered in high risk with 75% PCA use	████	██	2.50	████	████
Sotatercept administered in high risk with 50% PCA use	████	██	2.53	████	████
No treatment discontinuation at 24 weeks	████	██	2.61	████	████
Quality of life capped at general population values	████	██	2.53	████	████
WHO FC-based health state utility values	████	██	2.75	████	████
Removal of carer quality of life	████	██	2.40	████	████
Removal of transplant disutility	████	██	2.50	████	████
Mean weight of STELLAR ITT population	████	██	2.50	████	████
Sotatercept HCRU	████	██	2.50	████	████
Time horizon of 10 years	████	██	2.24	████	████
Time horizon of 20 years	████	██	2.50	████	████
Discount rate of 1.5%	████	██	2.79	████	████

ICER: incremental cost-effectiveness ratio; LY: life-years; QALY: quality-adjusted life-years

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

3.12 Subgroup analysis

No economic subgroup analyses were conducted as part of this submission, as the base case results already reflect a subgroup of the licensed population i.e. the expected position for sotatercept in UK clinical practice.

3.13 Benefits not captured in the QALY calculation

Sotatercept is expected to generate a range of health-related benefits that may be underestimated or missing entirely from the current cost-effectiveness analysis. Additionally, there may be certain health-related benefits that free up NHS resources or that have their cost impacts underestimated.

3.13.1 Patient and caregiver productivity

In line with the reference case, productivity losses for patients or their caregivers are not captured in the model. However, as discussed previously, 79% of PH patients have reported an impact on their ability to work or attend education (16). Furthermore, 60% of PH patients reported that their condition had an impact on their financial situation, with 25% reporting a “major impact” (16). There is also a reported impact for family members of PAH patients (26). In particular, given the age distribution of PAH patients, more effective treatments associated with a reduced monitoring burden could result in improved productivity for patients (through less absenteeism or dropping out of the workforce linked to achieving a lower risk status) and for caregivers (through less absenteeism necessitated by caring for the patient).

3.13.2 Caregiver quality of life utility

Caregiver quality of life is accounted for in the base case analysis, with utility gains linked to the patient’s improved risk status based on the Pennington et al. (2024) publication (24), which investigates various impacts on caregiver SF-6D. However, the results reported in this study are disease-agnostic, and therefore may underestimate the relationship between patient utility and caregiver utility in PAH. For more severe PAH patients, where the burden of continuous intravenous treatments has been widely described (20), the relationship between patient and caregiver utility may result in greater gains linked to the patient avoiding high risk status by introducing sotatercept earlier in the pathway than is accounted for in the base case QALY estimation.

The base case assumes that each patient has one informal caregiver, a conservative assumption necessitated by lack of data. However, given the age profile of PAH patients tends to skew to middle age, patients may have more than one informal caregiver across

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

© Merck Sharp & Dohme (UK) Limited (2025). All rights reserved

generations (parents, spouse and adult children). Therefore, improved patient risk status from more effective treatment, which translates into improvements in caregiver quality of life, could result in greater QALY gains overall if more than one familial relation is required to provide informal caregiving to the PAH patient.

3.13.3 Increased capacity on transplant list

Despite being rare in UK clinical practice, lung and heart-lung transplants are important treatment options for PAH patients with uncontrolled disease and are captured in the base case. However, the current analysis does not account for the impact of being on a waiting list prior to receiving the transplant, or the impact of this waiting list on the demand for donor organs in the healthcare system, specifically for the outcomes of patients in other disease areas where lung and heart/lung transplants are required. Given the transplant eligibility criteria, the improved risk status associated with sotatercept treatment should reduce the number of PAH transplant-eligible patients, and hence the demand for donor organs; not only reducing resources associated with the transplant procedure but also improving patient outcomes throughout the wider system due to shorter waiting times.

3.13.4 Psychological benefit of an innovative treatment option

Data from the 2023 PHA UK survey recorded that 84% of people with PH say the condition has an impact on their mental or emotional wellbeing, with 34% saying it has a “major impact” (16). It is expected that PAH patients whose disease is uncontrolled by dual therapy will feel more optimistic if they have an additional treatment option. Furthermore, sotatercept when added to dual therapy has been shown to improve disease severity significantly more so than the other available option for this group of patients, selexipag (see details in Section 2.10), providing patients with hope of risk status and WHO FC improvement, not just maintenance. Other treatment options for this group of patients focus on symptom management, which although necessary to mitigate negative QoL impacts, does not provide patients much hope for the future of their disease. Therefore, the availability of sotatercept could psychologically benefit PAH patients by providing hope for long-term improved disease outcomes. Furthermore, better disease control is likely to negate the need to initiate intravenous therapies, which is linked with poorer QoL and a psychological strain (20). These potential improvements could have an additional benefit for NHS mental health services capacity.

3.13.5 Stable disease can lead to more remote management

There are a limited number of specialist centres for the treatment of PAH in the UK and their diffuse geographical spread can often result in patients travelling long distances for their

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

© Merck Sharp & Dohme (UK) Limited (2025). All rights reserved

appointments. Since patients with better disease control can be managed more remotely, the improved risk status associated with sotatercept treatment should result in not only less resource use, but also less frequent travel; which reduces transportation costs and improves QoL for the patient and their family.

3.13.6 Wider cardiovascular health benefits from improved exercise capacity

PAH is associated with a reduced exercise capacity, and the primary endpoint of the STELLAR trial showed marked improvements in the participants' 6MWD at week 24 compared to baseline (see section 2.6.1), reflecting a greater capacity to be physically active while on-treatment. Reduced physical activity is a known risk factor for other diseases often associated with the cardiovascular system, the presence of which can compound the poor overall health of a PAH patient. Improving a patient's ability to take exercise is expected with a more effective treatment like sotatercept, and this is likely to reduce their susceptibility to develop or exacerbate comorbidities. This is a benefit which has not been captured in the base case QALY estimation.

3.13.7 Reduced use of anticoagulation medicines

In certain UK centres, PCA treatments are co-prescribed with anticoagulant medicines to reduce the risk of catheter-associated thrombosis. These treatments are associated with their own AEs, such as bleeding, and ongoing monitoring. The improvements in risk status associated with sotatercept treatment are likely to lead to a reduced initiation of PCA and consequently of anticoagulants, with a benefit for the patient and the healthcare budget. The inconsistent pattern of this co-prescribing across the UK has led to its exclusion from the base case QALY estimation.

3.13.8 Insensitivity of utility measure in PAH

EQ-5D is a well-established, commonly accepted utility instrument, however the general nature of its questions can result in limited sensitivity, especially to mental health impacts, that can be more pronounced in chronic diseases (although anxiety/depression is included in the questionnaire). Instruments have been developed for eliciting QoL information in PAH patients, such as PAH-SYMPACT, for which sotatercept demonstrated a statistically significant improvement in the Physical Impacts and Cardiopulmonary Symptoms components (see Section 2.6.1 for details). A mapping algorithm to EQ-5D was not identified. Therefore, the treatment-related improvements in patient utility through risk status change (from the EQ-

5D) may not detect other improvements that can occur while the patient maintains their same risk status.

3.13.9 Survival benefit beyond time spent in improved health states

Sotatercept's impact on all-cause mortality is currently captured in the base case through improvements in risk status, since risk status is defined by 1-year mortality rates (4). However, the causes of mortality in PAH are not completely understood and may be influenced by patient risk factors not captured in the risk status calculation, leading to a potential underestimate in the survival benefit of sotatercept. The STELLAR trial reported a marked difference in mortality rate between patients on sotatercept + background therapy and those in the placebo + background therapy arm (see section 2.6.1), although this effect is associated with uncertainty due to few mortality events during the STELLAR trial. This between-arm difference is not incorporated into the base case QALY estimate.

3.13.10 Selexipag AE impact on quality of life

As discussed in section 3.4.4, the base case model does not account for utility impacts due to AEs. Feedback from clinical experts indicated that the titration phase following selexipag initiation can be associated with the same AEs reported in the GRIPHON trial (24) e.g. headache, nausea, diarrhoea, myalgia, jaw pain etc. These events have been associated with reduced QoL in the literature, however given the severity of these events is a data gap, the conservative decision was made to exclude them from the base case QALY estimate. In STELLAR, the rate of serious AEs experienced by participants taking sotatercept was not significantly greater compared to that experienced by participants taking placebo (29).

3.14 Validation

3.14.1 Validation of cost-effectiveness analysis

Clinical validation

Expert clinical input was sought during the development of the cost-effectiveness model to ensure that the inputs and assumptions used in the analysis were relevant to UK clinical practice and to validate the clinical plausibility of the outcomes predicted by the model. Feedback was obtained in two advisory boards and in total, input was gathered from seven UK clinical experts. As detailed throughout the submission, the clinical experts supported the approaches and assumptions taken in the development of the cost-effectiveness model. Expert clinical opinion was sought to validate the following model inputs: use of risk-based

health state definitions, treatment discontinuation rules, long-term treatment efficacy, HCRU for patients with PAH. Further details of expert clinical opinion are presented in Appendix R.

Technical validation

The model was subjected to a thorough validation process in accordance with guidelines for validation put forth by the International Society for Pharmacoeconomics and Outcomes Research Society and the Society for Medical Decision-Making Joint Task Force for Modelling Good Research Practices (142). These guidelines stress the importance of face validity (confirming the model approach, data sources and assumptions with experts), internal validity (quality-checking of parameter values and calculations) and external validity (comparing model results with other published studies).

Face validity was tested throughout model development with external health economic and clinical experts during two advisory boards conducted by MSD. Internal validity was tested by researchers not involved in model development checking the accuracy of all data extracted from the literature, the logical structure of the model and the accuracy of all calculations and programming. Additionally, the researchers conducting the quality control review, in collaboration with the model developers, subjected the model to a series of diagnostic tests to ensure that the model reacted as expected. External validation was not possible as this is the first cost-effectiveness model utilising the risk-based categorisation of health states. All other cost-effectiveness models identified used a WHO FC-based approach, which was deemed inappropriate for this decision problem as outlined in section 3.2

3.15 Interpretation and conclusions of economic evidence

3.15.1 Conclusions from the cost-effectiveness analysis

Patients with PAH have a substantial clinical, humanistic and economic burden. There is a high unmet need for a targeted, efficacious, and safe new treatment for these patients. STELLAR provided a robust body of evidence to support the efficacy and impact of sotatercept on overall health of participants presenting with PAH.

The economic analysis described in this section presents a robust methodology for estimating the value of sotatercept as a treatment for PAH, utilising a novel approach that reflects current UK clinical practice. Sotatercept is cost-effective when compared to selexipag for the treatment of adult patients with intermediate-low risk of death due to PAH, with an incremental QALY gain of 2.50 and an ICER of █████ per QALY █████ (when accounting for the severity modifier). In particular, sotatercept demonstrates value for the healthcare system through retaining patients in lower risk states for longer than the current SoC, which are associated Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

© Merck Sharp & Dohme (UK) Limited (2025). All rights reserved

with a reduced resource use burden. The value of these lower risk states to the patient manifests as improved QoL and a lower mortality risk.

The model structure and inputs were validated by UK clinicians. Assumptions included in the model were tested with extensive scenario and sensitivity analyses which demonstrated that the base case ICER is stable to variation in inputs and structural assumptions. Based on the PSA, there is ■■■ likelihood that sotatercept is cost-effective versus selexipag at a willingness-to-pay (WTP) threshold of £30,000 per QALY as an addition to background therapy in intermediate-low risk.

Sotatercept represents a cost-effective use of NHS resources when initiated in patients with intermediate-low risk of death due to PAH compared with current therapy and would address a considerable unmet need for this rare but highly burdened population.

3.15.2 Generalisability to clinical practice

Clinical efficacy data is taken directly from randomised controlled evidence for sotatercept in PAH (section 2.3.1), with comparative efficacy versus the relevant comparator on the key endpoint derived from the only available source. A subpopulation of the patient populations included in the STELLAR and SOTERIA trials were included in the economic analysis, namely patients with intermediate-low risk of death due to PAH in order to reflect the expected point of sotatercept initiation in the clinical pathway.

3.15.3 Strengths

The model used for the analysis has been developed to reflect the current treatment pathway for PAH patients in the UK, which relies on the recommendations in the 2022 ESC/(ERS guidelines) (4).

This is also the first CEM presented in the UK which utilises risk-based health state definitions, a departure from all identified CEMs that previously relied on grouping patients based on WHO FC (refer to section 3.1), and the first ever being evaluated by NICE. As discussed in section 3.2, a model structure based on WHO FC is no longer congruent with current clinical practice. The analysis presented is also the first to present utility values defined based on risk groups. This provides a more comprehensive understanding of patient QoL in accordance with the most recent clinical definitions.

Additionally, the analysis benefits from the comprehensive utilisation of data from the STELLAR trial, including baseline characteristics, treatment efficacy, comparative efficacy and HRQoL measures, ensuring a robust and comprehensive assessment. Where external

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

© Merck Sharp & Dohme (UK) Limited (2025). All rights reserved

sources for inputs were required e.g., relationship between risk status and mortality, hospitalisation rates, HCRU associated with risk status or treatment, data was retrieved from the most appropriate sources available at the time of submission and validated with UK clinicians.

3.15.4 Limitations

The key efficacy outcomes in the STELLAR trial were measured up to 24 weeks which is considerably shorter than the 30-year time horizon covered by the economic model.

Due to the utilisation of a risk-based approach to health state definition, conducting an ITC versus selexipag for this endpoint using data from the latter's pivotal clinical trial is not feasible. Therefore, a non-randomised within-trial analysis associated with limitations was conducted to address the data gap. In a scenario analysis, the results of a trial-based ITC on the WHO FC change endpoints were applied to understand the impact.

Clinicians face challenges when it comes to accurately estimating survival rates as the worsening of the disease involves treatment escalation while also discontinuation of selexipag, a treatment with which UK clinicians do not yet have long-term experience with (it was commissioned in the UK in 2018). This makes validation of the model's longer-term survival estimates challenging, so plausibility criteria were applied during curve selection. (e.g. hazard pattern expectations, relative survival between risk groups). However, alternative assumptions were explored in scenario analyses.

During the follow-up of patients in the STELLAR and SOTERIA studies, the transition from intermediate-high risk to high risk states could not be determined. To address this data gap, the WHO FC classification of patients from the STELLAR study was used, specifically transitions from FC III. To maintain the 100% sum of transition probabilities, transitions from intermediate-high risk were replaced by transitions from WHO FCIII. Recalling clinical expert comments about the subjectivity of FC, this is likely to be a conservative assumption about sotatercept's efficacy in this risk group. While FC III and IV cannot be equated directly to intermediate-high risk and high risk, they are considered imperfect but acceptable proxies based on clinical feedback, in the absence of alternative evidence.

When patients discontinued treatment with sotatercept or selexipag, the breadth of available evidence was limited to data involving FC-based classification of health states. An external source presenting transition probabilities using FC III and IV as proxies (Roman, 2012 (41)) was utilised to fill the data gap, similar to the transition from intermediate-high to high risk.

Lastly, the SLR of HCRU and costs identified no studies whereby resource use and costs in patients with PAH were stratified by risk status. Extensive discussions with clinical experts in the UK demonstrated that HCRU is dependent on the treatment regimen a patient receives rather than their risk status alone. Therefore, the approach in the economic model is to assign HCRU based on the treatment regimen received by a patient in that health state: which includes the resource use and costs associated with treatment initiation, administration, and the ongoing monitoring costs.

3.15.5 Summary

Sotatercept addresses the high unmet need for patients with intermediate-low risk PAH who incur substantial clinical, humanistic and economic burden. Sotatercept demonstrated clinically and statistically significant improvements of PAH, as well as HRQoL with an acceptable benefit-risk profile.

The value of sotatercept was established in a robust economic analysis where sotatercept was shown to be a cost-effective treatment for PAH at a WTP threshold of £30,000 for the expected population, compared to the current SoC. The sensitivity and scenario analyses ICERs indicate that sotatercept is cost-effective when tested against a range of key model inputs and assumptions. Incremental QALY gains were generally in the range of 2.06 and 2.82 with ICERs ranging from [REDACTED] to [REDACTED] per QALY.

Given the rarity of PAH and the evolving nature of the management of PAH (particularly in terms of treatment decisions being made based on a patient's risk status rather than WHO FC), there is inevitable uncertainty in the economic case. However, efforts have been made to optimise the use of the available data and to conduct extensive scenario analysis to inform the decision problem. Given the substantial unmet need of this rare patient population, the potential for sotatercept to offer a step-change in treatment and the potential value that is not captured in the cost-effectiveness analysis a greater acceptance of such uncertainty must be considered.

4 References

1. Galiè N, Humbert M, Vachiery JL, Gibbs S, Lang I, Torbicki A, et al. 2015 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension: The Joint Task Force for the Diagnosis and Treatment of Pulmonary Hypertension of the European Society of Cardiology (ESC) and the European Respiratory Society (ERS): Endorsed by: Association for European Paediatric and Congenital Cardiology (AEPC), International Society for Heart and Lung Transplantation (ISHLT). *Eur Respir J*. 2015;46(4):903-75.
2. Prins KW, Thenappan T. World Health Organization Group I Pulmonary Hypertension: Epidemiology and Pathophysiology. *Cardiology Clinics*. 2016;34(3):363-74.
3. Delcroix M, Howard L. Pulmonary arterial hypertension: the burden of disease and impact on quality of life. *Eur Respir Rev*. 2015;24(138):621-9.
4. Humbert M, Kovacs G, Hoeper MM, Badagliacca R, Berger RMF, Brida M, et al. 2022 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension. *Eur Heart J*. 2022;43(38):3618-731.
5. Hoeper MM, Badesch DB, Ghofrani HA, Gibbs JSR, Gombert-Maitland M, McLaughlin VV, et al. Phase 3 Trial of Sotatercept for Treatment of Pulmonary Arterial Hypertension. *N Engl J Med*. 2023;388(16):1478-90.
6. Preston IR, Badesch D, Ghofrani H-A, Gibbs JSR, Gombert-Maitland M, Hoeper MM, et al. A Long-Term Follow-Up Study of Sotatercept for Treatment of Pulmonary Arterial Hypertension: Interim Results of SOTERIA. *European Respiratory Journal*. 2025:2401435.
7. NHS. National Pulmonary Hypertension Audit, 15th Annual Report2024. Available from: <https://digital.nhs.uk/data-and-information/publications/statistical/national-pulmonary-hypertension-audit/15th-annual-report>.
8. Humbert M, McLaughlin V, Gibbs JSR, Gombert-Maitland M, Hoeper MM, Preston IR, et al. Sotatercept for the treatment of pulmonary arterial hypertension. *New England Journal of Medicine*. 2021;384(13):1204-15.
9. Guignabert C, Savale L, Boucly A, Thuillet R, Tu L, Ottaviani M, et al. Serum and Pulmonary Expression Profiles of the Activin Signaling System in Pulmonary Arterial Hypertension. *Circulation*. 2023;147(24):1809-22.
10. Levine DJ. Pulmonary arterial hypertension: updates in epidemiology and evaluation of patients. *Am J Manag Care*. 2021;27(3 Suppl):S35-s41.
11. Thenappan T, Ormiston ML, Ryan JJ, Archer SL. Pulmonary arterial hypertension: pathogenesis and clinical management. *Bmj*. 2018;360:j5492.
12. CADTH. CADTH Therapeutic Review Report: Drugs for Pulmonary Arterial Hypertension: Comparative Efficacy, Safety, and Cost-Effectiveness — Recommendations Report2015. Available from: https://www.cadth.ca/sites/default/files/pdf/TR0006_PAH_Recs_Report.pdf.
13. Ivarsson B, Johansson A, Kjellström B. The Odyssey from Symptom to Diagnosis of Pulmonary Hypertension from the Patients and Spouses Perspective. *J Prim Care Community Health*. 2021;12:21501327211029241.

14. Guillevin L, Armstrong I, Aldrighetti R, Howard LS, Ryfstenius H, Fischer A, et al. Understanding the impact of pulmonary arterial hypertension on patients' and carers' lives. *Eur Respir Rev.* 2013;22(130):535-42.
15. Mao J, Bennett A, Harley J, Small M, Nadipelli VR, Butler J, Phatak H. Pulmonary Arterial Hypertension Treatment Patterns and Patient Health-Related Quality of Life: Findings from a Multinational Real-World Database. *American Journal of Respiratory and Critical Care Medicine.* 2022;205:A3050.
16. PHA UK. Breathless, not voiceless: What it means to live with PH today2023 2024. Available from: <https://www.phauk.org/research/pha-uk-led-research/breathless-not-voiceless-what-it-means-to-live-with-ph-today/>.
17. NHS. National Audit of Pulmonary Hypertension Great Britain, 2018-19 (10th annual report). 2019 October, 24.
18. PHA UK. The true emotional impact of pulmonary hypertension2019 2024. Available from: <https://www.phauk.org/living-with-pulmonary-hypertension/the-emotional-impact-of-ph-survey-results/>.
19. PHA UK. What it means to live with PH today2017. Available from: <https://www.phauk.org/research-survey-work/living-ph-report/>.
20. MSD. Patient Support Research in Preparation for Sotatercept Launch - Full Report. 2024.
21. Exposto F, Hermans R, Nordgren Å, Taylor L, Sikander Rehman S, Ogley R, et al. Burden of pulmonary arterial hypertension in England: retrospective HES database analysis. *Therapeutic Advances in Respiratory Disease.* 2021;15:1753466621995040.
22. Zozaya N, Abdalla F, Casado Moreno I, Crespo-Diz C, Ramírez Gallardo AM, Rueda Soriano J, et al. The economic burden of pulmonary arterial hypertension in Spain. *BMC Pulmonary Medicine.* 2022;22(1):105.
23. Studer S, Hull M, Pruett J, Koep E, Tsang Y, Drake W, 3rd. Treatment patterns, healthcare resource utilization, and healthcare costs among patients with pulmonary arterial hypertension in a real-world US database. *Pulm Circ.* 2019;9(1):2045894018816294.
24. Pennington BM, Alava MH, Strong M. Unpaid Caring and Health-Related Quality of Life: Longitudinal Analysis of Understanding Society (the UK Household Longitudinal Survey). *Value Health.* 2025;28(1):138-47.
25. Joish VN, Kreilick C, Germino R, Muccino D. PCV95 - Estimation Of Pulmonary Arterial Hypertension On Productivity Losses In The United States. *Value in Health.* 2014;17(3):A118.
26. Budhram BM, J.; Alemayehu, W.; Westerhout, C.; Gubbels, C.; Budd, E.; Abraham, T.; Weatherald, J., editor *Analyzing the Impact of Pulmonary Arterial Hypertension on Patients' Need for Caregivers and Caregiver Work Patterns: Results from a Canadian National Patient Survey.* World Symposium on Pulmonary Hypertension; 2024; Barcelona, Spain: MSD data on file.
27. Ruopp NF, Cockrill BA. Diagnosis and Treatment of Pulmonary Arterial Hypertension: A Review. *Jama.* 2022;327(14):1379-91.
28. Chin KM, Gaine SP, Gerges C, Jing ZC, Mathai SC, Tamura Y, et al. Treatment algorithm for pulmonary arterial hypertension. *Eur Respir J.* 2024;64(4).

29. MSD. STELLAR (MK-7962-003) Clinical Study Report (06-DEC-2022 LPLV) - data on file. 2024.
30. Preston IR, Badesch D, Ghofrani H-A, Gibbs JSR, Gombert-Maitland M, Hoeper MM, et al. Late Breaking Abstract - A long-term follow-up (LTFU) study of sotatercept for pulmonary arterial hypertension (PAH). *European Respiratory Journal*. 2023;62(suppl 67):OA739.
31. MSD. STELLAR (MK-7962-003) subgroup and post-hoc analyses (06-DEC-2022 LPLV) - data on file. 2024.
32. Hoeper MM, Pausch C, Olsson KM, Huscher D, Pittrow D, Grunig E, et al. COMPERA 2.0: a refined four-stratum risk assessment model for pulmonary arterial hypertension. *Eur Respir J*. 2022;60(1).
33. MSD (Confidential). [Confidential Data on File] Confidential STELLAR post-hoc analyses results. 2024.
34. Chin KM, Gombert-Maitland M, Channick RN, Cuttica MJ, Fischer A, Frantz RP, et al. Psychometric Validation of the Pulmonary Arterial Hypertension-Symptoms and Impact (PAH-SYMPACT) Questionnaire: Results of the SYMPHONY Trial. *Chest*. 2018;154(4):848-61.
35. DuBrock HM, Reddy YN, Durst LA, Schroeder DR, Park G, Cajigas HR, et al. The feasibility and value of assessing patient-reported outcomes in pulmonary arterial hypertension. *Pulm Circ*. 2022;12(4):e12143.
36. MSD. SOTERIA (MK-7962-004) subgroup and post-hoc analyses (08-NOV-2023 cumulative data cut-off date) - data on file. 2024.
37. Preston IR, editor SOTERIA: A long-term follow-up study of Sotatercept for pulmonary arterial hypertension. Pulmonary Vascular Research Institute 2024 Annual Congress; 2024 31 January – 03 February 2024; London: MSD data on file.
38. MSD (Confidential). [Confidential Data on File] Confidential SOTERIA ad-hoc analyses results. 2024.
39. Mathai SC, Puhon MA, Lam D, Wise RA. The minimal important difference in the 6-minute walk test for patients with pulmonary arterial hypertension. *Am J Respir Crit Care Med*. 2012;186(5):428-33.
40. NICE. Appendix I Quality appraisal checklist – economic evaluations NICE process and methods [PMG4]2012b [Available from: <https://www.nice.org.uk/process/pmg4/chapter/appendix-i-quality-appraisal-checklist-economic-evaluations>].
41. Roman A, Barberà JA, Escibano P, Sala ML, Febrer L, Oyagüez I, et al. Cost effectiveness of prostacyclins in pulmonary arterial hypertension. *Appl Health Econ Health Policy*. 2012;10(3):175-88.
42. Wlodarczyk J, Hunsche E, Silberberg B, Kemp D, Anell B, Corcoran K, et al. Cost-Effectiveness Analysis of Selexipag Versus Inhaled Iloprost In The Treatment of Pulmonary Arterial Hypertension (PAH) In Sweden. *Value in Health*. 2016;19(7):A556.
43. Coyle K, Coyle D, Blouin J, Lee K, Jabr MF, Tran K, et al. Cost Effectiveness of First-Line Oral Therapies for Pulmonary Arterial Hypertension: A Modelling Study. *Pharmacoeconomics*. 2016;34(5):509-20.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

© Merck Sharp & Dohme (UK) Limited (2025). All rights reserved

44. Lilyasari O, Subekti Y, Atika N, Dinarti LK, Putri S, Opitasari C, et al. Economic evaluation of sildenafil for the treatment of pulmonary arterial hypertension in Indonesia. *BMC Health Serv Res.* 2019;19(1):573.
45. Thongsri W, Bussabawalai T, Leelahavarong P, Wanitkun S, Durongpisitkul K, Chaikledkaew U, Teerawattananon Y. Cost-utility and budget impact analysis of drug treatments in pulmonary arterial hypertension associated with congenital heart diseases in Thailand. *Expert Review of Pharmacoeconomics & Outcomes Research.* 2016;16(4):525-36.
46. Ekhlasi M, Sheikhi S, Majd ZK, Peiravian F, Yousefi N. Cost-Effectiveness Analysis of Macitentan in Comparison With Bosentan in the Treatment of Pulmonary Arterial Hypertension in Iran. *Value in Health Regional Issues.* 2023;34:78-85.
47. Nosrati M, Shahmirzadi NA, Afzali M, Zaboli P, Rouhani H, Hamedifar H, Hajimiri M. Cost-utility analysis of Macitentan Vs. Bosentan in pulmonary atrial hypertension. *Journal of Family Medicine and Primary Care.* 2020;9(7):3634-8.
48. Tran-Duy A, Morrisroe K, Clarke P, Stevens W, Proudman S, Sahhar J, Nikpour M. Cost-Effectiveness of Combination Therapy for Patients With Systemic Sclerosis-Related Pulmonary Arterial Hypertension. *J Am Heart Assoc.* 2021;10(7):e015816.
49. Dong W, Zhang Z, Chu M, Gu P, Hu M, Liu L, et al. Cost-effectiveness analysis of selexipag for the combined treatment of pulmonary arterial hypertension. *Front Pharmacol.* 2023;14:1122866.
50. ICER. Sotatercept for Pulmonary Arterial Hypertension - Final Evidence Report. 2024 January 8.
51. CADTH. Treprostinil (Remodulin®) Canadian Agency for Drugs and Technologies in Health; 2004 November 2004.
52. PBS. Epoprostenol (Flolan®). The Pharmaceutical Benefit Scheme; 2006 March 2006.
53. AWMSG. Selexipag (Uptravi®). All Wales Medicines Strategy Group; 2018 June 2018.
54. CADTH. Selexipag (Uptravi®). Canadian Agency for Drugs and Technologies in Health; 2016 October 2016.
55. NCPE. Selexipag (Uptravi®). National Centre for Pharmacoeconomics, Ireland; 2018 May 2018.
56. SMC. Selexipag (Uptravi®). Scottish Medicines Consortium; 2018 May 2018.
57. AWMSG. Sildenafil (Revatio®) for the treatment of pulmonary arterial hypertension (WHO functional class II or III) to improve exercise capacity All Wales Medicines Strategy Group; 2010 June 2010.
58. CADTH. Sildenafil Citrate (Revatio® - Pfizer Canada Inc.). Canadian Agency for Drugs and Technologies in Health; 2007 February 2007.
59. PBS. Sildenafil citrate (Revatio®). The Pharmaceutical Benefit Scheme; 2006 November 2006.
60. SMC. Sildenafil citrate 20mg tablets (Revatio®). Scottish Medicines Consortium; 2006 March 2006.
61. SMC. Sildenafil citrate (Revatio®) Scottish Medicines Consortium; 2010 February 2010.
62. CADTH. Tadalafil (Adcirca®) Canadian Agency for Drugs and Technologies in Health; 2010 July 2010.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

63. SMC. Tadalafil (Adcirca®) Scottish Medicines Consortium; 2012 July 2012.
64. AWMSG. Riociguat (Adempas®). All Wales Medicines Strategy Group; 2015 September 2015.
65. CADTH. Riociguat (Adempas®). Canadian Agency for Drugs and Technologies in Health; 2014 March 2014.
66. SMC. Riociguat (Adempas®). Scottish Medicines Consortium; 2015 July 2015.
67. CADTH. Ambrisentan (Volibris®). Canadian Agency for Drugs and Technologies in Health; 2008 December 2008.
68. NCPE. Ambrisentan (Volibris®). National Centre for Pharmacoeconomics, Ireland; 2009 February 2009.
69. SMC. Ambrisentan (Volibris®). Scottish Medicines Consortium; 2008 November 2008.
70. PBS. Bosentan (Tracleer®) The Pharmaceutical Benefit Scheme; 2008 March 2008.
71. AWMSG. Macitentan (Opsumit®). All Wales Medicines Strategy Group; 2016 January 2016.
72. CADTH. Macitentan (Opsumit®). Canadian Agency for Drugs and Technologies in Health; 2015 January 2015.
73. SMC. Macitentan (Opsumit®). Scottish Medicines Consortium; 2014 April 2014.
74. CADTH. Macitentan and Tadalafil (Opsynvi®). Canadian Agency for Drugs and Technologies in Health; 2021 December 2021.
75. PBAC. Public Summary Report - July 2020 PBAC Meeting: SELEXIPAG, Tablets, 200, 400, 600, 800, 1000, 1200, 1400, 1600 microgram, Upravi®, Janssen-Cilag Pty Ltd.; 2020.
76. National Institute for Health and Care Excellence (NICE). Health Technology Evaluation. Sotatercept for treating pulmonary arterial hypertension - Final scope. 2024.
77. Boucly A, Weatherald J, Savale L, de Groote P, Cottin V, Prevot G, et al. External validation of a refined four-stratum risk assessment score from the French pulmonary hypertension registry. Eur Respir J. 2022;59(6).
78. NICE. Clinical Commissioning Policy: Selexipag for treating pulmonary arterial hypertension (adults). 2018.
79. Rosenkranz S, Pausch C, Coghlan JG, Huscher D, Pittrow D, Grunig E, et al. Risk stratification and response to therapy in patients with pulmonary arterial hypertension and comorbidities: A COMPERA analysis. J Heart Lung Transplant. 2023;42(1):102-14.
80. MSD. [Data on file] COMPERA registry: Association between long-term clinical outcomes and risk status in PAH patients (time to transitions). 2022.
81. Bernstein EJ, Bathon JM, Lederer DJ. Survival of adults with systemic autoimmune rheumatic diseases and pulmonary arterial hypertension after lung transplantation. Rheumatology (Oxford). 2018;57(5):831-4.
82. MSD. [Data on File] COMPERA registry: Association between long-term clinical outcomes and risk status / WHO-FC in PAH patients (survival FC risk change). 2023.
83. MSD. STELLAR (MK-7962-003) Clinical Study Report (26-AUG-2022 data cut-off date) - data on file.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

© Merck Sharp & Dohme (UK) Limited (2025). All rights reserved

84. Galie N, Humbert M, Vachiery JL, Gibbs S, Lang I, Torbicki A, et al. 2015 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension: The Joint Task Force for the Diagnosis and Treatment of Pulmonary Hypertension of the European Society of Cardiology (ESC) and the European Respiratory Society (ERS): Endorsed by: Association for European Paediatric and Congenital Cardiology (AEPC), International Society for Heart and Lung Transplantation (ISHLT). *Eur Heart J*. 2016;37(1):67-119.
85. 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.
86. Thenappan T, Shah SJ, Rich S, Tian L, Archer SL, Gomberg-Maitland M. Survival in pulmonary arterial hypertension: a reappraisal of the NIH risk stratification equation. *Eur Respir J*. 2010;35(5):1079-87.
87. Hoeper MM, Simon RGJ. The changing landscape of pulmonary arterial hypertension and implications for patient care. *Eur Respir Rev*. 2014;23(134):450-7.
88. Hoeper MM, Huscher D, Ghofrani HA, Delcroix M, Distler O, Schweiger C, et al. Elderly patients diagnosed with idiopathic pulmonary arterial hypertension: Results from the COMPERA registry. *International Journal of Cardiology*. 2013;168(2):871-80.
89. Office of National Statistics (ONS). National Life Tables: UK 2017-2019 2024 [Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables>].
90. Sitbon O, Channick R, Chin KM, Frey A, Gaine S, Galie N, et al. Selexipag for the Treatment of Pulmonary Arterial Hypertension. *N Engl J Med*. 2015;373(26):2522-33.
91. Reeves GEM, Shepherd J, Collins NJ, Twaddell S, Harjit Singh R. Assessing quality of life in pulmonary arterial hypertension: An independent prognostic marker. *Pulm Circ*. 2024;14(2):e12380.
92. Hernández Alava M, Pudney S, Wailoo A. Estimating EQ-5D by age and sex for the UK. 2022.
93. NICE. Developing NICE guidelines: the manual. 2024.
94. Fuge J, Park D-H, von Lengerke T, Richter MJ, Kamp JC, Olsson KM. Impact of pulmonary arterial hypertension on employment, work productivity, and quality of life- results of a cross-sectional multi-center study. *Frontiers in psychiatry*. 2022;12:781532.
95. Park DH, Fuge J, Kahl KG, Kamp JC, Olsson KM. Impact of SARS-CoV-2-pandemic on mental disorders and quality of life in patients with pulmonary arterial hypertension. *Frontiers in psychiatry*. 2021;12:668647.
96. Martínez Meñaca A, Blanco Vich I, López Meseguer M, López Reyes R, Bedate Díaz P, Elías Hernández T, et al. From Health-Related Quality of Life (HRQoL) of Patients with Pulmonary Hypertension to Patient Experience with the Care Received: Should We Be More Aware of Current Patient Needs? *Advances in therapy*. 2021;38:1860-75.
97. Hidayati F, Gharini PP, Hartopo AB, Anggrahini DW, Dinarti LK. The effect of oral sildenafil therapy on health-related quality of life in adults with pulmonary arterial hypertension related to uncorrected secundum atrial septal defect: A quasi experimental study. *Health and Quality of Life Outcomes*. 2020;18(1):1-9.

98. Dytrych V, Kuchar J, Ambroz D, Kunstyr J, Grus T, Lindner J, et al. Long-term experience with implantable treprostinil pumps in pulmonary arterial hypertension. *Bratisl Lek Listy*. 2020;121(5):356-61.
99. Mugii N, Someya F, Noto S, Hamaguchi Y, Matsushita T, Takehara K. Availability of EuroQol-5-Dimensions-5-Level (EQ-5D-5L) as health-related QOL assessment for Japanese systemic sclerosis patients. *Modern Rheumatology*. 2020;30(4):681-6.
100. Ivarsson B, Hesselstrand R, Rådegran G, Kjellström B. Health-related quality of life, treatment adherence and psychosocial support in patients with pulmonary arterial hypertension or chronic thromboembolic pulmonary hypertension. *Chronic respiratory disease*. 2018;16:1479972318787906.
101. Bussotti M, Gremigni P, Pedretti R, Kransinska P, Di Marco S, Corbo P, et al. Effects of an outpatient service rehabilitation programme in patients affected by pulmonary arterial hypertension: an observational study.. *Cardiovascular & Haematological Disorders-Drug Targets (Formerly Current Drug Targets-Cardiovascular & Hematological Disorders)*. 2017;17(1):3-10.
102. Davies EW, Llewellyn S, Beaudet A, Kosmas CE, Gin-Sing W, Doll HA. Elicitation of health state utilities associated with the mode of administration of drugs acting on the prostacyclin pathway in pulmonary arterial hypertension. *Patient preference and adherence*. 2018:1079-88.
103. Safdar Z, Tamez E, Frost A, Guffey D, Minard CG, Entman ML. Collagen metabolism biomarkers and health related quality of life in pulmonary arterial hypertension. *International journal of cardiovascular research*. 2015;4(2).
104. Awerbach J, Small M, Scott M, Harley J, Worden C, Doad G, Panjabi S. PCR223 Functional Class and Health-Related Quality of Life in Patients With Pulmonary Arterial Hypertension Associated With Congenital Heart Disease: Findings From a Real-World Study in the US. *Value in Health*. 2022;25(12):S433.
105. Pineda T, Bennett A, Bluro IM, Harley J, Hernandez-Oropeza JLJ, Lescano A, et al. WORLD HEALTH ORGANIZATION FUNCTIONAL CLASS AND HEALTH-RELATED QUALITY OF LIFE IN PATIENTS WITH PULMONARY ARTERIAL HYPERTENSION: REAL-WORLD FINDINGS IN LATIN AMERICA. *Chest*. 2022;162(4):A2323-A4.
106. Vizza CD, Klok R, Harley J, Small M, White RJ, Lautsch D. Health related quality of life in pulmonary arterial hypertension in the US, Europe and Japan. *Eur Respiratory Soc*; 2023.
107. Peñate GP, Parra NO, Morera JAD, Meñaca AM, Ramón ML, Menéndez SC, et al. Assessing the clinical benefit, safety, and patient-reported outcomes with the use of the PAHcare™ digital platform in pulmonary arterial hypertension: a pilot study. *Front Public Health*. 2024;12:1335072.
108. Szende A, Janssen B, Cabases J. Self-reported population health: an international perspective based on EQ-5D. 2014.
109. A Study to Evaluate Efficacy and Safety of Oral BAY63-2521 in Patients With Pulmonary Arterial Hypertension (PAH) [Internet]. 2008 [cited 25th March 2024]. Available from: <https://ClinicalTrialsgov/show/NCT00810693>.
110. BAY63-2521:Long-term Extension Study in Patients With Pulmonary Arterial Hypertension [Internet]. 2009 [cited 25th March 2024]. Available from: <https://clinicaltrials.gov/study/NCT00863681>.

111. Bunclark K, Doughty N, Michael A, Abraham N, Ali S, Cannon JE, et al. A minimal clinically important difference measured by the Cambridge Pulmonary Hypertension Outcome Review for patients with idiopathic pulmonary arterial hypertension. *Pulmonary circulation*. 2021;11(2):2045894021995055.
112. Speich R, Ulrich S, Domenighetti G, Huber LC, Fischler M, Treder U, Breitenstein A. Efficacy and safety of long-term imatinib therapy for pulmonary arterial hypertension. *Respiration*. 2015;89(6):515-24.
113. Small M, Piercy J, Pike J, Cerulli A. Incremental burden of disease in patients diagnosed with pulmonary arterial hypertension receiving monotherapy and combination vasodilator therapy. *Advances in therapy*. 2014;31:168-79.
114. Pellino K, Kerridge S, Church C, Peacock AJ, Crowe T, Jayasekera G, et al. Social deprivation and prognosis in Scottish patients with pulmonary arterial hypertension. *European Respiratory Journal*. 2018;51(2).
115. Hendriks PM, van Thor MC, Wapenaar M, Chandoesing P, van den Toorn LM, van den Bosch AE, et al. The longitudinal use of EmPHasis-10 and CAMPHOR questionnaire health-related quality of life scores in patients with pulmonary arterial hypertension and chronic thromboembolic pulmonary hypertension. *Respiratory Medicine*. 2021;186:106525.
116. Keen C, Smith I, Hashmi-Greenwood M, Sage K, Kiely DG. Pulmonary Hypertension and Measurement of Exercise Capacity Remotely: Evaluation of the 1-min Sit-to-Stand Test (PERSPIRE)—a cohort study. *ERJ Open Research*. 2023;9(1).
117. Lewis RA, Thompson AR, Billings CG, Charalampopoulos A, Elliot CA, Hamilton N, et al. Mild parenchymal lung disease and/or low diffusion capacity impacts survival and treatment response in patients diagnosed with idiopathic pulmonary arterial hypertension. *European Respiratory Journal*. 2020;55(6).
118. Alabed S, Dwivedi K, Durrington C, Alandajani F, Condliffe R, Elliot C, et al. P37 Correlation of emPHasis-10 with clinical tests: insights from the ASPIRE registry. *BMJ Publishing Group Ltd*; 2022.
119. Constantine A, Condliffe R, Clift P, Jansen K, Wort S, Moledina S, Dimopoulos K. Macitentan for pulmonary arterial hypertension related to repaired congenital heart disease: real-world UK experience. *European Heart Journal*. 2022;43(Supplement_2):ehac544. 1928.
120. Lewis RA, Armstrong I, Bergbaum C, Brewis MJ, Cannon J, Charalampopoulos A, et al. EmPHasis-10 health-related quality of life score predicts outcomes in patients with idiopathic and connective tissue disease-associated pulmonary arterial hypertension: results from a UK multicentre study. *European Respiratory Journal*. 2021;57(2).
121. Arvanitaki A, Mouratoglou SA, Evangeliou A, Grosomanidis V, Hadjimiliadiades S, Skoura L, et al. Quality of life is related to haemodynamics in precapillary pulmonary hypertension. *Heart, Lung and Circulation*. 2020;29(1):142-8.
122. Lawrie A, Hamilton N, Wood S, Exposto F, Muzwidzwa R, Raiteri L, et al. Association of risk assessment at diagnosis with healthcare resource utilization and health-related quality of life outcomes in pulmonary arterial hypertension. *Pulm Circ*. 2024;14(3):e12399.
123. Hernandez Alava M, Pudney S, Wailoo A. Estimating the Relationship Between EQ-5D-5L and EQ-5D-3L: Results from a UK Population Study. *Pharmacoeconomics*. 2023;41(2):199-207.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension
[ID6163]

© Merck Sharp & Dohme (UK) Limited (2025). All rights reserved

124. Di Tanna GL, Urbich M, Wirtz HS, Potrata B, Heisen M, Bennison C, et al. Health State Utilities of Patients with Heart Failure: A Systematic Literature Review. *Pharmacoeconomics*. 2021;39(2):211-29.
125. McMurray JJV, Trueman D, Hancock E, Cowie MR, Briggs A, Taylor M, et al. Cost-effectiveness of sacubitril/valsartan in the treatment of heart failure with reduced ejection fraction. *Heart*. 2018;104(12):1006-13.
126. Mott DJ, Schirmacher H, Al-Janabi H, Guest S, Pennington B, Scheuer N, et al. Modelling Spillover Effects on Informal Carers: The Carer QALY Trap. *Pharmacoeconomics*. 2023;41(12):1557-61.
127. Montpart Viñas F, Pomares E, Armengol Escolà J, Fernández Delgado M, Wu B, Classi P. EE46 High Doses of Treprostinil Could Entail Lower Hospitalization Costs in Patients With Pulmonary Arterial Hypertension in Nine European Countries. *Value in Health*. 2022;25(12):S62.
128. Sabashnikov A, Mohite PN, Zeriuoh M, Zych B, García-Sáez D, Maier J, et al. The role of extracorporeal life support in the management with severe idiopathic pulmonary artery hypertension undergoing lung transplantation: Are those patients referred too late? *J Thorac Dis*. 2019;11.
129. National Health Service. National Audit of Pulmonary Hypertension Great Britain 2022-2023. 2022-2023.
130. MSD. Winrevair® MHRA SmPC 60 mg powder and solvent for solution for injection. Dec 2024.
131. MSD. Winrevair® MHRA SmPC 60 mg powder for solution for injection.; Dec 2024.
132. MSD. Winrevair® MHRA SmPC 45 mg powder and solvent for solution for injection. Dec 2024.
133. MSD. Winrevair® MHRA SmPC 45 mg powder for solution for injection.; Dec 2024.
134. Jones K, Weatherly H, Birch S, Castelli A, Chalkley M, Dargan A, et al. Unit Costs of Health and Social Care 2023 Manual. 2024.
135. MSD [Data on File]. Interim CSR: A Phase 3, Randomized, Double-Blind, Placebo-Controlled Study to Compare the Efficacy and Safety of Sotatercept Versus Placebo When Added to Background Pulmonary Arterial Hypertension (PAH) Therapy for the Treatment of PAH. 2022.
136. Wailoo A. TECHNICAL SUPPORT DOCUMENT 23: A GUIDE TO CALCULATING SEVERITY SHORTFALL FOR NICE EVALUATIONS. School of Health and Related Research (SchHARR), University of Sheffield, Sheffield, UK; 2024 January.
137. MSD. [Data on File] Patient weight data from STELLAR trial. 2023.
138. Office of National Statistics (ONS). National life tables – life expectancy in the UK: 2018 to 2020 2021 [Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/nationallifetablesunitedkingdom/2018to2020>].
139. MSD. BARDS HTA Statistical Report - Protocol PAH 003: Four-stratum Risk Assessment: Descriptive Summaries and Transition Probabilities 2023.
140. MSD. BARDS HTA Report - Protocol PAH 003: Matching Adjusted Indirect Comparison (MAIC) of Sotatercept in STELLAR Study vs. Selexipag in GRIPHON Study. 2023.
141. NHS. National Schedule of NHS Costs 2023/24. 2024.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

142. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB, Force I-SMGRPT. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. Med Decis Making. 2012;32(5):733-43.

5 Appendices

- Appendix A: Summary of product characteristics (SmPC) and UK public assessment report
- Appendix B: Identification, selection and synthesis of clinical evidence
- Appendix C: Subgroup analysis
- Appendix D: Adverse reactions
- Appendix E: Published cost-effectiveness studies
- Appendix F: Health-related quality-of-life studies
- Appendix G: Cost and healthcare resource identification, measurement and valuation
- Appendix H: Clinical outcomes and disaggregated results from the model
- Appendix I: Price details of treatments included in the submission
- Appendix J: Current UK clinical commissioning policies in PAH
- Appendix K: STELLAR study additional methodological details
- Appendix L: STELLAR study additional statistical analysis details
- Appendix M: SOTERIA study additional methodological details
- Appendix N: SOTERIA study additional statistical analysis details
- Appendix O: Feasibility assessment of TRIPHIC data in indirect treatment comparison with STELLAR data – sotatercept versus selexipag
- Appendix P: Feasibility assessment and indirect treatment comparison of the clinical effectiveness of sotatercept-containing versus selexipag-containing regimens
- Appendix Q: Feasibility assessment for indirect treatment comparison of the clinical effectiveness of sotatercept-containing regimens versus infused prostacyclin
- Appendix R: Clinical expert validation elicitation

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Summary of Information for Patients (SIP)



March 2025

File name	Version	Contains confidential information	Date
ID6163_Sotatercept_SIP	1.0	No	10-MAR-2025

Summary of Information for Patients (SIP):

The pharmaceutical company perspective

What is the SIP?

The Summary of Information for Patients (SIP) is written by the company who is seeking approval from NICE for their treatment to be sold to the NHS for use in England. It is a plain English summary of their submission written for patients participating in the evaluation. It is not independently checked, although members of the public involvement team at NICE will have read it to double-check for marketing and promotional content before it is sent to you.

The **Summary of Information for Patients** template has been adapted for use at NICE from the [Health Technology Assessment International – Patient & Citizens Involvement Group](#) (HTAi PCIG). Information about the development is available in an open-access [JTAHC journal article](#)

SECTION 1: Submission summary

1a) Name of the medicine (generic and brand name):

Generic Name: Sotatercept

Brand name: Winrevair

1b) Population this treatment will be used by. Please outline the main patient population that is being appraised by NICE:

Sotatercept is initiated in adults with pulmonary arterial hypertension (PAH, World Health Organization [WHO] Group 1), who are at intermediate-low risk according to the European Society of Cardiology (ESC)/European Respiratory Society (ERS) four-strata risk rating system.

1c) Authorisation: Please provide marketing authorisation information, date of approval and link to the regulatory agency approval. If the marketing authorisation is pending, please state this, and reference the section of the company submission with the anticipated dates for approval.

This submission is based on a subset of the company's full marketing authorisation which is:

Winrevair, in combination with other pulmonary arterial hypertension (PAH) therapies, is indicated for the treatment of PAH in adult patients with WHO Functional Class (FC) II to III, to improve exercise capacity.

The marketing authorisation for sotatercept in the indication detailed in this submission was approved by the Medicines and Healthcare products Regulatory Agency (MHRA) on 27-DEC-2024.

1d) Disclosures. Please be transparent about any existing collaborations (or broader conflicts of interest) between the pharmaceutical company and patient groups relevant to the medicine. Please outline the reason and purpose for the engagement/activity and any financial support provided:

None.

SECTION 2: Current landscape

2a) The condition – clinical presentation and impact

Please provide a few sentences to describe the condition that is being assessed by NICE and the number of people who are currently living with this condition in England.

Please outline in general terms how the condition affects the quality of life of patients and their families/caregivers. Please highlight any mortality/morbidity data relating to the condition if available. If the company is making a case for the impact of the treatment on carers this should be clearly stated and explained.

What is Pulmonary Arterial Hypertension (PAH)?

Pulmonary hypertension (PH) is a condition characterised by high blood pressure in the vessels supplying the lungs, which can result from various underlying health issues. A specific type of PH is PAH, a rare and serious condition that affects the blood vessels leading from the heart to the lungs. In PAH, the pulmonary arteries become narrow, stiff, or blocked due to the thickening and narrowing of the blood vessel walls. This restricts blood flow, increases pressure within the arteries, and places added strain on the right side of the heart (1).

The symptoms of PAH typically have a gradual onset and worsen over time. Because oxygen is not getting around the body properly, people can often suffer with breathlessness. Other common symptoms include fatigue, chest pain, a racing heartbeat, dizziness, fainting, and swelling in the legs and ankles (2).

The exact cause of PAH is often unknown, and it can occur on its own or as a result of an underlying condition such as connective tissue disorders, heart defects, liver disease, or chronic lung diseases. Some medications and genetic factors may also contribute to the development of PAH (3).

How many people have PAH in the United Kingdom (UK)?

The most recent National Health Services (NHS) National Pulmonary Hypertension Audit, published on 12th December 2024 recorded a total of 4,626 active patients (3,510 with an active referral) in the United Kingdom (UK) currently being treated for PAH (4).

What is the impact of PAH on a person's quality of life?

PAH can have a substantial impact on the quality of life of patients, caregivers and family members. In the UK's largest survey into the lived experience of PH conducted by Pulmonary Hypertension Association (PHA) UK in 2023, 96% of participants reported saying that PH has had an impact on their overall quality of life, with 56% saying it has had a "major impact" (5). Furthermore, 85% reported an impact on their mental or emotional wellbeing, 79% reported an impact on their ability to work or attend education, and 78% reported saying it has had an impact on their relationships or family.

One of the main symptoms of PAH is shortness of breath, which can occur even with minimal physical activity or during rest. This breathlessness can make it difficult for people to engage in activities they once enjoyed, such as walking, climbing stairs, or participating

in sports (6). It can lead to limitations in daily activities and a decreased ability to perform tasks that require physical exertion.

Fatigue is another common symptom of PAH. Individuals may experience a constant feeling of tiredness and lack of energy, which can impact their ability to carry out daily activities (2). Depending on the distance to their nearest PH centre or hospital, patients often spend considerable time travelling for medical appointments, often requiring a carer or arranged transportation.

What is the impact of PAH on carers?

The burden of PAH impacts not only the patients but also their caregivers. A survey conducted by the European Pulmonary Hypertension Association in 2015 found that over half (57%) of carers surveyed reported that caring for people with PAH had a very high impact (rating of 8–10 on a scale of 10) on their own daily lives (7). Some of the major impacts on carers include:

1. **Physical and emotional burden:** Caring for someone with PAH can be physically demanding, involving assistance with daily activities, transportation, and managing medical needs. Almost half of the respondents (43%) from the European Pulmonary Hypertension Association survey in 2015 reported that they find the extra tasks and responsibilities physically exhausting. This, combined with the emotional stress of seeing a loved one suffering, can take a toll on the carer's well-being.
2. **Financial strain:** The high costs associated with PAH treatment and the potential loss of income if the carer has to leave work or reduce hours can create financial hardship.
3. **Social isolation:** Carers may have to sacrifice social activities and relationships due to the time and energy required for caregiving responsibilities.
4. **Disruption of family life:** The demands of caregiving can disrupt normal family routines, roles, and relationships, leading to stress and conflict.
5. **Lack of respite:** Finding adequate respite care for a PAH patient can be challenging, leaving carers with little break from their caregiving duties.
6. **Career implications:** Carers may have to make career sacrifices, such as leaving jobs, reducing hours, or passing up promotions to accommodate caregiving needs.
7. **Health risks:** The physical and emotional strain of caregiving can increase the risk of carers developing their own health problems, such as depression, anxiety, or physical injuries from lifting and assisting the patient.

What is the prognosis of PAH?

The prognosis of PAH can vary based on several factors, including the underlying cause, the severity of the disease, the individual's response to treatment, and their unique characteristics.

The five-year survival for people with PAH in the UK is approximately 51% according to the NHS National PH Audit from 2024 (8).

2b) Diagnosis of the condition (in relation to the medicine being evaluated)

Please briefly explain how the condition is currently diagnosed and how this impacts patients. Are there any additional diagnostic tests required with the new treatment?

How is PAH is diagnosed?

Most people will have several tests before a formal diagnosis can be made. This typically starts with a detailed medical history and physical examination (9). This is often followed by several tests including (but not limited to):

- Echocardiogram: This is a non-invasive test that uses ultrasound to create images of the heart. It helps assess the structure and function of the heart and can provide signs of possible PAH.
- Pulmonary function tests & lung imaging scans: These tests assess the lung anatomy, function and capacity.
- Blood tests: Blood samples can be analysed to check for specific markers associated with strain on the heart or genetic mutations associated with PAH. These tests help identify underlying causes and rule out other conditions.
- Right heart catheterisation: This is an invasive procedure where a thin tube is inserted into a blood vessel and threaded into the heart. It measures the blood pressure in the heart and lungs and helps confirm the diagnosis of PAH.

Many of these tests are repeated during treatment to assess response to treatment, disease progression and monitor side effects (9).

Speed of diagnosis

Receiving a diagnosis of PAH can be a frustrating and lengthy process. A survey conducted by PHA UK in 2023 revealed that more than 51% of individuals living with PH experienced a delay of over a year in receiving a diagnosis after first noticing symptoms, while 24% waited more than three years for diagnosis. (5). Almost a third of people saw four or more doctors before receiving their diagnosis.

There are no additional diagnostic tests anticipated before starting sotatercept, however there is a requirement to monitor haemoglobin and platelet levels through a series of blood tests during the medicine initiation.

2c) Current treatment options:

The purpose of this section is to set the scene on how the condition is currently managed:

- What is the treatment pathway for this condition and where in this pathway the medicine is likely to be used? Please use diagrams to accompany text where possible. Please give emphasis to the specific setting and condition being considered by NICE in this review. For example, by referencing current treatment guidelines. It may be relevant to show the treatments people may have before and after the treatment under consideration in this SIP.
- Please also consider:
 - if there are multiple treatment options, and data suggest that some are more commonly used than others in the setting and condition being considered in this SIP, please report these data.
 - are there any drug–drug interactions and/or contraindications that commonly cause challenges for patient populations? If so, please explain what these are.

Management of PAH involves a multidisciplinary approach and includes medical treatments, lifestyle modifications, and regular monitoring (9).

Medical treatments include conventional treatments (often called supportive treatments) such as oxygen and diuretics, and targeted therapies. These targeted therapies can be categorised according to their mechanism of action (Table 1).

Table 1: Targeted therapies for PAH available for use in the UK

Targeted therapy type	Medication generic name
PDE5i	- Sildenafil - Tadalafil
ERA	- Ambrisentan - Bosentan - Macitentan
sGC	Riociguat
PRA	Selexipag
PCA	- Epoprostenol - Iloprost

ERA: endothelin receptor antagonist; PAH: pulmonary arterial hypertension; PCA: prostacyclin analogue; PDE5i: phosphodiesterase type 5 inhibitor; PRA: prostacyclin receptor agonist; sGC: soluble guanylate cyclase; UK: United Kingdom

The current treatment pathway for PAH patients in the UK reflects that recommended in the 2022 ESC/ERS guidelines (9). Their treatment algorithm centres on a patient's risk status (predictive of one-year risk of mortality), with a goal of achieving and maintaining low-risk status. Key factors that are used to determine the risk status include the WHO FC (a measure of the impact of symptoms on physical activity), 6 Minute Walking Distance (6MWD – a measure of how far someone can walk in 6 minutes), and N-terminal pro-brain natriuretic peptide (NT-proBNP)/brain natriuretic peptide (BNP – blood markers indicating the strain on the heart) levels.

According to guidelines, risk status is assessed at diagnosis using a 3-strata system, with patients classified as low-, intermediate-, or high-risk. Patients should be followed closely and re-evaluated after 3 to 6 months regardless of initial therapy choice. At follow-up, patients should be assessed using a 4-strata system: low, intermediate-low, intermediate-high, or high-risk (described in Table 2).

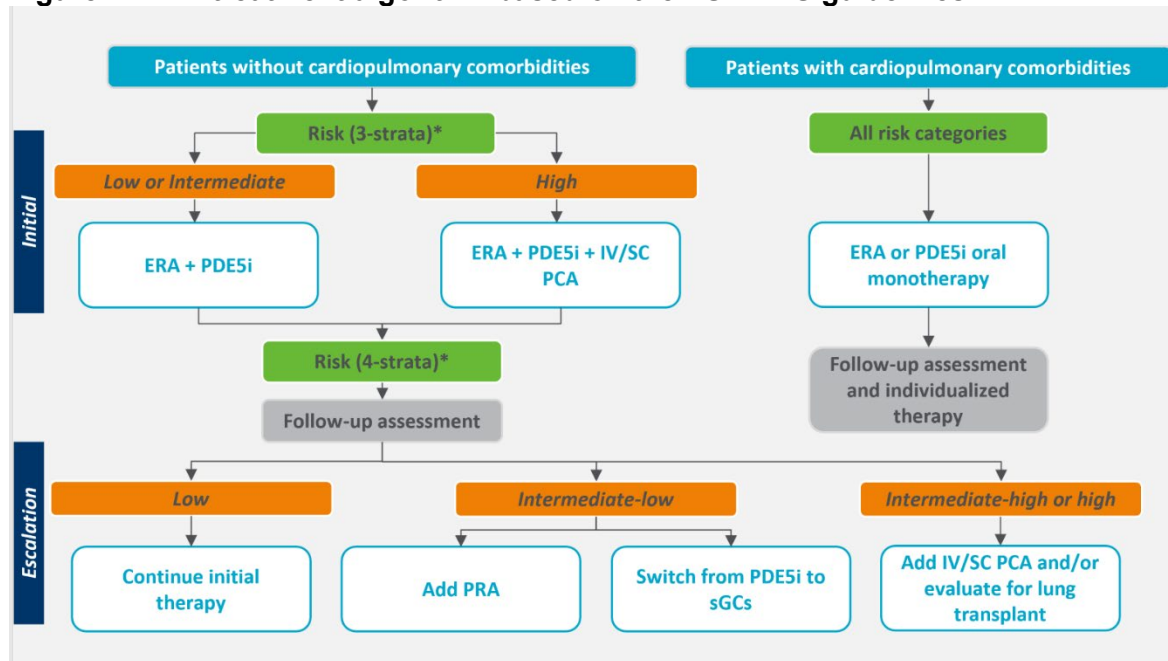
Table 2: 2022 ESC/ERS simplified four-strata risk-assessment tool for guidance of therapeutic decision-making

Determinants of prognosis	Low risk	Intermediate–low risk	Intermediate–high risk	High risk
Points assigned	1	2	3	4
WHO-FC	I or II*	-	III	IV
6MWD, m	>440	320–440	165–319	<165
BNP or NT-proBNP, ng/L	<50 <300	50–199 300–649	200–800 650–1100	>800 >1100

6MWD: 6-minute walk distance; BNP: brain natriuretic peptide; ESC: European Society of Cardiology; ERS: European Respiratory Society; FC: functional class; NT-proBNP: N-terminal pro-brain natriuretic peptide; WHO: World Health Organization
Risk is calculated by dividing the sum of all grades by the number of variables and rounding to the next integer.

The treatment pathway recommended in the 2022 ESC/ERS guidelines (and followed in the UK) is summarised in Figure 1.

Figure 1: PAH treatment algorithm based on the ESC/ERS guidelines



ERA: endothelin receptor antagonist; ERS: European Respiratory Society; ESC: European Society of Cardiology; IV: intravenous; PAH: pulmonary arterial hypertension; PCA: prostacyclin analogue; PDE5i: phosphodiesterase type-5 inhibitor; PRA: prostacyclin receptor antagonist; SC: subcutaneous; sGCs: soluble guanylate cyclase stimulator.

*Risk assessment using 3-strata model to determine risk stratification at diagnosis; 4-strata model based on determinants of prognosis.

Source: Humbert et al. 2022 (9).

The expected position of sotatercept in the treatment pathway is initiation in intermediate-low risk patients (according to the ERS/ESC four-strata system) who are currently treated with phosphodiesterase type-5 inhibitor (PDE5i) + endothelin receptor antagonist (ERA) + selexipag and instead should receive PDE5i + ERA + sotatercept to enable them to reach a low-risk state.

2d) Patient-based evidence (PBE) about living with the condition

Context:

- **Patient-based evidence (PBE)** is when patients input into scientific research, specifically to provide experiences of their symptoms, needs, perceptions, quality of life issues or experiences of the medicine they are currently taking. PBE might also include carer burden and outputs from patient preference studies, when conducted in order to show what matters most to patients and carers and where their greatest needs are. Such research can inform the selection of patient-relevant endpoints in clinical trials.

In this section, please provide a summary of any PBE that has been collected or published to demonstrate what is understood about **patient needs and disease experiences**. Please include the methods used for collecting this evidence. Any such evidence included in the SIP should be formally referenced wherever possible and references included.

Living with PAH involves a range of challenges and experiences that impact carers' and patients' daily lives. Patient based evidence research from people living with all PH types has identified several key areas of focus:

1. **Quality of life/daily activity:** Patients with PAH experience symptoms such as shortness of breath, fatigue, chest pain, and dizziness, which significantly affect their quality of life and daily activities. Data from the 2019 NHS National Audit of PH revealed that over 70% of people with PAH were at least WHO FC III, meaning

that the majority have marked limitation of physical activity, with less than ordinary activity causing undue breathlessness and fatigue (10). Moreover, 96% of people living with PH surveyed by the PHA UK in 2023 stated that the disease has had an impact on their overall quality of life, with 50% saying it has had a “major impact” (5).

2. **Psychological impact:** Living with PAH can lead to emotional distress, anxiety, and depression due to the chronic nature of the disease, uncertainty about the future, and limitations in physical activity. Data from the 2023 PHA UK survey recorded that 84% of people with PH say the condition has an impact on their mental or emotional wellbeing, with 34% saying it has a “major impact” (5). Similarly, a survey conducted by PHA UK in 2019 identified that 53% of people with PH have experience or been diagnosed with anxiety or depression (7). In addition, 78% of respondents reported that PH has had an impact on their relationships or family (11).
3. **Financial & social impact:** Living with PAH can also have a profound impact on personal finances. Data from the PHA UK 2023 survey recorded that 60% of people with PH reported that their condition had an impact on their financial situation, with 25% reporting a “major impact” (5). In addition, 79% of respondents reported that PH had an impact on their ability to attend work or education, with a number reporting that it caused them to have to stop working altogether (5). More specifically, 34% reported that their PH treatment and managements had a major impact on their ability to work or attend education.

When asked specifically about the outcomes most important to people living with PH, 52% of respondents from the 2023 PHA UK survey reported quality of life as the factor which matters to them most (5).

SECTION 3: The treatment

3a) How does the new treatment work?

What are the important features of this treatment?

Please outline as clearly as possible important details that you consider relevant to patients relating to the mechanism of action and how the medicine interacts with the body.

Where possible, please describe how you feel the medicine is innovative or novel, and how this might be important to patients and their communities.

If there are relevant documents which have been produced to support your regulatory submission such as a summary of product characteristics or patient information leaflet, please provide a link to these.

About sotatercept – its key features

Activin is a molecule that helps control cell growth and is found in higher-than-normal quantities in people with PAH (12). Activin may play a role in PAH by altering the normal growth and development of the pulmonary arterial walls.

Sotatercept is a new treatment for PAH that is designed to block Activin, thereby reducing the thickening of the pulmonary arterial walls. This treatment therefore aims to target a key underlying cause of the disease, while other PAH medications mainly work by relaxing and widening the blood vessels.

Sotatercept is an injection that needs to be administered under the skin in the stomach (abdomen) or upper thigh once every three weeks.

Sotatercept has already received marketing authorisation in Europe in combination with other PAH therapies, for the treatment of PAH in adult patients with WHO FC II to III, to

improve exercise capacity. Respective summary of product characteristics and information for patients can be found here: [winrevair-epar-product-information_en.pdf](#) (12)

3b) Combinations with other medicines

Is the medicine intended to be used in combination with any other medicines?

- Yes / No

If yes, please explain why and how the medicines work together. Please outline the mechanism of action of those other medicines so it is clear to patients why they are used together.

If yes, please also provide information on the availability of the other medicine(s) as well as the main side effects.

If this submission is for a combination treatment, please ensure the sections on efficacy (3e), quality of life (3f) and safety/side effects (3g) focus on data that relate to the combination, rather than the individual treatments.

Sotatercept is used in combination with background PAH therapy.

Sotatercept is intended to be used in combination with other background targeted PAH therapies including PDE5i and ERAs. Whilst these other therapies work to dilate the blood vessels in the lungs, sotatercept works differently by aiming to target activin and thereby reduce thickening of the pulmonary arterial walls.

Since these treatments work through different mechanisms in the body, using them together may provide added benefits for patients. This combined effect was shown in the STELLAR phase 3 clinical trial (13).

ERA and PDE5i therapies are widely available throughout treating centres in the UK. A list of their mechanisms of action and main side effects are outlined below, however this is not exhaustive. The full list can be found in the patient information leaflets (14-18).

PDE5 inhibitors (e.g. sildenafil, tadalafil)

Mechanism of action: PDE5i work by blocking an enzyme called phosphodiesterase type 5 (PDE5), which is found in the walls of blood vessels. PDE5 normally regulates blood flow in the pulmonary arteries. When PDE5 is blocked, these medications help the blood vessels to relax. This relaxation allows for better blood flow to the lungs, which can help reduce blood pressure.

Main side effects:

- Headache
- Flushing
- Indigestion
- Nasal congestion
- Visual disturbances (e.g., blurred vision, changes in colour perception)

Endothelin receptor antagonists (e.g., bosentan, ambrisentan, macitentan):

Mechanism of action: Endothelin receptor antagonists work by blocking the effects of endothelin, a substance that normally causes blood vessels to constrict or tighten. By preventing this constriction, these medications help blood vessels to stay relaxed, which improves blood flow.

Main side effects:

- Swelling of the ankles and legs
- Nasal congestion
- Low blood iron levels
- Liver enzyme abnormalities

- Headache
- Flushing

3c) Administration and dosing

How and where is the treatment given or taken? Please include the dose, how often the treatment should be given/taken, and how long the treatment should be given/taken for.

How will this administration method or dosing potentially affect patients and caregivers? How does this differ to existing treatments?

Sotatercept is an injection that is administered under the skin once every three weeks. People taking sotatercept will need to continue taking it as long as they receive clinical benefit from the medication as directed by their medical team.

Sotatercept will be supplied as a dry powder. It will need to be reconstituted with sterile water, followed by injection under the skin in the stomach (abdomen) or upper thigh. The volume of the injection to be administered will depend on the weight of the person receiving the medicine.

The initial dose of sotatercept is 0.3mg/kg. The maintenance doses are at 0.7mg/kg.

It is intended that people taking sotatercept will be trained on how to self-administer at home by their healthcare team. Training materials will be provided to support with this.

By self-administering the injections at home, patients will not have to travel to their centre of care to receive sotatercept every three weeks. Care givers can also receive training on how to administer the injection providing extra support for patients at home.

3d) Current clinical trials

Please provide a list of completed or ongoing clinical trials for the treatment. Please provide a brief top-level summary for each trial, such as title/name, location, population, patient group size, comparators, key inclusion and exclusion criteria and completion dates etc. Please provide references to further information about the trials or publications from the trials.

Phase 2 Trials

PULSAR - NCT03496207 (completed) (19)

PULSAR was a 24-week clinical trial, in which 106 adults who were receiving background therapy for PAH received subcutaneous sotatercept or placebo. The main outcome was the change from baseline to week 24 in pulmonary vascular resistance.

SPECTRA - NCT03738150 (completed) (20)

SPECTRA was a 24 week phase 2 exploratory study that evaluated the effects of sotatercept on exercise testing in participants with PAH. The main outcome was the change in oxygen uptake from baseline to week 24. Magnetic resonance imaging (MRI) was performed to assess the heart function. 21 participants completed the full 24 weeks of treatment.

Phase 3 Trials

STELLAR - NCT04576988 (completed) (13)

STELLAR was a phase 3 clinical trial in which adults with PAH who were receiving stable background therapy were randomly assigned to receive subcutaneous sotatercept or placebo every 3 weeks. The main outcome was the change in the 6-minute walk distance

after 24 weeks. Other outcomes included change in WHO FC, measures of heart function and pulmonary pressure, as well as patient reported outcomes.

ZENITH - NCT04896008 (ongoing – but decision to stop the trial early was made in November 2024)

This is a phase 3 study designed to evaluate sotatercept when added to maximum tolerated background PAH therapy compared to placebo. The main outcome is time to first event of all-cause death, lung transplantation, or PAH worsening related hospitalization of ≥ 24 hours. The estimated enrolment number is 166 participants.

HYPERION - NCT04811092 (ongoing – but decision to stop the trial early was made in January 2025)

This is a phase 3 study to evaluate sotatercept when added to background PAH therapy in newly diagnosed patients compared to placebo. The estimated enrolment number is 444 participants.

SOTERIA - NCT04796337 (ongoing – completion estimated September 2027)

This study is being conducted to assess the long-term safety, tolerability, and efficacy of sotatercept in participants with PAH. There is no control group.

3e) Efficacy

Efficacy is the measure of how well a treatment works in treating a specific condition.

In this section, please summarise all data that demonstrate how effective the treatment is compared with current treatments at treating the condition outlined in section 2a. Are any of the outcomes more important to patients than others and why? Are there any limitations to the data which may affect how to interpret the results? Please do not include academic or commercial in confidence information but where necessary reference the section of the company submission where this can be found.

A narrative summary of the clinical outcomes of the completed phase 2/3 clinical trials with sotatercept has been outlined below.

PULSAR (19)

Treatment with sotatercept resulted in a reduction in pulmonary vascular resistance in patients receiving background therapy for pulmonary arterial hypertension. The difference between the sotatercept 0.3-mg group and the placebo group in the change from baseline to week 24 in pulmonary vascular resistance was $-145.8 \text{ dyn}\cdot\text{sec}\cdot\text{cm}^{-5}$. Sotatercept was also associated with a significant decrease in N-terminal pro-B-type natriuretic peptide levels, as well as an increase in 6-minute walking distance.

STELLAR (13)

After 24 weeks, patients taking sotatercept could walk an average of 34.4 meters more in six minutes compared to just 1.0 meter for those on placebo. The overall difference between the two groups was 40.8 meters (with a confidence interval of 27.5 to 54.1 meters; $P < 0.001$), showing significant improvement for the sotatercept group.

Additionally, sotatercept led to improvements in several other health measures compared to the placebo. These included reduced lung blood vessel resistance, lower levels of NT-proBNP, improved WHO functional class, longer time without clinical worsening or death, and improved scores on the Pulmonary Arterial Hypertension-Symptoms and Impact (PAH-SYMPACT®) Physical Impacts and Cardiopulmonary Symptom scales. Eight out of nine of these secondary measures showed significant improvement with sotatercept.

3f) Quality of life impact of the medicine and patient preference information

What is the clinical evidence for a potential impact of this medicine on the quality of life of patients and their families/caregivers? What quality of life instrument was used? If the EuroQol-5D (EQ-5D) was used does it sufficiently capture quality of life for this condition? Are there other disease specific quality of life measures that should also be considered as supplementary information?

Please outline in plain language any quality of life related data such as **patient reported outcomes (PROs)**.

Please include any **patient preference information (PPI)** relating to the drug profile, for instance research to understand willingness to accept the risk of side effects given the added benefit of treatment. Please include all references as required.

STELLAR (13)

The STELLAR trial aimed to capture changes in quality of life measures using the PAH-SYMPACT® tool. The PAH-SYMPACT® tool is a validated patient-reported outcome measure specifically designed for individuals with PAH (21). It assesses the impact of PAH on peoples' daily lives, focusing on symptoms and their effects on physical and emotional well-being. The tool includes questions that cover:

- Physical impacts: Evaluates how PAH affects physical abilities and activities of daily living.
- Cardiopulmonary symptoms: Measures the frequency and severity of symptoms such as shortness of breath, chest pain, and fatigue.
- Emotional well-being: Assesses the emotional and psychological impact of living with PAH.

In the STELLAR trial, changes from baseline at week 24 in the Physical Impacts and Cardiopulmonary Symptoms domain scores showed greater improvements with sotatercept than with placebo. No significant between-group difference was observed for the PAH-SYMPACT® Cognitive/Emotional Impacts domain score.

3g) Safety of the medicine and side effects

When NICE appraises a treatment, it will pay close attention to the balance of the benefits of the treatment in relation to its potential risks and any side effects. Therefore, please outline the main side effects (as opposed to a complete list) of this treatment and include details of a benefit/risk assessment where possible. This will support patient reviewers to consider the potential overall benefits and side effects that the medicine can offer.

Based on available data, please outline the most common side effects, how frequently they happen compared with standard treatment, how they could potentially be managed and how many people had treatment adjustments or stopped treatment. Where it will add value or context for patient readers, please include references to the Summary of Product Characteristics from regulatory agencies etc.

What are the possible side effects of Sotatercept?

The most common side effects of sotatercept observed from the clinical trials are shown in Table 3. Very common is defined as occurring in at least 1 in 10 people, and common is defined as occurring in between 1 in 10 to 1 in 100 people.

Table 3: Very common and common adverse events of sotatercept

System organ class	Frequency	Adverse reaction
Blood and lymphatic system disorders	Very common	Thrombocytopenia (low platelet count) Increased haemoglobin
Nervous system disorders	Very common	Dizziness

		Headache
Respiratory, thoracic and mediastinal disorders	Very common	Epistaxis (nose bleeds)
Gastrointestinal disorders	Very common	Diarrhoea
	Common	Gingival (gum) bleeding
Skin and subcutaneous tissue disorders	Very common	Telangiectasia (condition where small blood vessels near the surface of the skin become widened and visible. These vessels can look like red or purple spider webs or lines).
		Rash
	Common	Erythema (redness)
General disorders and administration site conditions	Common	Injection site pruritus (itching)
Investigations	Common	Increased blood pressure

Serious side effects may include those listed in Table 4.

Table 4: Serious side effects that may be caused by taking sotatercept

Side effect	Description
High haemoglobin (a protein in red blood cells that carries oxygen)	This can raise the chance of a blood clot forming that can block a blood vessel. Blood tests are therefore taken regularly when starting sotatercept, and periodically on an ongoing monitoring basis. In the STELLAR study, adverse reactions of increased Hgb were reported in 8.6% of patients taking sotatercept.
Low platelet count (blood components that help clotting)	This can cause symptoms such as easy bruising, continued bleeding from cuts, and nosebleeds. Blood tests are therefore taken regularly when starting sotatercept, and periodically on an ongoing monitoring basis. Thrombocytopenia ('thrombocytopenia' and 'platelet count decreased') was reported in 10.4% of patients taking sotatercept in clinical studies and was managed by dose adjustments.
Decreased fertility	Sotatercept may decrease female and male fertility.
Serious bleeding	Signs and symptoms may include persistent headache, nausea, weakness, black or tarry stool, blood in your stool, bright red blood from vomiting or coughing, persistent abdominal cramps, severe back pain, or abnormally heavy menstrual bleeding.

3h) Summary of key benefits of treatment for patients

Issues to consider in your response:

- Please outline what you feel are the key benefits of the treatment for patients, caregivers and their communities when compared with current treatments.
- Please include benefits related to the mode of action, effectiveness, safety and mode of administration.

Key benefits of sotatercept for patients

Mechanism of action

- Targeted therapy: Sotatercept is a novel therapy which targets a key underlying cause of disease in PAH, helping to reduce the thickening of blood vessel walls. The mechanism of action is different from existing PAH medicines.

Effectiveness based on clinical trials:

- Better exercise capacity: Patients can walk further and feel less tired.
- Lower blood pressure in lungs: Sotatercept reduces the pressure in the lung arteries.
- Improved health markers: Lower levels of a heart-related protein indicate less strain on the heart.
- Better physical functioning: Many patients move to a better WHO functional class.
- Slower disease progression: Sotatercept helps keep the disease from getting worse for a longer time.

Safety:

- Side effects: Common side effects include mild reactions at the injection site, tiredness, headaches, and dizziness.
- Serious side effects: They are rare and manageable with regular check-ups.
- Regular monitoring: Helps to manage and minimise side effects.

Administration:

- Home injections: Sotatercept is given as an injection under the skin, which can be done at home following suitable training.

3i) Summary of key disadvantages of treatment for patients

Issues to consider in your response:

- Please outline what you feel are the key disadvantages of the treatment for patients, caregivers and their communities when compared with current treatments. Which disadvantages are most important to patients and carers?
- Please include disadvantages related to the mode of action, effectiveness, side effects and mode of administration
- What is the impact of any disadvantages highlighted compared with current treatments

While sotatercept can be effective for treating pulmonary arterial hypertension, it comes with potential side effects such as fatigue, nausea, and injection site reactions. Regular monitoring is required, which can be inconvenient, and the long-term safety is still under investigation.

3j) Value and economic considerations

Introduction for patients:

Health services want to get the most value from their budget and therefore need to decide whether a new treatment provides good value compared with other treatments. To do this they consider the costs of treating patients and how patients' health will improve, from feeling better and/or living longer, compared with the treatments already in use. The drug manufacturer provides this information, often presented using a health economic model.

In completing your input to the NICE appraisal process for the medicine, you may wish to reflect on:

- The extent to which you agree/disagree with the value arguments presented below (e.g., whether you feel these are the relevant health outcomes, addressing the unmet needs and issues faced by patients; were any improvements that would be important to you missed out, not tested or not proven?)
- If you feel the benefits or side effects of the medicine, including how and when it is given or taken, would have positive or negative financial implications for patients or their families (e.g., travel costs, time-off work)?
- How the condition, taking the new treatment compared with current treatments affects your quality of life.

MSD has submitted a health economic model for evaluation by NICE. The following sections describe how the model was structured and how it attempts to reflect the progression of PAH over time. The goal of the economic modelling is to quantify the differences in expected patient benefits and costs between initiating sotatercept in combination with dual therapy or selexipag in combination with dual therapy in patients with symptomatic PAH at intermediate-low risk. These differences are combined and used by NICE to determine whether sotatercept is cost-effective or not.

How the model reflects the condition

The model uses data from clinical trials to simulate what results would look like if the analysis was extended over a lifetime (30 years) in a real-world scenario.

The model starts with a close description of the clinical trials by using a “decision tree”, in which people with PAH at intermediate-low risk are allocated to sotatercept in combination with dual therapy or selexipag in combination with dual therapy. This is a common way of modelling a clinical trial period. For each treatment, people are assigned to one of four groups based on their risk category (low risk, intermediate-low risk, intermediate-high risk and high risk). At the end of the decision tree (24 weeks), outcomes for the people taking sotatercept in combination with dual therapy are assessed and, if they experience worsening of their cardiovascular function (assessed via the 6MWD, WHO FC or NT-proBNP levels), their treatment is stopped in the model and they revert back to what they were taking before. This rule is put in place in the model to reflect what may happen in reality to ensure that the NHS does not continue to spend money and resources on treating people who would not benefit.

After the decision tree, people then move to the longer-term part of the model that uses the clinical trial results to predict the difference between what would happen if people were treated with sotatercept in combination with dual therapy compared to what would happen if they received selexipag in combination with dual therapy. In this part of the model, people can move from one category to another (low risk, intermediate-low risk, intermediate-high risk, high risk, post-lung/heart transplant and death) to simulate the evolution of PAH.

Modelling how much a treatment extends life

Trials collect information about a treatment for a limited time period. The trials for sotatercept were not able to show that sotatercept could extend life because this is too short a time to observe a difference in the death rate between treatments. PAH progression is known to lead to shorter survival based on data from published studies.

Based on these survival estimates, the model suggests that because sotatercept in combination with dual therapy improved cardiovascular function in patients with symptomatic PAH at intermediate-low risk in the clinical trials, life may be extended. This includes both the total number of years lived and the number of years lived in better health.

Modelling how much a treatment improves quality of life

Quality of life was measured during the clinical trials using the PAH-SYMPACT[®] tool (as described in section 3f) and the EuroQol 5-Dimension 5-Level (EQ-5D-5L; a standard questionnaire measuring general health). The EQ-5D-5L data were transformed into EuroQol 5-Dimension 3-Level (EQ-5D-3L) data to be used in the model, as this is NICE's preferred measure for quality of life in economic models.

People with more severe PAH (worse cardiovascular function) have worse quality of life. Therefore, the model assigns different levels of quality of life to each risk group (low risk, intermediate-low risk, intermediate-high risk, high risk). When a person experiences a hospitalisation or receives a lung/heart transplant, this leads to a small decrease in quality of life.

Because people experience improved cardiovascular function whilst taking sotatercept in combination with dual therapy, they experience better quality of life in the model than people treated with selexipag in combination with dual therapy.

Modelling how the costs of treatment differ with the new treatment

The model also compares the cost of treating people with sotatercept in combination with dual therapy to selexipag in combination with dual therapy. It calculates the differences in the cost of the medicines and the cost of NHS services such as medical appointments or hospital stays.

Nevertheless, the cost burden of PAH on patients, their caregivers and the NHS is substantial, and sotatercept has the potential to save money by preventing many of the costs that occur due to disease progression, poor quality of life and early death. This is because treatment with sotatercept in combination with dual therapy leads to better cardiovascular function and quality of life, and significantly slower disease progression (13). This could reduce medical visits and hospitalisations in people with PAH.

Uncertainty

Modelling the course of PAH and extrapolating data from clinical trials beyond the end of observed data (follow-up period) have the potential to create uncertainty in the predicted outcomes. Advice has been taken from UK clinical experts to confirm that the inputs of the models are representative of clinical practice. Where possible, the predicted outcomes have been checked against other available sources (real-world evidence and other clinical trials) to see if they are plausible and reflect the expected clinical reality. Specifically, these challenges include:

- Limitations of the analysis comparing sotatercept in combination with dual therapy to selexipag in combination with dual therapy in terms of risk category improvement
- Insufficient published evidence on the mortality of patients with PAH based on their risk categories, which makes it difficult to estimate their life expectancy
- Insufficient evidence on the long-term improvement of risk categories with sotatercept.

Additional factors: benefits of the treatment not captured in the model

Some health benefits of sotatercept could not be considered in the health economic model. These additional health benefits of sotatercept include:

- Reducing the burden on more than just one caregiver (family members)
- Allowing patients and their family to continue working
- Reducing the travel requirements to attend specialist appointments for patients by slowing down disease progression
- Increasing survival beyond the increase associated with improved risk status
- Psychological benefit from the availability of a new effective treatment
- Reducing NHS waiting list for transplant by reducing the need for lung/heart transplant in patients with PAH
- Improving the ability to exercise could lead to reducing the risk of other cardiovascular diseases
- Reducing the need to receive other medications like anticoagulants

3k) Innovation

NICE considers how innovative a new treatment is when making its recommendations.

If the company considers the new treatment to be innovative please explain how it represents a 'step change' in treatment and/ or effectiveness compared with current treatments. Are there any QALY benefits that have not been captured in the economic model that also need to be considered (see section 3f).

Sotatercept is a new therapeutic for PAH which targets a key underlying cause of the disease. This represents a step change in PAH, since prior therapy has primarily focussed on dilating the vessels.

3l) Equalities

Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged.

Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics.

More information on how NICE deals with equalities issues can be found in the NICE equality scheme.

Find more general information about the Equality Act and equalities issues [here](#)

No equity or equality considerations are anticipated.

SECTION 4: Further information, glossary and references

4a) Further information

Feedback suggests that patients would appreciate links to other information sources and tools that can help them easily locate relevant background information and facilitate their effective contribution to the NICE assessment process. Therefore, please provide links to any relevant online information that would be useful, for example, published clinical trial data, factual web content, educational materials etc.

Where possible, please provide open access materials or provide copies that patients can access.

Further information on NICE and the role of patients:

- Public Involvement at NICE [Public involvement | NICE and the public | NICE Communities | About | NICE](#)
- NICE's guides and templates for patient involvement in HTAs [Guides to developing our guidance | Help us develop guidance | Support for voluntary and community sector \(VCS\) organisations | Public involvement | NICE and the public | NICE Communities | About | NICE](#)
- EUPATI guidance on patient involvement in NICE: <https://www.eupati.eu/guidance-patient-involvement/>
- EFPIA – Working together with patient groups: <https://www.efpia.eu/media/288492/working-together-with-patient-groups-23102017.pdf>
- National Health Council Value Initiative. <https://nationalhealthcouncil.org/issue/value/>
- INAHTA: <http://www.inahta.org/>
- European Observatory on Health Systems and Policies. Health technology assessment - an introduction to objectives, role of evidence, and structure in Europe: http://www.inahta.org/wp-content/themes/inahta/img/AboutHTA_Policy_brief_on_HTA_Introduction_to_Objectives_Role_of_Evidence_Structure_in_Europe.pdf

4b) Glossary of terms

N/A

4c) References

Please provide a list of all references in the Vancouver style, numbered and ordered strictly in accordance with their numbering in the text:

1. Bousseau S, Sobrano Fais R, Gu S, Frump A, Lahm T. Pathophysiology and new advances in pulmonary hypertension. BMJ Med. 2023;2(1):e000137.
2. NHS. Pulmonary hypertension: National Health Service; 2023 [updated 02-MAY-2023. Available from: <https://www.nhs.uk/conditions/pulmonary-hypertension/>.
3. NHS. Pulmonary hypertension - Causes: National Health Service; 2023 [updated 02-MAY-2023. Available from: [https://www.nhs.uk/conditions/pulmonary-hypertension-causes/](https://www.nhs.uk/conditions/pulmonary-hypertension/causes/).
4. NHS. National Pulmonary Hypertension Audit, 15th Annual Report2024. Available from: <https://digital.nhs.uk/data-and-information/publications/statistical/national-pulmonary-hypertension-audit/15th-annual-report>.

5. PHA UK. Breathless, not voiceless: What it means to live with PH today2023 2024. Available from: <https://www.phauk.org/research/pha-uk-led-research/breathless-not-voiceless-what-it-means-to-live-with-ph-today/>.
6. PHA Europe. The impact of pulmonary arterial hypertension (PAH) on the lives of patients and carers: results from an international survey2012. Available from: https://www.phaeurope.org/wp-content/uploads/PAH_Survey_FINAL.pdf.
7. PHA UK. The true emotional impact of pulmonary hypertension2019 2024. Available from: <https://www.phauk.org/living-with-pulmonary-hypertension/the-emotional-impact-of-ph-survey-results/>.
8. NHS. National Audit of Pulmonary Hypertension, 14th Annual Report2023 2024. Available from: <https://digital.nhs.uk/data-and-information/publications/statistical/national-pulmonary-hypertension-audit/14th-annual-report>.
9. Humbert M, Kovacs G, Hoeper MM, Badagliacca R, Berger RMF, Brida M, et al. 2022 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension. Eur Heart J. 2022;43(38):3618-731.
10. NHS. National Audit of Pulmonary Hypertension Great Britain, 2018-19 (10th annual report). 2019 October, 24.
11. PHA UK. What it means to live with PH today2017. Available from: <https://www.phauk.org/research-survey-work/living-ph-report/>.
12. Bogaard HJ, Goncharova EA, de Jesus Perez VA. Editorial: Pulmonary hypertension in the modern era: Science and clinical practice, Volume II. Front Med (Lausanne). 2023;10:1136157.
13. Hoeper MM, Badesch DB, Ghofrani HA, Gibbs JSR, Gombert-Maitland M, McLaughlin VV, et al. Phase 3 Trial of Sotatercept for Treatment of Pulmonary Arterial Hypertension. N Engl J Med. 2023;388(16):1478-90.
14. EMC. Bosentan 62.5 mg Film-coated Tablets: Electronic Medicines Compendium; 2021 [updated 20-SEP-2021. Available from: <https://www.medicines.org.uk/emc/product/820/smpc>.
15. EMC. Tadalafil 10mg Film Coated Tablets: Electronic Medicines Compendium; 2022 [updated 24-MAR-2022. Available from: <https://www.medicines.org.uk/emc/product/13431/smpc>.
16. EMC. Sildenafil 25 mg film-coated tablets: Electronic Medicines Compendium; 2023 [updated 24-AUG-2023. Available from: <https://www.medicines.org.uk/emc/product/15025/smpc>.
17. EMC. Opsumit 10mg film coated tablets (Great Britain): Electronic Medicines Compendium; 2023 [updated 15-MAY-2023. Available from: <https://www.medicines.org.uk/emc/product/5223/smpc>.
18. EMC. Ambrisentan 5 mg film-coated Tablets: Electronic Medicines Compendium; 2024 [updated 29-FEB-2024. Available from: <https://www.medicines.org.uk/emc/product/11696/smpc>.
19. Humbert M, McLaughlin V, Gibbs JSR, Gombert-Maitland M, Hoeper MM, Preston IR, et al. Sotatercept for the treatment of pulmonary arterial hypertension. New England Journal of Medicine. 2021;384(13):1204-15.
20. Waxman AB, Systrom DM, Manimaran S, de Oliveira Pena J, Lu J, Rischard FP. SPECTRA Phase 2b Study: Impact of Sotatercept on Exercise Tolerance and Right Ventricular Function in Pulmonary Arterial Hypertension. Circ Heart Fail. 2024;17(5):e011227.
21. McCollister D, Shaffer S, Badesch DB, Filusch A, Hunsche E, Schuler R, et al. Development of the Pulmonary Arterial Hypertension-Symptoms and Impact (PAH-SYMPACT(R)) questionnaire: a new patient-reported outcome instrument for PAH. Respir Res. 2016;17(1):72.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal

Sotatercept for pulmonary hypertension [ID6163]

Clarification questions

March, 2025

File name	Version	Contains confidential information	Date
ID6163 Sotatercept EAG clarification letter to NICE – MSD response (17- APR-2025)	1.0	Yes	17-APR-2025

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Notes for company

Highlighting in the template

Square brackets and grey highlighting are used in this template to indicate text that should be replaced with your own text or deleted. These are set up as form fields, so to replace the prompt text in [grey highlighting] with your own text, click anywhere within the highlighted text and type. Your text will overwrite the highlighted section.

To delete grey highlighted text, click anywhere within the text and press DELETE.

Section A: Clarification on effectiveness data

A1. Appendix B, Section B1.1.1. Please clarify how many reviewers conducted the risk of bias assessment of the included studies and whether reviewers worked independently

MSD response:

The risk of bias assessment was conducted by two independent reviewers, with any disagreements resolved by a third senior reviewer.

A2. Appendix P, Section P.1.2.4. If available, can the company provide any ITC results using subgroups from the published GRIFFON trial that are more aligned with the intermediate-low risk subgroup of the STELLAR trial.

MSD response:

Subgroup data from the GRIPHON trial are available from the Sitbon et al. 2015 publication (1), the subgroups for which data are available are shown in the publication

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

(and also provided in Appendix P section P1.2.5 [Figure 6 - Figure 8] of the company evidence submission). From these it can be seen that:

- These do not include subgroups based on ESC/ERS risk status (which is expected in a study designed and conducted before the ESC/ERS guidelines on the management of pulmonary hypertension began recommending this form of risk stratification in 2022).
- Based on the company's discussion with clinical experts about the subjective nature of WHO functional class (FC), none of these subgroups presented in the publication reasonably constitute one that may be a justifiable proxy for ESC/ERS intermediate-low risk patients, as ESC/ERS risk status is based on a scoring system taking into account several patient characteristics (WHO FC, 6-minute walk distance [6MWD], and natriuretic peptide test results) (2), whereas GRIPHON results are only reported for subgroups based on differences in single variables/characteristics (1).

Therefore, data to inform an ITC using subgroups from the published GRIPHON trial that are more aligned with the intermediate-low risk subgroup of the STELLAR trial are not available.

Furthermore, as can be seen in forest plots of the subgroup analyses of the GRIPHON study and as noted by the GRIPHON investigators in the publication (1), treatment effect with respect to the key endpoints were consistent between all patients in the trial and the analysed subgroups, so it is unlikely that an ITC using any of the subgroup data will provide findings meaningfully more appropriate to the decision problem of this appraisal than the analyses based on data from all patients.

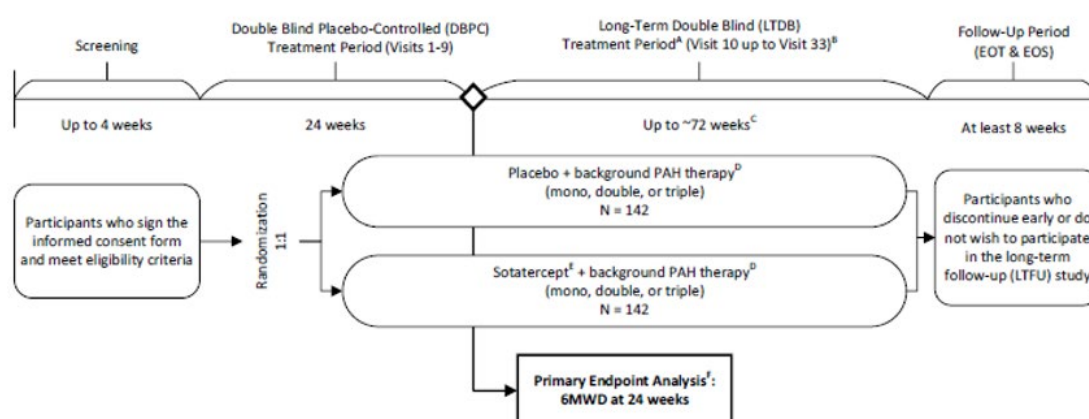
A3. Section 2.6.1. Please explain the flow of patients from STELLAR into SOTERIA. What is responsible for the attrition in numbers available to inform outcomes during the 72-week double blind phase of the STELLAR?

MSD response:

An overview of the STELLAR study design is illustrated in Figure 1 below.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Figure 1: STELLAR study design overview



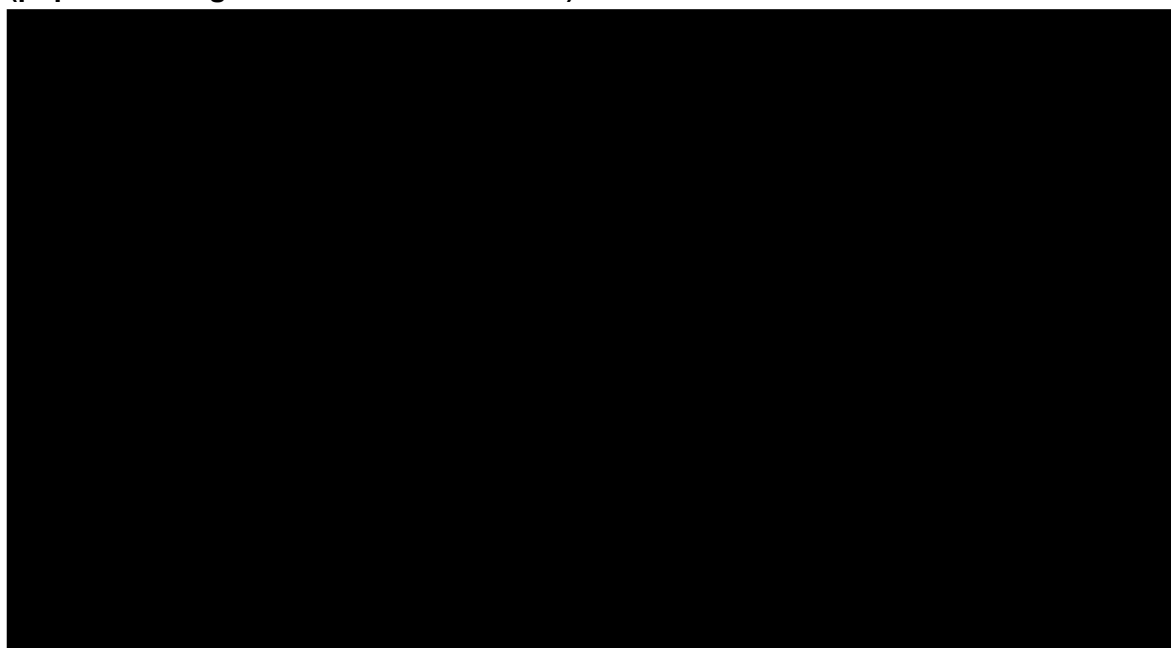
Participants were randomly assigned in a 1:1 ratio to receive, for the duration of the Double-Blind Placebo-Controlled (DBPC) and Long-Term Double-Blind (LTDB) treatment periods:

- Arm 1: Placebo administered subcutaneously (SC) every 21 days plus background PAH therapy
- Arm 2: Sotatercept at a starting dose of 0.3mg/kg with a target of 0.7mg/kg administered subcutaneously (SC) every 21 days plus background PAH therapy

DBPC and LTDB treatment periods were expected to be completed at 24 weeks and up to 72 weeks, respectively. SOTERIA is a phase 3, optional open-label long-term follow-up study i.e. STELLAR participants could choose to enrol in SOTERIA once they had completed 24 weeks of treatment in either arm of the parent study. The most common source of attrition in participant numbers observed during the LTDB phase of STELLAR is patients opting to switch to SOTERIA and is presented in the Kaplan-Meier curve shown in Figure 2. As would be expected, the rate of attrition is greater in the placebo arm, as entry to SOTERIA offered all participants the opportunity to receive the active treatment, and despite remaining blinded during the LTDB phase, patients in this arm may have been more motivated to enrol based on their clinical response in STELLAR.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Figure 2: Kaplan-Meier curve of time to discontinuation from the STELLAR study (population eligible to enrol in SOTERIA)



For a minority of patients, exit from STELLAR is attributable to other reasons, as outlined below:

323 patients were randomised in the STELLAR study:

- 163 patients randomised to the sotatercept arm.
- 160 patients randomised to the placebo arm.

307 of the 323 randomised patients completed 24 weeks of treatment in the DBPC phase of STELLAR and were eligible to roll over into the SOTERIA study:

- 159 of the 163 patients randomised to the sotatercept arm of STELLAR were eligible to roll over into the SOTERIA study:
 - 4 of the 163 patients randomised to the sotatercept arm of STELLAR discontinued the study:
 - 2 patients discontinued due to participant request (withdrawal of consent).
 - 1 patient discontinued due to participant unwillingness or inability to comply with the protocol.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

- 1 patient discontinued due to adverse event.
- 148 of the 160 patients randomised to the placebo arm of STELLAR were eligible to roll over into the SOTERIA study:
 - 12 of the 160 patients randomised to the sotatercept arm of STELLAR died or discontinued the study:
 - 5 patients died.
 - 3 patients discontinued due to participant request (withdrawal of consent).
 - 2 patients discontinued due to clinical worsening event.
 - 1 patient discontinued due to adverse event.
 - 1 patient discontinued due to participant unwillingness or inability to comply with the protocol.

297 of the 307 patients eligible to roll over into the SOTERIA study actually enrolled into SOTERIA:

- 154 of the 159 patients originally in the sotatercept arm eligible to roll over into the SOTERIA study actually enrolled into SOTERIA:
 - 5 of the 159 patients eligible to roll over into the SOTERIA study died or discontinued the LTDB phase of STELLAR:
 - 2 patients died.
 - 2 patients discontinued due to adverse event
 - 1 patient did not choose to enrol into SOTERIA.
- 143 of the 148 patients originally in the placebo arm eligible to roll over into the SOTERIA study actually enrolled into SOTERIA:
 - 5 of the 148 patients eligible to roll over into the SOTERIA study died or discontinued the LTDB phase of STELLAR:
 - 1 patient died.

- 3 patients discontinued due to participant request (withdrawal of consent).
- 1 patient discontinued due to participant unwillingness or inability to comply with the protocol.

A4. Section 2.6.2. Analysis of long-term effect of sotatercept in STELLAR through SOTERIA, Table 52. Table 52 looks like it refers to the whole cohort of SOTERIA, rather than those recruited from STELLAR. Please provide the table for the subgroup of interest.

MSD response:

A version of the table for the subgroup of SOTERIA patients whose parent study is STELLAR is provided in Table 1 below.

Table 1: SOTERIA study summary of clinical worsening events – SOTERIA participants whose parent study is STELLAR

	Placebo crossed (N=143)	Continued sotatercept (N=153)	All participants (N=296)
Participants with at least one clinical worsening event, n (%)			
Total no. of clinical worsening events	1	1	1
Median (range) time to first event, weeks			
Assessment of first clinical worsening event^a, n (%)			
Death			
Worsening-related listing for lung and/or heart transplant			
Initiation of rescue therapy or increase infusion prostacyclin dose by $\geq 10\%$			
Need for atrial septostomy			
PAH-related hospitalisation (≥ 24 h)			
Deterioration of PAH [†]			
Clinical worsening related to Adverse Event/SAE			

6MWD: 6-minute walking distance; FC: functional class; NR: not reported; PAH: pulmonary arterial hypertension; WHO: World Health Organization

SOTERIA cumulative data cutoff date: 08 November 2023.

^a No participant needed to have atrial septostomy as assessment of first clinical worsening event. A participant could have more than one assessment as the first clinical worsening event.

[†] Defined by both worsened WHO FC and decrease in 6MWD by $\geq 15\%$ (confirmed by two 6-minute walk tests ≥ 4 hours apart but ≤ 1 week), compared with baseline values, even if these two events began at different times.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Section B: Clarification on cost-effectiveness data

Economic model structure

B1. PRIORITY. Document B, Section 3.2.2, P.150-157. The EAG acknowledges the company's justification for developing a model with health states defined consistently with the ECS / ERS guidelines but also notes limitations with the approach taken to generate a comparative assessment using STELLAR data for the economic model. Please clarify:

- Whether it was possible to recreate the ECS / ERS health states using other studies from data available in the public domain.

MSD response:

The company explored multiple options for addressing the decision problem in this appraisal i.e. a comparison vs. selexipag for the treatment of intermediate-low risk PAH patients using a risk status-based cost-effectiveness model.

Patient transitions between ESC/ERS risk states for the sotatercept arm were estimated using patient-level data from the STELLAR study. Relevant studies reporting risk status change at comparable timepoints for selexipag were not identified by the literature review described in the CS, mainly due to the fact they reported results prior to the publication of the ESC/ERS risk status categories (2). Given the composite nature of the risk status classification, patient-level data for multiple endpoints at appropriate timepoints are required. Studies which reported the required level of granularity were not identified.

The feasibility of conducting an ITC of sotatercept vs. selexipag using real-world patient-level data on the risk status inputs for selexipag was assessed, as reported in Appendix O of the CS. The TRIPHIC research database was identified as a source of relevant individual data on patients with PAH treated with selexipag in the UK (other UK datasets were either unavailable or did not record data in an appropriate or complete format). The conclusion of the feasibility assessment was that relevant data were available for too few patients from the TRIPHIC research database to allow for a Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

credible and meaningful indirect comparison to be feasible. In conclusion, it was not possible to conduct the risk status-based cost-effectiveness analysis from available external data.

- If not, have the company attempted to obtain a categorisation of ECS / ERS states from relevant selexipag study authors that would enable an assessment of comparative effectiveness (via an ITC) to populate the economic model?

MSD response:

The relevant studies reporting selexipag efficacy, which met the inclusion criteria for the ITC were GRIPHON and TRACE. Both trials were sponsored by Actelion Pharmaceuticals Ltd, a Janssen Pharmaceutical Company of Johnson & Johnson. Given that the ESC/ERS risk status categorisations have been published more recently than the publication of these studies, data reporting risk status for trial participants is unlikely to have been estimated at the time of study read-out. Risk status categorisations would need to be estimated retrospectively at relevant timepoints (baseline and week 24), using individual patient-level data for the constituent endpoints: 6MWD, WHO FC and NT-proBNP or BNP. The company have not contacted the study authors or sponsors to request the patient-level data.

B2. Document B, Section 3.2.2, P.150-157. To fully understand the impact of the lack of concordance between available data and the ideal model structure, please consider providing:

- An alternative economic model structured around WHO functional class health states parameterised with appropriate transition probabilities, relative risks, costs and utilities derived from the STELLAR trial data and ITC findings.

MSD response:

In developing the CS, the company sought to present the most clinically relevant model structure and discussion with UK clinical experts indicates this to be one which uses ESC/ERS risk status-based health states. The subjective nature of WHO FC

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

categorisation and uncertain link with PAH-related mortality were amongst the reasons cited. As stated in the CS, a model which exclusively categorises patients according to WHO FC risks:

- diverging from the current treatment pathway (i.e., treating patients in the model with the wrong treatment regimen)
- failing to characterise the clinical objective of treatment (i.e., achieving low risk status)
- producing clinically implausible outcomes.

The appropriateness of categorising patients according to ESC/ERS risk status rather than WHO FC was also endorsed by stakeholders (NHS England) in the scope document of this submission (ID6163). As a result, the company are not providing a model constructed using WHO FC-based health states and inputs.

B3. Would it have been possible to build a cost-effectiveness case in a broader population, aligned with the STELLAR trial ITT Population, if health states were defined according to WHO functional class rather than ECS / ERS risk strata? If so, would the company consider providing a case that includes both low and high-intermediate risk populations, with parameterisation based on the ITT analysis set of STELLAR and comparative effectiveness inputs derived from the ITC.

MSD response:

The STELLAR ITT population consisted of WHO FCII and FCIII patients. If a FC-based model were presented, appropriate comparators would be informed by the current NHS clinical commissioning policies, about which the company has received feedback indicating they no longer reflect optimal clinical practice. These sources indicate a differing standard of care (and hence comparator) depending on the patient's WHO FC.

For FCII patients, the policies state the comparator would be monotherapy with a PDE5i (see CS Appendix J). Given that just 4% of the STELLAR population were

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

receiving a background monotherapy at baseline, the comparative efficacy data for this comparison would be limited, and using the ITT results as a proxy for this small monotherapy population was deemed clinically inappropriate. For FCIII patients, the policies state the potential comparators would include selexipag, parenteral PCA and dual therapy including riociguat. As discussed in the CS, ITC versus selexipag is feasible with limitations, ITC versus parenteral PCA is not feasible and riociguat is not used in a population which overlaps with the expected sotatercept patient group. Selexipag is the relevant comparator in the intermediate-low risk population; using the ITC results would not be appropriate for the low risk population (who are achieving their clinical target with dual therapy) or for the intermediate-high risk population (for whom guidelines indicate parenteral PCA added to PDE5i+ERA to be the standard of care).

In conclusion, the company do not believe a FC-based model parameterised using the ITT analysis set of STELLAR and comparative effectiveness inputs derived from the ITC would expand the cost-effectiveness case to the low and intermediate-high risk PAH populations.

B4. PRIORITY. Document B, Section 3.2.2, P.154, Figures 14 & 15. The EAG

Note that there is a disparity between the discontinuation rules for sotatercept and selexipag upon exiting the intermediate-low risk health state beyond 24 weeks. The sotatercept cohort remain on treatment in the intermediate high-risk state, but selexipag cohort are discontinued from treatment, instead receiving IV PCA. Please provide the following additional information:

- Further justification, including detailed responses from multiple UK clinical experts regarding why selexipag treated patients would progress more quickly to IV PCA in the intermediate-high risk state than sotatercept treated patients.

MSD response:

Current standard of care for intermediate-low risk PAH patients is selexipag added to dual therapy with PDE5i + ERA. In order to understand the treatment pathway for these patients once they progress to intermediate-high risk, the company consulted current treatment guidelines, World Symposium on Pulmonary Hypertension (WSPH)

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

consensus guidelines on the future pathway and clinical experts treating patients in the UK. The ESC/ERS guidelines presented in CS Figure 2 indicate selexipag should be used as an option in intermediate-low risk but is not indicated in intermediate-high risk. This is corroborated by the WSPH consensus guidelines presented in CS Figure 1.

In order to understand the real-world application of these guidelines, four UK clinical experts were asked about the use of selexipag in their patients as part of individual consultation meetings. The following response was provided in one of these meetings:

“When a patient progresses from intermediate-low to intermediate-high, the next treatment depends on their co-morbidity profile and PAH type i.e. is the 6MWD truly deteriorating for PAH-related reasons? e.g. those with arthritis or cardiac issues can’t walk far anyway. I would typically offer IV PCA, and inhaled iloprost is another potential option, however the short half-life means very frequent administration (up to 7 times per day) and this is not appealing. Potential forthcoming availability of inhaled treprostinil, which would require less frequent administration, could provide a better option.

If patients are not suitable for parenteral PCA, selexipag would be stopped and options include sGC or entering a clinical trial; there is no value in continuing selexipag in these patients.”

Another clinical expert stated:

“The principal reason for stopping an oral PCA like selexipag is escalation to intravenous PCA, and selexipag is withdrawn over 1-2 days. This happens when a patient deteriorates to intermediate-high risk, and we have the conversation with the patient immediately”.

The positions where clinicians expect to offer sotatercept is also detailed in the WSPH consensus guidelines. In contrast to selexipag, this indicates the addition of sotatercept to PDE5i + ERA in both intermediate-low and intermediate-high risk PAH, and as a second-line treatment in high risk patients, following triple therapy with IV PCA.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

UK clinical experts were also asked about their expected use of sotatercept were it to be available for use in the NHS:

“We know that if mild PAH is treated well, we get better survival outcomes, so I would prescribe in accordance with the World Symposium guidelines.”

“My intended use of sotatercept aligns with the WSP statement”

This divergent use of sotatercept vs. selexipag in the intermediate-high risk population relates to differing perceptions of the treatments’ efficacy and capacity to improve a patient’s risk status. Given the limited capacity for selexipag to improve risk status (stated by multiple UK experts), when a patient has deteriorated to intermediate-high risk, they require a treatment with potential to return that patient to intermediate-low risk, a benefit of sotatercept demonstrated by the results of the STELLAR trial.

- A table detailing the proportion, as n/N (%), of patients receiving IV PCA in all health states, categorised by treatment arm (i.e. sotatercept, selexipag), and pooled across treatment arms, using data from the STELLAR study.

MSD response:

As clarified in the call with the EAG on April 9th, the patients in the sotatercept + ERA + PDE5i vs. selexipag + ERA + PDE5i subgroup analysis were, by definition, not receiving any IV PCA treatments at baseline. The question refers to the proportion of patients in whom IV PCA were initiated during the STELLAR study. This was only observed in the selexipag subgroup, in 2 participants, as presented in Table 2 below.

Table 2: Participants who were initiated on IV PCA during the STELLAR study

	Sotatercept: ERA+PDE5i, █	Placebo: ERA+PDE5i + selexipag, █	Pooled, █
Participants who were initiated on IV PCA during STELLAR: n (%)	0 (0.0%)	█	█

- Please provide functionality within the economic model, along with the results of a scenario analysis, that equalises the treatment discontinuation rules across treatment arms in the model.

MSD response:

A scenario analysis where sotatercept treatment is discontinued upon progression to intermediate-high risk and patients are escalated to receive parenteral PCA in addition to ERA + PDE5i (consistent with stopping criteria for selexipag) is presented within the revised results at the end of the document, in Section C – Scenario Analysis.

B5. Priority. Section 3.3.2, Long-term transition probabilities, and initiation of IV PCA (Table 84 and 85). Please justify why patients in the selexipag arm have no chance of improving their risk status upon initiation of IV PCA as they do in the sotatercept arm? This seems inconsistent with clinical expert advice received by the EAG. Please add functionality for a scenario that allows for improvement in risk status following initiation of IV PCA in the selexipag arm.

MSD response:

MSD acknowledge the clinical advice received by the EAG. The capacity for a patient previously receiving selexipag to improve risk status following initiation of IV PCA was not included due to:

- the complexity this would add to the model structure: this would require the creation of an alternative intermediate-low risk health state characterised by patients receiving PDE5i + ERA + IV PCA (clinical expert opinion indicates that a PCA would not be withdrawn if an intermediate-high risk patient improves). This would include an alternative transition probability matrix to those receiving selexipag in intermediate-low risk.
- the uncertain evidence which would inform transitions for the intermediate-low risk patients receiving (it is uncertain how the efficacy of PDE5i + ERA + IV PCA in intermediate-low risk differs from that of sotatercept or selexipag in combination with dual therapy).

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

However, the company note that this is possible for one cycle when sotatercept patients enter the high risk health state and initiate IV PCA, and the inconsistency that the EAG have raised. In order to equalise the capacity for risk status improvement following progression from the index treatment, functionality to remove this capacity for sotatercept patients has been added to the model, as proposed during the call with the EAG on April 9th. The company agrees to update their base case with this inconsistency removed. Along with other updates and corrections, the updated company base case results are presented at the end of the document, in Section C.

B6. Section 3.2.3, Intervention technology and comparators. Expert opinion to the EAG suggests that in clinical practice there will be patients with intermediate-low risks status who have previously progressed and achieved intermediate-low risk status following initiation of treatment with IV PCA.

- a) Is it the company's intention that these intermediate-low risk groups (who may not be eligible for selexipag) should be excluded from the economic case?

MSD response:

Clinical expert advice received by the company is that patients who achieve intermediate-low risk status following initiation of treatment with IV PCA will remain on the IV treatment in this risk status, hence the standard of care for these patients is PDE5i + ERA + IV PCA. The addition of sotatercept to this combination is not indicated for intermediate-low risk patients by WSPH guidelines; this maximal therapy is indicated for high risk patients only. If a clinician would consider removing the IV PCA from these patients, and replacing it with sotatercept, an ITC would be required and the infeasibility of this has been reported in CS Appendix Q. As a result, there is insufficient evidence to make an economic case for these patients. The expected position for sotatercept is in addition to dual background therapy (PDE5i + ERA).

- b) What proportion of patients in each arm of the STELLAR trial had prior exposure to IV PCA at baseline, in the ITT, intermediate-low risk subgroup and the double therapy (sotatercept: ERA+PDE5i) or triple therapy (placebo: ERA+PDE5i+selexipag) subgroup

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

MSD response:

It was clarified in the call with the EAG on April 9th that the request is for proportions of patients receiving (1) any PCA administered intravenously (2) parenteral PCA available in UK practice (i.e. IV epoprostenol, IV treprostinil, inhaled iloprost). The proportions for the ITT and intermediate-low risk subgroup are presented in Table 3 below. The proportion receiving IV PCA at baseline (according to either definition) is comparable between the arms.

Please note that, by definition, no patients in the double therapy (sotatercept: ERA+PDE5i) or triple therapy (placebo: ERA+PDE5i+selexipag) subgroup were receiving PCA at baseline.

Table 3: Proportions of patients in each arm of the STELLAR trial who received PCA at baseline

ITT population (N, %)	Sotatercept (163, 100.0)	Placebo (160, 100.0)
Any IV PCA	██████	██████
UK parenteral PCA*	██████	██████
Intermediate-low risk† at baseline (N, %)	Sotatercept (73, 100.0)	Placebo (69, 100.0)
Any IV PCA	██████	██████
UK parenteral PCA*	██████	██████

* IV epoprostenol, IV treprostinil, or inhaled iloprost only

† Intermediate-low risk according to the 2022 ESC/ERS treatment guidelines for pulmonary arterial hypertension

Clinical input parameters

B7. Priority. Section 3.3.2, “Risk stratum transition probabilities”. Why are the 24-week relative risks (risk ratios) of improvement/deterioration, derived from the subgroup analysis of STELLAR, being treated as hazard ratios in the model? This appears incorrect and is leading to transition probabilities in the selexipag arm that are internally inconsistent with the data analysis used to inform them. Please correct this and provide updated results.

MSD response:

In the revised cost-effectiveness model submitted alongside these responses, the application of the relative risks (RR) has been updated, multiplying the sotatercept

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

transition probabilities by the inverse of the RR, rather than raising them to the power of the RR. The updated company base case has been reported in the results appendix.

B8. Section 3.3.2. Short-term transition probabilities (up to week 24). The EAG understands that a sub-group of █ intermediate-low risk patients has been used to inform the week 24 transition probabilities for sotatercept. Please provide the baseline characteristics of this sub-group, including background therapies at baseline, and comment on its suitability for informing the transition probabilities for sotatercept when added to PDE5i + ERAs.

MSD response:

The transition probabilities (TPs) for the first 24 weeks of the model were informed by the population who were recorded as having intermediate-low risk status at baseline, and for whom a risk status at baseline and at week 24 was recorded (i.e. in whom a transition probability was estimable). There were █ intermediate-low risk patients randomised to receive sotatercept and a transition probability was estimable for █ of these. The TPs from this population were used in the model in order to align with the population in whom sotatercept is expected to be initiated in UK clinical practice i.e. intermediate-low risk. Clinical advice received is that a patient's risk status is the key factor in deciding which PAH treatment to prescribe. This population does not fully align with those who received sotatercept in combination with PDE5i+ERA (█), as participants' background therapy was pre-determined prior to recruitment and trial investigators were not able to influence the stable background therapy regimens patients were receiving. In the cost-effectiveness model, the population under assessment is those who are intermediate-low risk so the transitions observed in STELLAR for these patients is considered the most appropriate, to understand how those patients' risk status is distributed after 24 weeks of sotatercept treatment. As shown in the table below, the largest proportion of the intermediate-low risk population at baseline were receiving background triple therapy. However, background therapy (dual vs. triple) has been demonstrated to not be an effect modifier for sotatercept efficacy (see CS Table 59). Furthermore, this approach enables the use of a larger sample size (█).

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

The baseline patient characteristics of the intermediate-low risk at baseline population are presented in Table 4 below. It was clarified in the call with the EAG on April 9th that the baseline background therapies request refers to the proportions receiving each regimen type (mono-/dual/triple), and this information is included in Table 4 below.

Table 4: Baseline characteristics of participants with intermediate-low risk status at baseline – STELLAR study full analysis set

	Sotatercept	
	n	(%)
Participants in population	■	
Age (years)		
Participants with data	■	
Mean	■	
SD	■	
Median	■	
Range	■	
Age Group, n (%)		
18 to <45	■	■
45 to <65	■	■
65 to <75	■	■
>=75	■	■
Sex, n (%)		
Male	■	■
Female	■	■
Race, n (%)		
WHITE	■	■
BLACK OR AFRICAN AMERICAN	■	■
ASIAN	■	■
MULTIPLE	■	■
OTHER	■	■
Missing	■	■
Ethnicity, n (%)		
HISPANIC OR LATINO	■	■
NOT HISPANIC OR LATINO	■	■
NOT REPORTED	■	■

	Sotatercept	
	n	(%)
Height (cm)		
Participants with data	■	
Mean	■	
SD	■	
Median	■	
Range	■	
Weight (kg)		
Participants with data	■	
Mean	■	
SD	■	
Median	■	
Range	■	
BMI (kg/m2)		
Participants with data	■	
Mean	■	
SD	■	
Median	■	
Range	■	
Length of time since PAH diagnosis (years)		
Participants with data	■	
Mean	■	
SD	■	
Median	■	
Range	■	
WHO Diagnostic pulmonary hypertension Group I: PAH, n (%)		
IDIOPATHIC PAH	■	■
HERITABLE PAH	■	■
DRUG/TOXIN-INDUCED PAH	■	■
PAH ASSOCIATED WITH CONNECTIVE TISSUE DISEASE	■	■
PAH ASSOCIATED WITH SIMPLE, CONGENITAL SYSTEMIC TO PULMONARY SHUNTS AT LEAST 1 YEAR FOLLOWING REPAIR	■	■
WHO functional classification for symptomatic pulmonary hypertension, n (%)*		
Class II	■	■
Class III	■	■
Type of background PAH therapy, n (%)**		
Mono	■	■

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

	Sotatercept	
	n	(%)
Double	■	■
Triple	■	■
Exercise program for cardiopulmonary rehabilitation, n (%)		
Yes	■	■
No	■	■

Database Cutoff Date: 06DEC2022

Number of participants: full analysis set population, participants with intermediate-low risk status at baseline.

Risk stratum assessment is based on COMPERA version 2.0

*One participant, in the sotatercept arm, is counted in the WHO FC II category in this table (in alignment with the CSR Shift from Baseline in WHO Functional Class by Visit entries). To note is that the same participant, in the CSR Disease History Associated with PAH entry is counted in the WHO FC III category. This discrepancy was caused by an incorrect entry in the WHO FC intake form in the IRT and was recorded to match the IRT entry by the site in the study database. For purposes of efficacy and safety analyses, this participant is counted as WHO FC II.

**One participant, in the placebo arm, is counted in the double therapy category in this table (in alignment with the CSR Disease History Associated with PAH entry). To note is that the same participant, in the CSR Concomitant Standard of Care Medications is counted in the triple therapy category. The discrepancy was caused by an incomplete end date record in the study database for a concomitant PAH therapy. For purposes of efficacy and safety analyses, this participant is counted as being on double therapy.

B7. Document B, Section 3.3.2, P159 onwards. The EAG note that the model is populated using parameters obtained from a mix of different subgroups from the STELLAR study, specifically:

- The full ITT analysis set (FAS) with available data, total N=■ (Sotatercept) for transition probabilities from 12-24 weeks extrapolated over the longer term in the model.
- The intermediate-low risk subgroup, total N=■ for health state distributions at 24 weeks.
- The double therapy (sotatercept: ERA+PDE5i) or triple therapy (placebo: ERA+PDE5i+selexipag) subgroup, total N=■ for comparative effectiveness inputs.

Please provide the following:

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

- Comment on any potential biases due to the use of different data sets to populate different sotatercept parameters in the model.

MSD response:

The company acknowledges the use of efficacy inputs from multiple STELLAR trial populations, in order to make best use of available data. Rationale for deriving the model inputs as outlined above is presented in the following sections.

Health state distributions at week 24

As outlined in the decision problem section of the CS, sotatercept in combination with PDE5i + ERA is expected to be initiated in intermediate-low risk patients, based on the WSPH guidelines and available comparative data. The most appropriate approach in the economic analysis is therefore to model a baseline population consisting of 100% intermediate-low risk patients, and to present how those baseline patients would be distributed following 24 weeks of treatment (informed by the transitions observed in STELLAR). If an alternative distribution of patients were used at baseline (e.g. that of the ITT population, which included patients in multiple risk statuses), the week 24 distribution would be uninformative about the impact of treatment with sotatercept on an intermediate-low risk population. For example, the week 24 risk status of the STELLAR patients who were low risk at baseline would be different to those in the target population, as by definition those patients could not have experienced a risk status improvement. The clinical objective in PAH management is to administer treatments which enable a patient to achieve a low risk status; sotatercept is unlikely to be initiated in patients who have already achieved this status receiving background dual therapy alone. Due to the international nature of the study recruitment, and the range of background therapies observed in the trial population, the risk status distribution of the within-trial analysis subgroups also included multiple risk status categories (i.e. selexipag is not just being administered to intermediate-low risk patients), so this issue also applies if using this distribution at baseline.

For the model cycles which reflected the observed data from STELLAR, the submitted approach makes best use of the available data.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Transitions beyond week 24

The STELLAR trial recorded transition probabilities for patients during the DBPC phase of the trial, aligned with the time period over which the primary and secondary endpoints were assessed. This also preserved sample size to the greatest extent possible, as participants were required to complete this phase of the trial before being permitted to enrol in SOTERIA. Given the rare nature of PAH, sample size available to inform individual transitions is limited and hence use was made of all the available data to inform the model cycles following the observed period. The full ITT analysis set was used to derive transition probabilities for the sotatercept arm from Week 12 to Week 24 in the STELLAR trial, and these were used to represent the ongoing transitions beyond the observed period. The analyses investigating a potential waning in treatment effect from the first interim analysis of SOTERIA (presented in CS Table 80) did not detect a wane in effect.

If the transitions for just the intermediate-low risk at baseline or within-trial analysis subgroups were used for the cycles after week 24, certain transitions would be informed by low and hence uncertain sample size as the patient distribution has shifted from the baseline mix. Given these transitions inform the majority of the model time horizon, week 12-24 data from the full analysis set has been used to mitigate the influence of any potential chance findings informed by uncertain data.

As presented in Table 59 of the CS, interaction tests were carried out for subgroup analyses and no significant difference was observed for outcomes (improvement, maintenance, or worsening) in change from baseline in risk status at Week 24 for any of the tested subgroups, including baseline risk status. Hence, any potential bias introduced by a differing capacity for sotatercept to change a patient's risk status was not detected.

Comparative effectiveness inputs

In order to compare the effectiveness of the intervention and comparator regimens that are relevant to the decision problem, an analysis was performed to determine the effect on risk status change between the subgroups receiving sotatercept + ERA +

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

PDE5i vs. selexipag + ERA + PDE5I + placebo. This was deemed to be the most appropriate method for understanding the relative effectiveness of these regimens, given the absence of randomised data. Using the full ITT dataset, or using the full background triple therapy or baseline intermediate-low risk subpopulations as a proxy was deemed to be inappropriate, based on the mix of background therapies reported for the trial. Selexipag is considered to have a different efficacy profile to other background therapies. Furthermore, given that patients receiving selexipag as a background therapy were not a pre-specified subgroup, the interaction tests mentioned above were not performed on this subgroup.

The company acknowledges that the use of inputs from multiple STELLAR populations inherently includes some bias: certain model parameters (e.g. mean starting age, mean patient weight, baseline gender split etc.) have not been matched to the population used for each efficacy input, however the company is not aware of any clinical rationale that such parameters would be systematically different in the intermediate-low risk or dual therapy populations (with or without selexipag). Given the rarity of PAH and the evolving nature of the management of PAH (particularly in terms of treatment decisions being made based on a patient's risk status rather than WHO FC), there is inevitable uncertainty in the economic case. However, efforts have been made to optimise the use of the available data and to conduct extensive scenario analysis to inform the decision problem.

- A full set of clinical input parameters derived solely from each of the three subgroups, where possible and provide functionality within the economic model to select any of these data sources to populate the model.

MSD response:

Given the rationale provided for the use of inputs from multiple STELLAR populations above, the company believe that the current approach is justified in the context of a rare disease and makes the best use of available data from the trial to inform the respective clinical parameters in the economic model. Using a full set of input parameters for the different populations would lead to implausible and uncertain outcomes.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

B6. Document B, Section 3.3.2, Long-term transition probabilities (post week 24), P158 onwards. Please clarify why long-term transition probabilities were not informed by the SOTERIA long-term extension study for sotatercept? Please consider providing an analysis that utilises all available data from SOTERIA to inform the long-term transitions in the model for sotatercept. For any parameters used from SOTERIA in an updated economic model, please ensure these data are obtained from the subset of SOTERIA participants that entered the study from STELLAR.

MSD response:

As previously presented in response to Clarification Question A3, 154 participants randomised to the sotatercept arm in STELLAR were both eligible and chose to enrol in the SOTERIA single-arm extension study, of which one participant discontinued before receiving a treatment dose in SOTERIA. Table 55 and Table 56 in the CS demonstrate that the sample size in SOTERIA for which a risk status assessment is available at each given timepoint decreases over time. Relative to STELLAR Week 24, only ■ participants had a completed risk status assessment at Week 108, with an unknown distribution. This decline in sample size at each timepoint, indicating that many of the enrolled participants had yet to reach the later timepoints, limits the robustness of estimating long-term transition probabilities from the SOTERIA trial. Thus, the available data from the first interim analysis of SOTERIA is unlikely to provide the required granular longitudinal data, and likely to produce transition probabilities that are highly uncertain.

Furthermore, additional consideration should be given to the potential post-randomisation bias introduced in the SOTERIA trial due to patients and investigators being unblinded to the treatment of each patient which could impact the observed transition patterns.

The company base case relies on the STELLAR trial for transition probabilities of the sotatercept arm beyond week 24 and assumes a constant treatment effect over time, supported by a treatment waning analysis (see Section 3.3.2 of the CS for further details, and clarification question B2, below). This report demonstrated that patients'

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

responses (measured by risk status/ functional class) were maintained over time, without needing to model discrete transition probabilities.

B2. Section 3.3.2, Table 80 (also Tables 55 and 56), SOTERIA - Analysis of the proportion with risk status improvement or maintenance relative to STELLAR baseline and Week 24 outcome. Please clarify the number of participants at each time point that: a) have reached the follow-up time point; b) are still on sotatercept treatment; c) have available risk status outcome data; and d) having missing risk status outcome data.

MSD response:

In these tables, the proportion of patients recording a risk status improvement or maintenance relative to a specified baseline is reported for multiple timepoints. In each of the rows, the denominator presented represents the number of patients who have reached that timepoint and are receiving sotatercept treatment, and for whom it was possible to determine a risk status (i.e. non-missing value). Hence the denominator value (e.g. 130 patients at week 36 in Table 56 and 80) is the answer to requests (a) to (c) above. The number of participants with missing risk status at each timepoint is the difference between the number of sotatercept patients who enrolled in SOTERIA (i.e. 153) and the denominator at that timepoint. However, it should be noted that, given these results are informed by the first interim analysis of SOTERIA, the main reason for a missing value is likely to be that the patient has not yet reached that duration of treatment in the study yet, particularly for the later timepoints. Hence these missing values should not be interpreted as true missing values i.e. not estimable due to an incomplete assessment, rather a reflection of the immaturity of the data.

B8. Document B, Section 3.3.2, Long-term transition probabilities (post week 24), P168, Table 80 Please re-produce Table 80 to provide these data for the double therapy (sotatercept: ERA+PDE5i) or triple therapy (placebo: ERA+PDE5i+ prostacyclin [selexipag]) subgroup.

MSD response:

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

It is only possible to re-produce this table for the double therapy (sotatercept: ERA+PDE5i) subgroup, as participants in the triple therapy (placebo: ERA+PDE5i+ prostacyclin [selexipag]) subgroup would have initiated sotatercept in SOTERIA, not continued it. The data is presented in Table 5 below, noting the low patient numbers.

Table 5: Analysis of the proportion of risk status improvement or maintenance relative to STELLAR Week 24 (subgroup)

	Sotatercept (N ^a =34)		
Endpoint	n/N ^b	%	(95% CI) ^c
Risk status improvement or maintenance relative to STELLAR Week 24			
Week 36	■	■	■
Week 52 (Year 1)	■	■	■
Week 108 (Year 2)	■	■	■

Database Cutoff Date: STELLAR (MK-7962-003) - 06DEC2022 / SOTERIA (MK-7962-004) - 08NOV2023

CI: confidence interval; n: number of participants with improvement or maintenance of the endpoint assessment at a given timepoint relative to Week 24

^a Participants who enrolled to SOTERIA from STELLAR sotatercept arm and continued receiving sotatercept. 154 participants from STELLAR sotatercept arm enrolled to SOTERIA, but one participant discontinued before receiving study dose in SOTERIA.

^b Participants with non-missing endpoint assessment at both given timepoint and STELLAR Week 24. Row percentages are based on this number.

^c Derived based on the Clopper-Pearson method (exact binomial distribution).

B9. PRIORITY. Document B, Section 3.3, Treatment Discontinuation. P.158

onwards. Please provide further details about the proportions, reported as n/N (%), and reasons for sotatercept and selexipag treatment discontinuation. Specifically, please provide the following:

- Proportions of patients discontinuing sotatercept treatment, and the reasons for treatment discontinuation, during the STELLAR study, up to 24 weeks, from the STELLAR study up to 72 weeks, and from the subgroup entering the SOTERIA study from STELLAR up to 108 weeks as well as full details of the reasons for treatment discontinuation.

MSD response:

As clarified in the call with the EAG on April 9th, selexipag was administered in the STELLAR study as a background therapy, rather than an investigational treatment. As a result, rate of, and reasons for, discontinuation of the placebo:

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

ERA+PDE5i+prostacyclin [selexipag]) subgroup were not recorded. The following responses are provided for the sotatercept arm only, as agreed during the call.

The proportion of patients discontinuing sotatercept (and reasons for discontinuation) at week 24 in the STELLAR study are presented in Table 6 below.

Table 6: Proportion of patients discontinuing sotatercept (and reasons for discontinuation) at week 24 in the STELLAR study

	Sotatercept arm: N (%)
Total randomised	████████
Discontinued study treatment (total)	██████
Primary reason for discontinuation of study treatment:	
Adverse event	██████
Participant's request (withdrawal of consent)	██████
Participant's unwillingness or inability to comply with the protocol	██████
Clinical worsening event	██████
Death	██████

The proportion of patients discontinuing sotatercept at week 72 in the STELLAR study (and reasons for discontinuation) are presented in Table 7 below.

Table 7: Proportion of patients discontinuing sotatercept at week 72 in the STELLAR study

	Sotatercept arm: N (%)
Total randomised	████████
Discontinued study treatment (total)	██████
Primary reason for discontinuation of study treatment:	
Adverse event	██████
Participant's request (withdrawal of consent)	██████

Participant's unwillingness or inability to comply with the protocol	
Clinical worsening event	
Death	
Other	

The proportion of patients discontinuing sotatercept in the subgroup entering the SOTERIA study from STELLAR up to 108 weeks (and reasons for discontinuation) are presented in Table 8 below.

Table 8: Proportion of patients discontinuing sotatercept in the subgroup entering the SOTERIA study from STELLAR up to 108 weeks

	Continued sotatercept arm ^a : N (%)
Total no. of subjects enrolled	
Discontinued study treatment (total)	
Primary reason for discontinuation of study treatment:	
Adverse event	
Participant's request (withdrawal of consent)	
Participant's unwillingness or inability to comply with the protocol	
Death	
Other	

^a: patients who were randomised to sotatercept in STELLAR and continued treatment upon enrolling in SOTERIA

- Using the double therapy (sotatercept: ERA+PDE5i) or triple therapy (placebo: ERA+PDE5i+prostacyclin [selexipag]) subgroup from the STELLAR study, please provide the proportion of sotatercept and selexipag patients that discontinued treatment at week 24.

The proportion of patients discontinuing sotatercept in the double therapy (sotatercept: ERA+PDE5i) subgroup (and reasons for discontinuation) at week 24 in the STELLAR study are presented in Table 9 below.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Table 9: Proportion of patients discontinuing sotatercept in the double therapy (sotatercept: ERA+PDE5i) subgroup

	Sotatercept arm: N (%)
Total randomised	██████████
Discontinued study treatment (total)	██████
Primary reason for discontinuation of study treatment:	
Adverse event	██████
Death	██████

- A breakdown of the proportion of sotatercept and selexipag patients that remained on treatment in each modelled health state at 24-weeks follow up from the STELLAR study.

The sotatercept arm patients who remained on treatment at week 24 and for whom a risk status was estimable are presented in Table 10 below, including a breakdown per risk status. This does not reflect their baseline risk status.

Table 10: Risk status distribution of sotatercept arm patients who remained on treatment at week 24 in the STELLAR study

	Sotatercept arm: N (%)
Total with non-missing risk status	██████████
Low risk	██████
Intermediate-low risk	██████
Intermediate-high risk	██████
High risk	██████

B10. PRIORITY. Document B, Section 3.3, Clinical expert opinion, P.158

onwards. Clinical expert opinion is used widely throughout the clinical inputs, parameterisation of the model and notes of the expert validation work are provided in Appendix R of the submission. Please provide further details of all clinical expert elicitation work, including detailed feedback obtained from the clinical experts, including number of experts consulted, their background, methods used to elicit their opinions (e.g. validation of model assumptions or active participation in development Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163])

and refining of assumptions), and full details of all numerical estimates and advice provided by each individual clinical expert consulted. Where variation of opinion occurred, please provide details of how consensus on base case assumptions was agreed.

MSD response:

In preparation of the CS, the company sought UK clinical expert opinion to ensure that the inputs and assumptions used in the analysis were relevant to UK clinical practice. This was considered particularly important, given the apparent inconsistency between current clinical practice and reimbursement policies, and the absence of a prior NICE TA in PAH. Feedback was obtained in two advisory boards and in total, input was gathered from seven UK clinical experts (consultant physicians and senior nursing staff), including a series of individual consultancy meetings. All clinicians consulted treat PAH patients at one of the six specialist centres in England (an appropriate geographical spread was ensured), and in addition, two nurse specialists were consulted regarding the HCRU associated with selexipag and parenteral PCA, in addition to the anticipated HCRU for administration and monitoring of sotatercept. Many expert-sourced inputs were qualitative in nature, relating to the current clinical pathway, appropriate assumptions to make in the absence of data, or the appropriateness of external data as model inputs. A summary of the key insights is presented in the CS Appendix R.

Where a numerical estimate informed by clinical expert opinion was used in the CS, this is indicated, and in the case of competing values, appropriate scenario analysis has been conducted. As is common with expert opinion, alternative views were expressed both in the advisory boards and across the individual meetings; in these cases, arbitration of the consensus view was conducted where possible and the prevailing output was used as a CS input, with other views addressed through sensitivity analysis.

B12. Document B, Section 3.3.2, Mortality, P173. Please provide further details describing how the Rosenkrantz 2023 and Boucly et al. 2022 studies were selected. The EAG note the submission describes a ‘targeted literature search’ and ‘clinical Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

expert opinion' to select studies. Please provide details of the targeted search. How many studies were identified? Did clinical experts review all studies? What criteria were used to prioritise these two studies?

MSD response:

The company cost-effectiveness model structure consists of health states stratified by risk status according to the COMPERA four-strata risk classification, as per UK clinical practice. The 2022 ESC/ERS treatment guidelines also recommend using four-strata risk stratification. Given that these guidelines were only published in 2022, there is only a short period between adoption of these guidelines and the timing of this appraisal, for the subsequent reporting of risk status-related outcomes in PAH registries and real-world databases. As such, few studies would be expected to have reported survival according to risk-status stratification to date. Therefore, an unprotocolled literature search was taken to identify relevant studies using keyword search terms. Only two studies were identified reporting survival in PAH patients stratified by ESC/ERS risk status; Boucly (2022) and Rosenkranz (2023) (3, 4).

Subsequently, both these studies were presented to two clinical experts for their review of the generalisability of the study populations to the population included in the STELLAR trial. In addition, clinical experts gave feedback that they were not aware of any further publications reporting long-term survival stratified by four-strata risk status in PAH, corroborating the findings of the company literature search. Both Boucly (2022) and Rosenkranz (2023) were subsequently presented to UK clinical experts, who agreed that survival data from the “no or fewer comorbidities” subgroup in both publications would be more applicable to patients recruited to the STELLAR trial, and to this decision problem population, given that co-morbid patients are less likely to be receiving dual background therapy due to tolerability (see CS Appendix R for further details).

Furthermore, a systematic literature review published by Reinders, Didden, and Ong (2024) identified studies reporting disease outcomes in PAH patients (5), confirmed through right-heart catheterisation, in real-world population-based observational studies. The search, conducted in November 2021, identified five studies in which both

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

survival and risk status was reported, of which only two studies - based on the French PAH Registry (Boucly 2022) and COMPERA (Hoeper 2021) - reported the four-strata stratification (3, 6). The Rosenkranz (2023) publication, used as the model base case for informing mortality, was not identified in this review as it was published after the search date, however it uses data from the same identified COMPERA registry and reports more granular long-term survival on sub-populations based on patient comorbidity. The company are unsure about how the Boucly study from 2022 was identified in the 2021 search. However, a pragmatic search of the literature since November 2021 using the search terms from the Reinders publication did not identify additional relevant studies.

At this time, the company is not aware of any alternative studies or publications in which survival in PAH patients is reported stratified by risk status using the COMPERA 2.0 four-strata risk classification to inform long-term survival in the economic model.

Adverse reactions

B13. Document B, Section 3.4.4. Adverse reactions, P.187. The EAG note the assumption that adverse reactions are similar for sotatercept and placebo. However, please provide commentary and a comparison of adverse event probabilities between sotatercept and selexipag treated groups from the STELLAR study. If differences in adverse event profiles exist, consider incorporating these into the economic model in terms of cost and utility implications.

MSD response:

Selexipag was administered in the STELLAR study as a background therapy, rather than an investigational treatment. As a result, adverse event rates and severity were not recorded for this subpopulation.

Sotatercept safety was monitored and reported for the FAS population and not for the subgroup in which sotatercept was added to background PDE5i + ERA. The company is not aware of any biological rationale by which a different safety profile would be

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

expected for this subgroup compared to the FAS population, and so makes the assumption that the adverse event probability for the overall population is generalisable to the subgroup. As stated in the CS, more serious adverse events have been reported for selexipag than for sotatercept, however in the absence of severity data from the GRIPHON study publication, this cannot be quantified, and the company considers this to be a benefit not captured in the QALY.

Utility parameters

B14. PRIORITY. Document B, Section 3.4.5, P.188, Table 90. Please provide tables detailing the following EQ-5D-5L mapped to 3L health state utility values, reported as mean (SD) N:

- Each health state, by measurement time point, for the full analysis set, the low-intermediate subgroup, and the double combination therapy group

MSD response:

The health state utility values at baseline and week 24 for the full analysis set are presented in Table 11 below.

Table 11: Health state utility values at baseline and week 24 of the STELLAR study

N ^a :323	Baseline		Week 24	
	N	Mean (SD)	N ^b	Mean (SD)
Low risk				
Intermediate-low risk				
Intermediate-high risk				
High risk				

a: Number of participants: full analysis set population

b: Number of participants: full analysis set population, participants with non-missing EQ-5D-5L Score at the given timepoint
EQ-5D-3L utility scores (mapped crosswalk) derived based on reference: Alava, H. et al. "Estimating EQ-5D by age and sex for the UK", DSU, January 2022.

Database Cutoff Date: 06DEC2022

The utility values at baseline and week 24 for the intermediate-low risk at baseline subgroup are presented in Table 12 below. This has not been estimated per health state due to the limited sample size available to inform individual risk status groups.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Table 12: Utility values at baseline and week 24 for the intermediate-low risk subgroup of the STELLAR study

	Total sample pooled across arms N=323 ^a
Baseline N ^b Mean (SD)	
Week 24 N ^b Mean (SD)	

a: Number of participants: full analysis set population

b: Number of participants: full analysis set population, participants in the intermediate-low risk stratum with non-missing EQ-5D-5L Score at the given timepoint

EQ-5D-3L utility scores (mapped crosswalk) derived based on reference: Alava, H. et al. "Estimating EQ-5D by age and sex for the UK", DSU, January 2022.

Database Cutoff Date: 06DEC2022

The utility values at baseline and week 24 for the dual background therapy subgroup at baseline (background therapy was only recorded at baseline in STELLAR) are presented in Table 13 below. This has not been estimated per health state due to the limited sample size available to inform individual risk status groups.

Table 13: Utility values at baseline and week 24 for the dual background therapy at baseline subgroup of the STELLAR study

	Total sample pooled across arms N=323 ^a
Baseline N ^b Mean (SD)	
Week 24 N ^b Mean (SD)	

a: Number of participants: full analysis set population

b: Number of participants: full analysis set population, participants on double background PAH therapy

EQ-5D UK crosswalk (from EQ-5D-5L to EQ-5D-3L) algorithm applied based on reference: Van Hout B et al. (2012) "Interim Scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L Value Sets", Value in Health Volume 15, Issue 5, July-August 2012, Pages 708-715.

EQ-5D-3L utility scores (mapped crosswalk) derived based on reference: Alava, H. et al. "Estimating EQ-5D by age and sex for the UK", DSU, January 2022.

Database Cutoff Date: 06DEC2022

- All available HSUV data, by health state and measurement time point observed in the double blinded follow-up beyond 24 weeks and from the SOTERIA study

MSD response:

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

In the STELLAR trial, EQ-5D-5L was included as an exploratory (non-alpha-controlled) endpoint, for the purpose of informing health state utility values in the cost-effectiveness analyses. The EQ-5D-5L is a generic patient-reported outcome (PRO) measure that does not assess PAH-specific health-related quality of life, but overall health status. Thus, it is not expected to be sensitive to smaller, more incremental changes in disease-specific health status.

As per the protocol for the STELLAR study, during the DBPC EQ-5D-5L surveys were only administered at baseline and at the end of the main trial period, week 24 (primary endpoint) as is common. As such, no EQ-5D-5L surveys were collected in the LTDB phase of the STELLAR trial.

The EQ-5D-5L questionnaire is not administered in SOTERIA in line with the study protocol, hence additional measures of health state utility are not available beyond week 24.

B15. Document B, Section 3.4.5, P.190, Table 92. The EAG are concerned that application of a utility decrement of 0.307 to all patients receiving IV infusion may double count the utility impact of higher-risk health states. Please clarify why the addition of an administration disutility for IV treatment has not already been captured within the health state utility values used for the model. Using STELLAR trial data, please provide a table detailing the proportion of patients receiving IV Infusion in each modelled health state (by treatment arm and pooled across arms). Should the company wish to maintain the assumption of an additional utility decrement, please provide results of an appropriately specified utility regression model, providing evidence of the magnitude of utility decrement, within state, using STELLAR trial data.

MSD response:

For patients whose PAH requires them to receive PCA treatment intravenously, an additional quality-of-life burden has been reported (7). The company base case sought to incorporate this additional disutility, given the model assumption that patients initiate Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

parenteral PCA upon entry to certain health states (high risk in the sotatercept arm, intermediate-high or high risk in the selexipag arm). Health state utility values were calculated based on the elicited EQ-5D-5L responses from the STELLAR study, where IV PCA was a background therapy (i.e. patients were established on the treatment prior to study recruitment) so the HRQoL impact of initiation of the pump and particularly the long hospital stay is not captured in the trial values.

Furthermore, as detailed in Table 14 below, the health state values are calculated from responses collected from participants both on- and off-IV PCA, and the majority of participants in the STELLAR study were not receiving an IV PCA. Given that IV PCA administration was limited to approximately one-quarter of the participants in any of the health states, the health state values are unlikely to adequately capture the impact of receiving treatment via this administration route. It is the company's view that to exclude this disutility (i.e. to use the reported health state value alone) is to overestimate the HRQoL of a patient receiving an IV PCA in the model. Similarly, a utility regression model using IV PCA as a covariate would rely on limited patient numbers to inform the "with IV PCA" group and the resulting decrement would be highly uncertain; the company has not conducted this analysis.

As a result, literature values reporting the impact of intravenous treatment on HRQoL were consulted in order to understand the independent effect. The Davies et al. (2018) study has the dual benefits of using health state descriptions informed by the lived experience of patients with PAH receiving the treatment intravenously and then eliciting values directly from members of the UK general population, reflecting their preferences (8).

Table 14: Proportions of participants receiving IV PCA per risk status at baseline

	Sotatercept (n, %)	Placebo (n, %)	Total (n, %)
Low risk at baseline			
On IV PCA			
Off IV PCA			
Intermediate-low risk at baseline			
On IV PCA			
Off IV PCA			

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Intermediate-high risk at baseline						
On IV PCA						
Off IV PCA						
High risk at baseline						
On IV PCA						
Off IV PCA						

To address the uncertainty associated with this model input and to understand the impact for cost-effectiveness, an additional scenario analysis is presented, applying a disutility for IV PCA that is half of the value reported by Davies et al. This is in line with the approach adopted by PBAC in the appraisal of selexipag (9), acknowledging the additional utility decrement to reflect the quality of life burden associated with the administration of IV PCA in PAH patients. Please see Table 28 for scenario analysis results. Under this assumption, the ICER remains below the willingness-to-pay threshold.

Cost parameters

B16. PRIORITY. Document B, Section 3.5.1, P.196, Table 98. Please provide the source for the treatment distribution applied in each treatment arm in Table 98.

Where these treatment distributions were not sourced from the STELLAR study, please provide a table comparing the base case distribution with the distribution of treatments from STELLAR for each group.

MSD response:

The treatments applied in each health state reflect the regimen specified for each risk status by the treatment guidelines. The distribution per drug within each drug class was derived from the UK National Audit database to reflect the distribution used in UK clinical practice. In the absence of further stratification of distribution by risk status, this distribution within each drug class is not assumed to differ by risk status and thus is applied to each health state where that drug class is applicable i.e. ERA + PDE5i in all

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

health states, PCA only in health states following treatment discontinuation of either sotatercept or selexipag as per the specified treatment stopping rules.

In the economic model, neither the dual background therapy treatment distributions nor the distribution of parenteral PCA impact the treatment efficacy. In order to reflect current clinical practice, and consequently drug prescription costs, the company consulted the UK National PAH Audit to obtain current market shares for each of the products within the drug class. These distributions are used in the base case model and are presented in Table 15 below.

Table 15: Base case treatment distributions based on UK clinical practice

Drug class and drugs per drug class	Number of drug prescriptions				Distribution per drug class
	All PH patients	CTEPH patients	PAH patients (All PH – CTEPH)	Total per drug class	
ERA					
Bosentan	366	16	350	2742	12.76%
Ambrisentan	1344	157	1187		43.29%
Macitentan	1473	268	1205		43.95%
PDE5i					
Sildenafil citrate	3165	618	2547	4117	61.87%
Tadalafil	1896	326	1570		38.13%
PCA					
Epoprostenol	284	9	275	444	61.94%
Treprostinil	24	2	22		4.95%
Iloprost (inhaled)	155	8	147		33.11%

The distribution of dual background therapies at baseline in the STELLAR trial for both the sotatercept + ERA + PDE5i and selexipag + ERA + PDE5i + placebo groups are presented in Table 16 below. The distribution of therapies within the PDE5i and ERA drug classes is comparable to the real-world prescription data in the UK.

Both treatment groups were by definition not receiving a PCA at baseline in the STELLAR trial. For the distribution of individual PCA products which these patients would receive following progression, the trial was not used because, upon discontinuation of sotatercept in the trial, patients were no longer followed up. Furthermore, the number of discontinuing sotatercept patients was very low. In order to reflect clinical practice and hence costs accurately, the National PAH Audit was also consulted and the market shares of the currently available PCA products were used.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

The company believe the real-world prescription data is the most appropriate source for informing the treatment mix of PCA in England and Wales clinical practice.

Table 16: Dual background therapy treatment distribution at baseline in STELLAR

Dual background therapy component	Sotatercept + ERA + PDE5i ■■■	Selexipag + ERA + PDE5i + placebo ■■■
Sildenafil	19 (52.8%)	13 (50.0%)
Tadalafil	17 (47.2%)	13 (50.0%)
Macitentan	16 (44.4%)	16 (61.6%)
Ambrisentan	11 (30.6%)	8 (30.8%)
Bosentan	9 (25.0%)	2 (7.7%)

B17. PRIORITY. Document B, Section 3.5.1, P.198, Table 100 & Section 3.3.1, P159, Table 74. Please provide further details regarding the different potential assumptions for weight-based dosing in the economic model. Specifically, please provide:

- a table detailing descriptive statistics (mean, SD, median, min, max, N) of patient weight for each of the following categories: Full ITT analysis set, low-intermediate risk group and the double/triple treatment subgroup.
- a table detailing descriptive statistics (mean, SD, median, min, max, N) of patient weight for each of the following groups from the STELLAR study: all countries, all countries without the USA, and for the UK only.
- Please update the economic model to incorporate the above and provide the results of scenario analyses using these different groups of STELLAR study trial participants. Please also provide an updated version of Table 100 that incorporates all these scenarios.

MSD response:

The descriptive statistics of patient weights is provided in Table 17 below for each of the following categories from the STELLAR trial: full ITT, intermediate-low risk at baseline, double background therapy at baseline, and triple background therapy at baseline.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Table 17: Descriptive statistics of patient weights

Participants	N	Patient weight (kg)		
		Mean (SD)	Median	Min, Max
ITT	323	72.3 (18.32)	68.2	38.0, 141.3
Intermediate-low risk	■	■	■	■
Double PAH background therapy	■	■	■	■
Triple PAH background therapy	■	■	■	■

ITT, intention to treat; kg, kilograms; PAH, pulmonary arterial hypertension; SD, standard deviation

The descriptive statistics of patient weights is provided in Table 18 below for participants from all countries (equivalent to the ITT population presented above) and for all countries excluding USA from STELLAR. Given the small number of UK patients enrolled in STELLAR (■), the company do not consider it feasible, or clinically meaningful, to perform descriptive statistics for these participants, as explained in the call with the EAG on April 9th.

Table 18: Descriptive statistics of patient weights, by country

	Patient weight (kg)			
	N	Mean (SD)	Median	Min, Max
All countries (ITT)	323	72.3 (18.32)	68.2	38.0, 141.3
■	■	■	■	■

ITT, intention to treat; kg, kilograms; SD, standard deviation; USA, United States of America

B18. Document B, Section 3.5.2, P.201, Table 104. Please clarify why clinical expert opinion was relied upon, rather than mapping to an appropriate HRG code within NHS reference costs to estimate the duration of hospitalisation and costs of catheterisation for IV prostacyclin. Similarly, for maintenance costs, please clarify

what consumables are included in the £450 per month cost estimate, and would these be paid for by the NHS?

MSD response:

The company understands that the process related to the catheterisation for PAH patients initiating IV PCA is increasingly complicated compared to the catheterisation for other conditions. This is largely due to the extensive training that is required for both patients and carers during the hospital stay to ensure that, upon discharge, patients and carers are adequately trained to self-administer and monitor the IV PCA treatment outside of the hospital setting. Therefore, based on conversations with experts, the company believe that the total cost to the healthcare system of the catheterisation process would be underestimated if using a HRG code informed by a broad range of catheterisation stays across multiple therapeutic purposes. To understand this complexity and how it is accounted for by NHS costs, the company consulted with 2 nurse specialists from PAH centres in England. The cost of a hospital stay during the initiation of IV PCA for PAH patients was quoted as £483 per night.

Attempts to identify a HRG tariff which may reflect the cost of the catheterisation for IV PCA initiation were made, and a proposed option is to use the weighted average of 'Complex cardiac catheterisation' elective inpatient currencies. To avoid over-costing, it is proposed to exclude patients with a CC score of 7+ as this likely represents patients with complex comorbidities admitted in an urgent context. The resulting weighted average gives a similar value to the expert-informed cost in the company base case (see Table 19 below).

Table 19: HRG tariffs reporting complex cardia catheterisation costs

Currency – Description	Activity	Unit Cost	Weighted Average
EY42B – Complex Cardiac Catheterisation with CC Score 4-6 (Elective Inpatient)	296	£ 4,626.11	£4,233.67
EY42C – Complex Cardiac Catheterisation with CC Score 2-3 (Elective Inpatient)	331	£ 3,868.89	
EY42D – Complex Cardiac Catheterisation with CC Score 0-1 (Elective Inpatient)	146	£ 4,264.52	

Source: NHS National Cost Collection 2023/24

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

The nurse specialists were also consulted about consumables involved in IV PCA initiation, and a monthly cost of £450 was provided. This consists of the cost of equipment required for a patient to manage their Hickman line with all equipment provided by the hospital at discharge following catheterisation, and thus would be paid for by the NHS, making it relevant to include in the economic evaluation as per the NICE reference case. In addition to drug vials and infusion pumps which have been costed separately, the monthly consumables cost includes, but is not limited to:

- Cassette
- Syringes
- Needles
- Tubing
- Dressing packs
- Sterile gloves
- Sterile wipes
- Sharps disposal bin
- Medical tape
- Sterile gauze packs

In the absence of available unit costs that would enable a micro-costing approach, the monthly cost of consumables was sourced from PAH nurse specialist opinion and used in the model base case.

B19. Document B, Section 3.5.4, P.205, Table 107. Please confirm whether PAH hospitalisation events are partially double counted in the model. For example, can the company confirm that the costs of IV prostacyclin catheterisation and re-catheterisation are not double counted within the HES analysis from Exposto 2021? Please consider appropriate scenario analyses to ensure that there is no double counting of costs in the economic model.

MSD response:

The cost per hospitalisation from Exposto 2021 using data from HES captures all hospitalisations for the identified PAH patients using the algorithm specified in the

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

study, and therefore the company are unable to confirm that IV prostacyclin catheterisation costs are not included in the weighted cost per PAH hospitalisation. The cost of IV prostacyclin catheterisation would only be applicable to patients initiated with IV PCA, who comprise a subset of the total PAH population within the Exposto 2021 study, and therefore the risk of double counting would only be applicable to an unknown proportion of the total study population. No baseline characteristics are provided in the study according risk status, functional class, or PAH therapy regimen and therefore the company is unable to determine the proportion of the study population on IV PCA to which catheterisation costs would apply.

To account for the potential for IV PCA initiation hospitalisation costs being double counted, the company have provided an additional scenario analysis where cost per hospitalisation is based on an HRG tariff using the unit cost for a generic, disease-agnostic, 'Daycase': £1,031 (NHS National Cost Collection 2023/24). Catheterisation and re-catheterisation requires an overnight hospital stay and thus using the day case tariff would logically ensure that the hospitalisation cost in the model does not double count for catheterisation.

The resulting ICER is presented within the revised results at the end of the document, in Section C – Scenario Analysis. This is a likely underestimate of the average PAH hospitalisation cost given the complicated needs of a patient with this disease. Nevertheless, this results in a minor impact on the ICER and is not a driver of the model results.

Economic model switches and parameters

B20. Submitted economic model, Tab: “indirect cost inputs”, Cell: C9.

Changing the “societal perspective” switch from “yes” to “no” amends QALY calculations. Please clarify what this switch in the model refers to, and please confirm that the submitted base case adheres to an NHS and PSS perspective as per the NICE reference case.

MSD response:

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

The company acknowledges that the “Societal perspective” switch in the submitted economic model controlled for both indirect costs and carer quality of life. Therefore, switching this setting on the indirect cost inputs sheet to “No” would remove the inclusion of the carer quality of life impact and thus affect the modelled QALYs in either treatment arm. Despite this, no indirect costs were included in the company economic evaluation and the submitted base case adheres to an NHS and PSS perspective as per the NICE reference case.

In the updated company economic model, the societal perspective controls have been amended so that the application of indirect costs and carer quality of life are controlled independently. This has no impact on the model results. Please refer to the change log for further details of the amendments made to the updated company model.

Severity weighting calculation

B21. Document B, Section 3.6, P.206, Table 109. Please also provide severity weighting calculations without application of carer utility impacts.

MSD response:

The company confirm that the calculation of the absolute shortfall presented in CS Table 109 (and hence QALY weight) does not include the application of carer utility impacts. Please refer to the formula in the model (Tab: Main results; Cell: K79) which derives the absolute shortfall. This table has been replicated and updated below to include economic model updates following clarification questions.

Table 20: Summary of QALY shortfall analysis (including clarification question model updates)

Expected total QALYs for the general population	Total QALYs that people living with a condition would be expected to have with current treatment*	Absolute QALY shortfall	Proportional QALY shortfall	QALY weight
16.74	■	■	■	1.2

QALY, quality-adjusted life-year

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Section C: Textual clarification and additional points

C1. PRIORITY. Document B, Section 3.4.5, P.188, Table 90. The reported health state utility value for intermediate-high risk does not align with the value used in the submitted economic model (Tab: Patient HRQoL inputs; Cell: E11). Please clarify the correct utility value and if necessary, please update all results tables accordingly.

MSD response:

The company acknowledge the inconsistency in the intermediate-high risk health-state utility value reported in Table 90 of the company submission and the submitted economic model, with the error being in the economic model. The health state utility value in 'Patient HRQoL inputs!E11' has been amended to align with the CS document, resulting in a minor change to the base case results. Updated full model results tables for the company revised base case are provided below.

Updated Results

In response to the clarification questions, the company has updated their base case. The revised base case incorporates the following:

- correcting the utility value of the intermediate-high risk health state
- updating the method for application of relative risks
- removing the one-cycle improvement for patients discontinuing sotatercept and initiating IV PCA (to be consistent with the selexipag arm).

In addition to the detailed deterministic results, the sensitivity and scenario analyses were also replicated using the new base case. For a detailed change log please refer to the front sheet of the CEM that includes a list of all changes that were applied to the model.

In the updated base case sotatercept in combination with dual therapy accumulated [REDACTED] of costs and [REDACTED] QALYs and selexipag in combination with dual therapy accumulated [REDACTED] of costs and [REDACTED] QALYs. The resulting ICER of [REDACTED] per QALY,

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

inclusive of the 1.2 QALY severity modifier for which sotatercept qualifies, is below the £30,000/QALY willingness-to-pay threshold and is consequently cost-effective.

Table 21: Deterministic base case results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)	ICER with severity modifier (£/QALY)
Sotatercept + dual therapy	████	██	██	████	██	2.53	████	████
Selexipag + dual therapy	████	6.37	██	–	–	–	–	–

ICER: incremental cost-effectiveness ratio; LYG: life-years gained; QALY: quality-adjusted life-years

Error! Reference source not found. Table 22 presents the incremental deterministic net health benefit (NHB) and net monetary benefit (NMB) vs. sotatercept.

Table 22: Deterministic net health benefit and net monetary benefit

Technologies	Incremental costs (£)	Incremental QALYs	NHB at £20,000	NHB at £30,000	NMB at £20,000	NMB at £30,000
Selexipag + dual therapy	████	2.53 * 1.2 = 3.04	██	██	████	████

NHB: net health benefit; NMB: net monetary benefit; LYG: life-years gained; QALY: quality-adjusted life-years

Exploring uncertainty

Probabilistic sensitivity analysis

A probabilistic sensitivity analysis (PSA) was run for 1,000 iterations.

The parameters varied in the PSA and the nature of the uncertainty around their value are described in the Company Submission. The mean summary results from the probabilistic analysis are presented in Table 23 below. Graphical results including the cost-effectiveness plane (incremental costs vs. incremental QALYs) and the cost-effectiveness acceptability curve (CEAC) are shown below in Figure 3 and Figure 4 respectively.

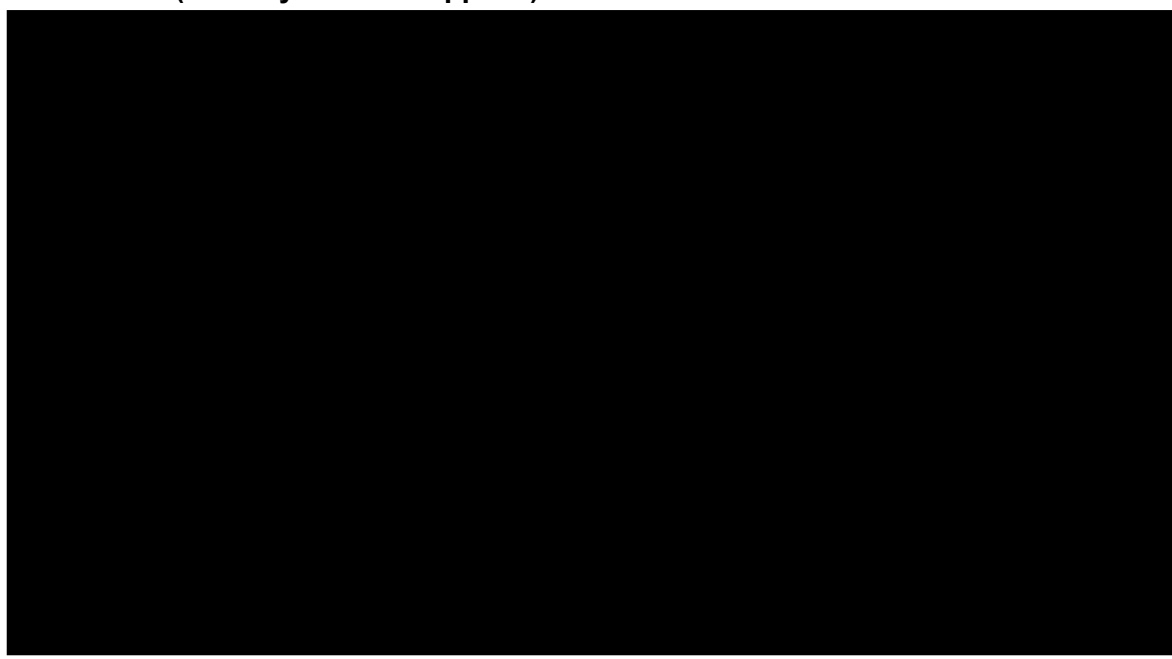
Table 23: Probabilistic base case results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)	ICER with severity modifier (£/QALY)	ICER weighted according to modifier (£/QALY)
Severity modifier									
Sotatercept + dual therapy	■	■	■	■	■	2.72	■	■	■
Selexipag + dual therapy	■	■	■	■	■	-	■	■	■

ICER: incremental cost-effectiveness ratio; LYG: life-years gained; QALY: quality-adjusted life-year

Based on the 1000 iterations, sotatercept compared to selexipag add as an add-on to dual background therapy in intermediate-low risk was cost-effective in ■ of iterations at a willingness-to-pay threshold of £30,000 per QALY, with more than ■ out of all iterations being dominant (i.e., sotatercept is more efficacious and less costly). There were no iterations where sotatercept was projected to be dominated by selexipag or less efficacious but less costly.

Figure 3: Cost-effectiveness plane scatter plot (1,000 iterations) – Incremental Costs and QALYs (severity modifier applied)



QALY: quality-adjusted life-year; WTP: willingness-to-pay

Figure 4: Cost-effectiveness acceptability curve (1,000 iterations) (severity modifier applied)



BGT: background therapy; QALY = quality-adjusted life-year

Deterministic sensitivity analysis

Please note that a correction was made in the DSA calculations due to the following reason: the dimensions of the range “DSA_out” were not large enough to hold every input tested within the DSA, causing some parameters to be excluded. This error affected parameters related to market share of PCA products in intermediate-high and high risk groups, which have the largest impact on the results. The company apologises for any inconvenience this may have caused. However, with the severity modifier applied, all DSA results remain below the £30,000/QALY threshold. The correction implemented did not impact calculations related to patient flow, quality of life, or cost accumulation and was limited to the functionality used to facilitate deterministic sensitivity analysis. Results of the updated analyses are presented in Table 24 and Figure 5.

Table 24: Ten most impactful parameters from one-way deterministic sensitivity analysis on ICER (severity modifier applied)

Parameter, (Base Case [Lower Bound, Upper Bound])	ICER for lower bound value	ICER for upper bound value
Base case	████	█
Epoprostenol, IV infusion - Selexipag + dual BGT - High (0.619 [0.37, 0.84])	████	████
Epoprostenol, IV infusion - Selexipag + dual BGT - Int-High (0.619 [0.37, 0.84])	████	████

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Parameter, (Base Case [Lower Bound, Upper Bound])	ICER for lower bound value	ICER for upper bound value
Drug acquisition cost/12 wks: Iloprost, Inhaled (9244.456 [5982.52, 13204.82])	████	████
RR, RS worsening (sotatercept vs. selexipag from within-STELLAR analysis) ██████████	████	████
Iloprost, Inhaled - Selexipag + dual BGT - High (0.331 [0.21, 0.47])	████	████
Macitentan, Oral - Selexipag + dual BGT - High (0.44 [0.27, 0.61])	████	████
Macitentan, Oral - Sotatercept + dual BGT - Int-High (0.44 [0.27, 0.61])	████	████
Iloprost, Inhaled - Selexipag + dual BGT - Int-High (0.331 [0.21, 0.47])	████	████
Age (years) (47.9 [46.33, 49.47])	████	████
Non-sotatercept PAH drugs: HCRU Cost Selexipag + dual BGT: High Risk ██████████	████	████

BGT = background therapy; HCRU = health-related resource use; ICER: incremental cost-effectiveness ratio; Int: Intermediate; RS = risk status

Figure 5:Tornado diagram for ICER impact for one-way deterministic sensitivity analysis (severity modifier applied)



BGT: background therapy; HCRU: healthcare resource use; ICER: incremental cost-effectiveness ratio; IV: intravenous; PAH: pulmonary arterial hypertension; QALY: quality-adjusted life-year; RR: relative risk; RS: risk status; SC: subcutaneous injection; wks: weeks

Scenario analysis

Various scenarios were explored to understand the impact of structural and input assumptions in the model base case. The list of scenarios and the rationale behind their inclusion in the analysis is presented in Table 25 and Table 26.

Table 25: Company Submission scenario analyses

Scenario	Rationale
Hospitalisation HR = 1	Given the inherent uncertainty associated with ITC, the impact of excluding the relative benefit of sotatercept in preventing PAH-related hospitalisations is tested.
Relative risk of improvement and worsening from WHO FC ITC	An alternative source was used, relying on the Bucher ITC conducted versus selexipag, based on the GRIPHON and TRACE trials (as reported in Error! Reference source not found. of CS section Error! Reference source not found.). Since GRIPHON and TRACE did not report risk-based outcomes, WHO FC change was used as a proxy, although the outcome is not directly comparable. WHO FC is considered a more subjective measure and is less sensitive to treatment efficacy. This scenario is conducted to provide an alternative to the only available source of relative efficacy versus selexipag that was included in the base case.
Mortality approach: dependent model using Gompertz	As described in CS section Error! Reference source not found. , the dependent model to estimate long-term overall survival is tested in scenarios. The evolution of risks over time, as provided by the Gompertz and Weibull distributions, were judged to be potential alternatives to the base case choice and therefore included in the scenario analyses.
Mortality approach: dependent model using Weibull	
Mortality approach: Boucly et al 2022 + UK age	An alternative source to estimate survival is based on Boucly et al., as described in CS section Error! Reference source not found. . The scenario included updated HRs reported by Boucly, applied to the projected low risk survival also reported by Boucly (reported HRs of 2.65, 5.65 and 11.08 for intermediate-low risk, intermediate-high, and high, relative to low risk, using a Cox proportional hazards model) (3). As the mean age in the source dataset was 61.0 years, the starting age of the cohort is set to the average UK age of 58.0 years, as reported in the PAH audit. Age adjustment is applied up to the age of 61.0 years.
Sotatercept administered in high risk with 100% PCA use	Based on the uncertain evidence in high risk patients, it was assumed in the base case that patients would not receive sotatercept in the high risk state. In a scenario analysis, the impact of keeping patients on sotatercept while concurrently initiating PCA therapy is tested.
Sotatercept administered in high risk with 75% PCA use	In addition to the above assumption, based on clinician feedback, it is hypothesised that for patients receiving a

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Scenario	Rationale
	combination of sotatercept and PCA, the use of PCA would decrease to 75% of the original uptake.
Sotatercept administered in high risk with 50% PCA use	In addition to the above assumption, based on clinician feedback, it is hypothesised that for patients receiving a combination of sotatercept and PCA, the use of PCA would decrease to 50% of the original uptake.
No treatment discontinuation at 24 weeks	Clinical experts proposed base case discontinuation criteria at 24 weeks for patients who exhibit an insufficient response, which is not stipulated in the product SmPC. The impact of removing this discontinuation assumption is explored in scenario analysis.
Quality of life capped at general population values	The base case age adjustment of utility values is removed and health state utilities are capped at general population normal values if the model estimates them to be higher. This is only observed in the low risk health state.
WHO FC-based health state utility values	Given the sparsity of reported state values for risk status categories, an alternative approach using values according to WHO FC is explored, which reflects the more common approach in the literature. The scenario assumes that WHO FC-based QoL is a reasonable approximation of risk status, despite their known differences in predicting survival. Utilities based on WHO FC from the STELLAR study were derived in a similar manner to how they were derived based on risk groups.
Removal of carer quality of life	The impact of excluding the impact of carer utility is tested.
Removal of transplant disutility	The impact of excluding the impact of disutility due to transplantation is tested.
Mean weight of STELLAR ITT population	Sotatercept is a weight-based therapy, and USA trial participants were excluded from the base case mean, for reasons of generalisability. The impact of this exclusion is tested in a scenario, where the mean weight of the STELLAR ITT population is used. The mean (SD) weight was 72.3 (18.3) kg.
Sotatercept HCRU	The sotatercept SmPC specifies blood tests are required alongside the first 5 doses. The base case assumption is that these tests will be conducted remotely, however to account for difference in monitoring approaches across PAH centres, the scenario tests the impacts of this monitoring taking place face-to-face. The 5 appointments would follow an initial decision-making meeting and training session, leading to a total of 7 face-to-face appointments for all patients.
Time horizon of 10 years	An alternative time horizon of 10 years is tested.
Time horizon of 20 years	An alternative time horizon of 20 years is tested.
Discount rate of 1.5%	A discount rate of 1.5% for both costs and health benefits is used, reflecting the scenario discussed in the NICE Methods Manual. The extent and timing of the longer-term

Scenario	Rationale
	benefits of sotatercept are uncertain but the innovative mechanism of action merits this exploration.

FC: functional class; HCRU: healthcare resource utilisation; HR: hazard ratio; ICER: incremental cost-effectiveness ratio; ITC: indirect treatment comparison; ITT: intention-to-treat; LY: life-years; QALY: quality-adjusted life-years; PAH: pulmonary arterial hypertension; PCA: prostacyclin analogue; UK: United Kingdom

Table 26: Additional scenario analyses from Clarification Questions

Scenario	Rationale
Sotatercept discontinuation upon entering intermediate-high risk	The EAG requested that the company explore the impact of equivalent stopping rules between sotatercept and selexipag. Given that selexipag is only used in intermediate-low risk per treatment guidelines, the appropriate scenario is the discontinuation of sotatercept upon progression to intermediate-high risk or high risk, in alignment with the current stopping rules with selexipag.
Allowing for one-cycle improvement of sotatercept patients upon initiating parenteral PCA	The EAG noted that in the original company base case, the model allowed for a one-cycle improvement of sotatercept patients upon progression to high risk status, discontinuing sotatercept and initiating parenteral PCA, but did not offer the same improvement opportunity for patients on selexipag, which they felt was inconsistent. As a response, the company updated their base case to exclude improvement of sotatercept patients who initiate PCA in high risk. This scenario investigates the impact of the original base case assumption (patients discontinuing sotatercept have an opportunity to improve after progressing to high risk and initiating parenteral PCA).
Hospitalisation cost – Generic day admission	The EAG noted a concern that the source for hospitalisation costs (Exposto 2021) may include IV prostacyclin administration costs, resulting in double counting of the administration costs in the model. The company present an alternative hospitalisation cost scenario: replacing the Exposto value with a disease-agnostic inpatient day case cost (£1,031). Since IV prostacyclin catheterisation requires an overnight stay, this day case cost should not by definition include IV prostacyclin administration.
Hospitalisation cost – PAH inpatient admission	
IV administration disutility – halved (per selexipag PBAC submission)	The EAG expressed a concern that the administration utility decrement applied to IV infusions may be double counted i.e. already captured within the health state utility values. The company notes that the health state utility includes patients that were established on prostacyclin prior to entering the trial and so is unlikely to capture the QoL impact of PCA initiation. Also, each HSUV is derived from a mixture of patients on- and off-IV PCA (see Table 14). To address this uncertainty, this scenario uses a lower disutility value (half of the disutility reported in Davies study), which was accepted as an uncertain but plausible value in the PBAC submission for selexipag.

Company evidence submission template for Sotatercept for treating pulmonary arterial hypertension [ID6163]

Scenario	Rationale
Mean weight of intermediate-low risk population	To understand the implications of the assumptions regarding weight-based dosing, the EAG requested scenarios using the weight distribution of various subgroups in the STELLAR trial (base case: STELLAR non-USA population). The company presents scenario analyses using the mean weights of the following trial populations: (1) ITT [already presented in the CS], (2) intermediate-low risk subgroup, (3) dual background therapy subgroup, (4) triple background therapy subgroup
Mean weight of patients on dual background therapy	
Mean weight of patients on triple background therapy	

The results of the original scenario analyses and additional scenario analyses are presented in Table 27 and Table 28, respectively. In almost all scenarios, the ICER remains below the threshold of £30,000/QALY indicating the base case conclusion to be robust. The limited number of scenarios in which the scenario analysis ICER exceeds the threshold are those which employ alternative approaches to modelling mortality.

The additional scenarios requested by the EAG generally slightly increase the ICER, most notably the scenarios focused on population weight distributions or the administration utility decrement for IV infusions. However, none of the new scenarios produced an ICER above £20,000/QALY, indicating the limited impact of the parameters tested. A few of the additional scenarios slightly decreased the ICER, namely the scenario testing equivalent stopping rules for sotatercept and selexipag and the scenario with an alternative hospitalisation cost using the day case HRG code.

Table 27: Original scenario analyses deterministic results

Scenario	Incremental costs	Incremental LYs	Incremental QALYs	ICER (£/QALY)	ICER with severity modifier (£/QALY)
Updated base case	████	██	2.53	████	████
Hospitalisation HR = 1	████	██	2.49	████	████
Relative risk of improvement and worsening from WHO FC ITC	████	██	2.08	████	████
Mortality approach: dependent model using Gompertz	████	██	2.8	████	████
Mortality approach: dependent model using Weibull	████	██	2.85	████	████
Mortality approach: Boucly et al 2022 + UK age	████	██	2.37	████	████
Sotatercept administered in high risk with 100% PCA use	████	██	2.53	████	████
Sotatercept administered in high risk with 75% PCA use	████	██	2.56	████	████
Sotatercept administered in high risk with 50% PCA use	████	██	2.58	████	████
No treatment discontinuation at 24 weeks	████	██	2.64	████	████
Quality of life capped at general population values	████	██	2.56	████	████
WHO FC-based health state utility values	████	██	2.78	████	████
Removal of carer quality of life	████	██	2.43	████	████
Removal of transplant disutility	████	██	2.53	████	████
Mean weight of STELLAR ITT population	████	██	2.53	████	████
Sotatercept HCRU	████	██	2.53	████	████
Time horizon of 10 years	████	██	2.28	████	████

Time horizon of 20 years	██████	██	2.52	██████	██████
Discount rate of 1.5%	██████	██	2.82	██████	██████

ICER: incremental cost-effectiveness ratio; LY: life-years; QALY: quality-adjusted life-years

Table 28: Additional scenario analyses deterministic results

Scenario	Incremental costs	Incremental LYs	Incremental QALYs	ICER (£/QALY)	ICER with severity modifier (£/QALY)
Updated base case	██████	██	2.53	██████	██████
Sotatercept discontinuation upon entering intermediate-high risk	██████	██	1.79	██████	██████
Allowing for one-cycle improvement of sotatercept patients upon initiating parenteral PCA	██████	██	2.56	██████	██████
Hospitalisation cost – Generic day admission	██████	██	2.53	██████	██████
IV administration disutility – halved (per selexipag PBAC submission)	██████	██	2.06	██████	██████
Mean weight of intermediate-low risk population	██████	██	2.53	██████	██████
Mean weight of patients on dual background therapy	██████	██	2.53	██████	██████
Mean weight of patients on triple background therapy	██████	██	2.53	██████	██████

ICER: incremental cost-effectiveness ratio; LY: life-years; QALY: quality-adjusted life-year

References

1. Sitbon O, Channick R, Chin KM, Frey A, Gaine S, Galie N, et al. Selexipag for the Treatment of Pulmonary Arterial Hypertension. *N Engl J Med*. 2015;373(26):2522-33.
2. Humbert M, Kovacs G, Hoeper MM, Badagliacca R, Berger RMF, Brida M, et al. 2022 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension. *Eur Heart J*. 2022;43(38):3618-731.
3. Boucly A, Weatherald J, Savale L, de Groote P, Cottin V, Prevot G, et al. External validation of a refined four-stratum risk assessment score from the French pulmonary hypertension registry. *Eur Respir J*. 2022;59(6).
4. Rosenkranz S, Pausch C, Coghlan JG, Huscher D, Pittrow D, Grunig E, et al. Risk stratification and response to therapy in patients with pulmonary arterial hypertension and comorbidities: A COMPERA analysis. *J Heart Lung Transplant*. 2023;42(1):102-14.
5. Reinders S, Didden EM, Ong R. Survival, morbidity, and quality of life in pulmonary arterial hypertension patients: a systematic review of outcomes reported by population-based observational studies. *Respir Res*. 2024;25(1):373.
6. Hoeper MM, Pausch C, Olsson KM, Huscher D, Pittrow D, Grunig E, et al. COMPERA 2.0: a refined four-stratum risk assessment model for pulmonary arterial hypertension. *Eur Respir J*. 2022;60(1).
7. MSD. Patient Support Research in Preparation for Sotatercept Launch - Full Report. 2024.
8. Davies EW, Llewellyn S, Beaudet A, Kosmas CE, Gin-Sing W, Doll HA. Elicitation of health state utilities associated with the mode of administration of drugs acting on the prostacyclin pathway in pulmonary arterial hypertension. Patient preference and adherence. 2018:1079-88.
9. PBAC. Public Summary Report - July 2020 PBAC Meeting: SELEXIPAG, Tablets, 200, 400, 600, 800, 1000, 1200, 1400, 1600 microgram, Upravi®, Janssen-Cilag Pty Ltd.; 2020.

Single Technology Appraisal

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Clinical expert statement

Information on completing this form

In [part 1](#) we are asking for your views on this technology. The text boxes will expand as you type.

In [part 2](#) we are asking you to provide 5 summary sentences on the main points contained in this document.

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. 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.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. 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.

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 as '**confidential [CON]**' in turquoise, and all information submitted as '**depersonalised data [DPD]**' in pink. If confidential information is submitted, please also

Clinical expert statement

Sotatercept for treating pulmonary arterial hypertension [ID6163]

1 of 18

send a second version of your comments with that information redacted. See [Health technology evaluations: interim methods and process guide for the proportionate approach to technology appraisals](#) (section 3.2) for more information.

The deadline for your response is **5pm** on **<insert deadline>**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received, 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 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.

Part 1: Treating pulmonary arterial hypertension and current treatment options

Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality

1. Your name	Daniel Knight
2. Name of organisation	Royal Free London NHS Foundation Trust
3. Job title or position	Consultant Cardiologist and Lead for the National Pulmonary Hypertension Service, Royal Free Hospital, London
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 pulmonary arterial hypertension? ☒ A specialist in the clinical evidence base for pulmonary arterial hypertension 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 ☒ Other (they did not submit one, I do not know if they submitted one etc.) I am aware of the submission but I have not been privy to its contents and hence am unable to answer this question.
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. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	I have no disclosures with respect to the tobacco industry.

Clinical expert statement

Sotatercept for treating pulmonary arterial hypertension [ID6163]

3 of 18

8. What is the main aim of treatment for pulmonary arterial hypertension? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	Pulmonary arterial hypertension (PAH) is a progressive disease with high morbidity and mortality. The main aim of treatment for PAH is to improve symptoms and exercise capacity, enhance quality of life, prevent clinical worsening and improve survival.
9. 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)	The primary end-point in most pivotal PAH clinical trials has been an improvement in 6-minute-walk distance. The minimal clinically important difference in 6-minute-walk distance for patients with PAH is 33 metres (Moutchia et al. Am J Respir Crit Care Med. 2023;207(8):1070-1079. doi: 10.1164/rccm.202208-1547OC). Time to clinical worsening (TTCW) is also often used as a primary or secondary endpoint in PAH clinical studies, especially in pivotal phase III trials (e.g. AMBITION, SERAPHIN, GRIPHON). The European Society of Cardiology/European Respiratory Society (ESC/ERS) 2022 guidelines recognise TTCW as a relevant and validated composite endpoint, which reflects disease progression. Delaying TTCW has been shown in trials to correlate with better long-term outcomes. A statistically significant increase in TTCW (i.e. patients take longer to reach clinical worsening) is clinically significant, especially in this population which has a high mortality and hospitalisation rate. Other clinically significant treatment responses include improvements in functional class, symptoms and quality of life, reduction in NT-proBNP/BNP levels, favourable changes in right heart size and/or function on cardiac imaging, and improved haemodynamics at right heart catheter (e.g. mean pulmonary artery pressure and pulmonary vascular resistance). In clinic, we assess for treatment response in PAH using composites of the above non-invasive metrics, such as the ESC/ERS 2022 simplified four-strata risk-assessment tool and the REVEAL Lite 2 clinical risk score (Benza et al. Chest 2021;159(1):337-346. [doi: 10.1016/j.chest.2020.08.2069]).

Clinical expert statement

Sotatercept for treating pulmonary arterial hypertension [ID6163]

4 of 18

10. In your view, is there an unmet need for patients and healthcare professionals in pulmonary arterial hypertension?	There are many unmet needs for patients and healthcare professionals in PAH. Specifically, related to therapeutics: - There has been a longstanding unmet need for the development of targeted pulmonary hypertension (PH) therapies that have a mechanism of action other than pulmonary vasodilation - this has been the primary mechanism of action of PH therapies to date. - Moreover, there has previously been no therapy for PAH that has demonstrable additional clinical benefit in patients in the intermediate-high or high PH risk strata when added to intravenous prostanoid. Sotatercept addresses both of these unmet therapeutic needs.
11. How is pulmonary arterial hypertension currently treated in the NHS? • Are any clinical guidelines used in the treatment of the condition, and if so, which? • 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.) • What impact would the technology have on the current pathway of care?	Adult patients with PAH are treated in one of seven UK specialist centres. Treatment decisions depend on PAH severity/risk stratification, underlying cause, vasoreactivity and patient factors. In NHS specialist PH centre practice, diagnostic and treatment pathways generally follow the European Society of Cardiology/European Respiratory Society (ESC/ERS) 2022 guidelines. The pathway of care is well defined. Only PH specialist centres can initiate and continue long-term prescriptions of commissioned targeted PH therapies. The National Audit of Pulmonary Hypertension (NAPH) annually benchmarks referral and treatment standards across the seven UK specialist centres for adults with PH along with the National Paediatric PH Service. This supports nationally agreed quality indicators and effectively operate as a de facto NHS pathway for PAH, ensuring consistency across PH centres in key areas such as referral to treatment, time from referral to initial consultation, and performing core diagnostic tests (including right heart catheterisation) before starting targeted PH therapies. A UK PH treatment algorithm, which broadly mirrors the ESC/ERS 2022 guidelines, is currently under review with NHS England and has had extensive input from the UK National PH Physicians Group. Variation in the pathway of care of PH in England is relatively small. The ESC/ERS 2022 guidelines are well embedded in clinical practice, and NAPH

Clinical expert statement

	benchmarking and NHS England commissioning policies (e.g. riociguat, selexipag) create consistency. Minor differences may exist in terms of sequence of therapy escalation, shared-care arrangements, and expert discretion (e.g. assessment of comorbidity, borderline eligibility cases). However, overarching international guidelines minimise variation. Furthermore, the PH specialist centres have recently started to participate in a national MDT forum to discuss difficult clinical cases along with a biannual UK PH educational conference. Sotatercept would not impact the current pathway of care prior to the treatment stage for patients diagnosed with PAH, since only specialist centres can initiate and continue long-term prescriptions of commissioned targeted PH therapies. Each centre would need to explore the logistical requirements to effectively deliver this medication, with an impact as follows: (i) At least 5 extra clinic appointments during the initial 3-month monitoring phase; (ii) A priori additional full blood count (FBC) for those extra clinic appointments in the initial 3-month monitoring phase in order to ensure the next dose can be given; (iii) Extra pharmacy capacity to screen prescriptions (requiring up-to-date body weight and FBC); (iv) Clinical nurse specialist (CNS) resource and space for training patients in medication reconstitution and injection in the clinics during the initial 3-month monitoring phase; (v) Cold chain delivery via homecare (the medication needs to be kept between 2 to 8 degrees).
12. Specific treatment pathway questions • How are patients in the NHS treated when they escalate from intermediate-low to intermediate-high or high-risk states? • In the company model, patients who progress from the intermediate-low to intermediate-high risk state on	Currently, when patients with PAH treated in the NHS escalate from intermediate-low to intermediate-high or high-risk states, they are considered for intravenous Epoprostenol and suitability for referral for lung transplant assessment.

Clinical expert statement

sotatercept, are assumed to remain on sotatercept and not require initiation of PGI2 analogues, despite progressing on treatment. Is this likely in NHS practice? If not what proportion would discontinue sotatercept? • In the company model, patients who start PGI2 treatment in the high-risk state (following progression on sotatercept) would stop receiving sotatercept. Is this likely in NHS practice? If not what proportion would continue sotatercept?	If patients progress from intermediate-low to a persistently intermediate-high risk state whilst receiving sotatercept, we would consider initiating PGI2 analogues as an add-on therapy (and not discontinue sotatercept, since their risk status has progressed on PH therapies). I would not wish to withdraw sotatercept, similarly to how I would not withdraw PDE5 inhibitors or endothelin receptor antagonists in this scenario. Moreover, the PAH treatment algorithm from the 7th World Symposium on Pulmonary Hypertension (Chin KM et al. Eur Respir J 2024; 64: 2401325 [DOI: 10.1183/13993003.01325-2024]) recommends that parenteral prostacyclin therapy is considered as add-on therapy to PDE5 inhibitors, endothelin receptor antagonists and sotatercept (rather than a treatment substitution) in patients with PAH at intermediate-high or high risk at follow-up. I think it is unlikely that patients who start PGI2 treatment in the high-risk state (following progression whilst on sotatercept) would stop receiving sotatercept. This is based upon: - The evidence base from the STELLAR trial (Hoeper MM et al. N Engl J Med 2023;388:1478-1490 [DOI: 10.1056/NEJMoa2213558]) and the ZENITH trial (Humbert M et al. N Engl J Med 2025;392:1987-2000 [DOI: 10.1056/NEJMoa2415160]), in which patients at high risk on therapy including PGI2 treatment show benefit of sotatercept as an add-on therapy. - Personal experience of 23 patients receiving sotatercept in a free-of-charge patient access program that I have led at the Royal Free Hospital. All patients had PAH, were intermediate-high or high risk (by ESC/ERS 2022 four-strata PH risk criteria) and already on triple targeted PH therapies (including intravenous Epoprostenol for at least 3 months). Out of the 17 patients who have so far completed 3 months of sotatercept treatment as part of quadruple targeted PH therapy (i.e. no therapies discontinued), there was added value of add-on therapy as demonstrated by changes in their REVEAL Lite 2 clinical risk score (a PAH risk assessment that provides discrimination among patients at low, intermediate, and high risk of 1-year mortality, Benza et al. Chest 2021;159(1):337-346. [doi: 10.1016/j.chest.2020.08.2069]):

Clinical expert statement

	- 14 patients (82%) improved their REVEAL Lite 2 score and 3 (18%) had no change in their score. - Overall, sotatercept significantly reduced REVEAL Lite 2 score at 3 months in these 17 patients ($Z = -3.32$, $p < 0.001$ using Wilcoxon signed-rank test), indicating a real and meaningful treatment effect. These are unpublished data but are forming part of our single centre registry study; I can share the anonymised specific components of the dataset should the panel wish to scrutinise them.
13. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice? • How does healthcare resource use differ between the technology and current care? • In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic) • What investment is needed to introduce the technology? (for example, for facilities, equipment, or training)	Sotatercept is a first-in-class activin-signalling inhibitor in PAH. This pathway is not targeted by any other currently licenced targeted PH therapies in the UK. Sotatercept should be prescribed and managed in the long-term thereafter by designated PH specialist centres, in the same way as currently commissioned targeted PH therapies. Reasons include the rarity of PAH, comprehensive oversight required for the risk of serious adverse events (such as increased risk bleeding with sotatercept), integrated multidisciplinary team approach with specialist knowledge and training, and access to registries. Investment required to facilitate sotatercept being introduced in UK PH care might include: - Additional clinical nurse specialist (CNS) resource (both nursing staff and clinic space) to teach patients how to reconstitute and inject the medication (and assess their safety and suitability for home injection). For each patient, this is at least 5 clinic appointments across an initial 3-month period (with further appointments if deemed unsafe to self-administer). PH centres already have CNS resource, but this would have to be restructured with some centres possibly needing to consider some resource expansion in this area. - Utilisation of homecare for cold chain medication delivery. - Pharmacy resource, which includes checking the current weight and full blood count of the patient in order to screen the prescription.

Clinical expert statement

Sotatercept for treating pulmonary arterial hypertension [ID6163]

8 of 18

	- Full blood count prior to each dose for the first five doses to monitor for erythrocytosis or thrombocytopaenia. This may require assistance from shared care centres +/- primary care to facilitate for patients living far away from PH centres, so that the drug can be screened and ready for them when they travel to the centre.
14. Do you expect the technology to provide clinically meaningful benefits compared with current care? • Do you expect the technology to increase length of life more than current care? • Do you expect the technology to increase health-related quality of life more than current care?	- I expect sotatercept to increase length of life more than current care based upon: (i) The secondary endpoint of significant improvement in the distribution of time to first occurrence of death or non-fatal clinical worsening event in patients randomised to sotatercept being met in the STELLAR trial (Hoepfer MM et al. N Engl J Med 2023;388:1478-1490 [DOI: 10.1056/NEJMoa2213558]). (ii) High-risk adults with PAH who already receiving the maximum tolerated dose of background PH therapies and randomised to treatment with sotatercept had a lower risk of a composite of death from any cause, lung transplantation, or hospitalization for worsening PAH in the ZENITH trial (Humbert M et al. N Engl J Med 2025;392:1987-2000 [DOI: 10.1056/NEJMoa2415160]). - I expect sotatercept to increase health-related quality of life more than current care based upon the STELLAR trial secondary endpoints of: (i) Significant improvements in 6-minute walk distance (and thus exercise capacity); (ii) Improvement in WHO functional class; (iii) Greater improvements in quality of life as measured by the Physical Impacts and Cardiopulmonary Symptoms domain scores of the PAH-SYMPACT quality-of-life questionnaire.
15. Is this treatment more beneficial to any specific patient groups? • Is this treatment more beneficial or more appropriate for low/intermediate/high risk patients and why?	Sotatercept has an evidence-based benefit for treatment in patients with PAH in the intermediate-low, intermediate-high and high risk strata, even on maximal currently available targeted PH therapies. I believe that the treatment is most beneficial in the intermediate-high and high risk groups. These are patients in whom we have exhausted all currently

Clinical expert statement

	available treatment options. Sotatercept lowered the risk of a composite of death from any cause, lung transplantation, or hospitalization for worsening PAH in the ZENITH trial when added to maximal background PH therapies in patients with PAH in WHO functional classes 3 or 4 and a high 1-year mortality risk (Humbert M et al. N Engl J Med 2025;392:1987-2000 [DOI: 10.1056/NEJMoa2415160]).
16. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Patients with idiopathic PAH, heritable PAH, drug-induced PAH, connective-tissue disease-associated PAH, or residual PAH after shunt correction.
17. 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)	Sotatercept and its additional monitoring requirements, especially during the first 3 months, will be add-on to existing care. The requirements are relatively minor comparable to the magnitude of treatment benefit in PAH but will, nevertheless, need consideration of the additional logistics and resource required. The main requirement is for at least 5 extra clinic appointments during the initial 3-month monitoring and training phase, plus the a priori full blood count (FBC) results: - A priori FBC monitoring on a three-weekly basis for at least the first five doses (the results of which are required in order to issue the next dose), followed by three- to six-monthly FBC thereafter (the latter can be integrated to routine clinical follow-up). These FBC tests could be undertaken either in collaboration with primary care and/or shared care hospitals. - Extra pharmacy capacity to screen prescriptions (requiring up-to-date body weight and FBC); - Clinical nurse specialist (CNS) resource and space for training patients in medication reconstitution and injection in the clinics during the initial 3-month monitoring phase; - Cold chain delivery via homecare (the medication needs to be kept between 2 to 8 degrees). Patient accessibility will be affected by the at least 5 extra clinic appointments during the initial 3-month monitoring phase in order to monitor patients and teach them how to reconstitute and safely inject the medication. This will involve travelling to the PH centre (rather than shared care centres) +/- staying nearby.

Clinical expert statement

	This will have financial and time burden impacts on patients. However, this is usually only during the first 3 months of treatment.
18. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	In the majority of cases, sotatercept would be add-on therapy for PAH. However, I believe that most PH physicians would be comfortable to consider stopping selexipag in patients in whom treatment is being escalated with sotatercept. In order to inform decision-making with starting or stopping sotatercept treatment, it is unlikely that this would involve any extra significant burden of additional testing (as PH risk stratification routinely takes place in PH follow-up anyway, and this is what would be used to inform therapy changes).
19. 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? Do the instruments that measure quality of life fully capture all the benefits of the technology or have some been missed? For example, the treatment regimen may be more easily administered (such as an oral tablet or home treatment) than current standard of care	Yes. In my opinion, sotatercept for the treatment of PAH is likely to result in substantial and important health-related benefits that may not be fully captured in the standard measures used for quality-adjusted life year (QALY) calculation. Specifically: - Improvements in exercise capacity (as evidenced by significantly improved 6-minute walk distance) and symptom burden can significantly enhance patients' daily functioning, independence and emotional well-being. However, these improvements may not be fully reflected in standard quality of life (QOL) measures, which often lack sensitivity to the physical limitations experienced by patients with PAH. - A reduction in PAH-related hospitalisations has important implications not only for health-related QOL (e.g. avoiding the physical and psychological stress of acute decompensations) but also for patient safety and long-term prognosis. Avoided hospital admissions can reduce the risk of hospital-acquired complications, preserve functional status, and reduce caregiver burden and impact, none of which are directly captured in QALY metrics. Moreover, reduction in PAH-related admissions will reduce financial burden on healthcare resource.

	Therefore, while the QALY model will capture some elements of benefit of sotatercept, it may underestimate the full value of this therapy from both a patient and health system perspective.
20. 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? - Is the technology a 'step-change' in the management of the condition? - Does the use of the technology address any particular unmet need of the patient population?	Sotatercept is a first-in-class activin-signalling inhibitor in PAH and has demonstrated significant health-related benefit even in the most high risk PAH patients on maximal targeted PH therapies. Therefore, sotatercept is a step-change in the management of PAH that fulfils the unmet therapy needs in PAH of: (i) providing additional benefit beyond current commissioned therapies even in those patients with PAH who are at the highest risk of mortality; (ii) acting on reversing the underlying mechanism of disease (i.e. pulmonary arterial remodelling) rather than solely targeting pulmonary arterial dilatation.
21. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	The main side-effects of sotatercept are a higher rate of bleeding (mostly non-serious such as epistaxis or gingival bleeding), raised haemoglobin, thrombocytopaenia and telangiectasia. Raised haemoglobin or thrombocytopenia are generally encountered in the first 5 doses during the initial 3-month period of monitoring at the PH centre and can be managed as per dosing protocol. Most bleeding side-effects are minor. However: (i) In the STELLAR study, patients who developed serious bleeding were more likely to be on background prostacyclin therapy, taking anticoagulation, or have a low platelet count. Hence caution should be applied in these settings and patients should be advised accordingly. (ii) The overall risk of serious bleeding events (including gastrointestinal and intracranial haemorrhage) is not trivial (approximately 4-5%) and seems to accumulate over prolonged treatment. Hence ongoing clinical monitoring should be performed by PH specialist centres.

Clinical expert statement

22. Do the clinical trials on the technology reflect current UK clinical practice? • 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? • 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?	The two main clinical trials of sotatercept to consider are the STELLAR trial (Hoeper MM et al. N Engl J Med 2023;388:1478-1490 [DOI: 10.1056/NEJMoa2213558]) and the ZENITH trial (Humbert M et al. N Engl J Med 2025;392:1987-2000 [DOI: 10.1056/NEJMoa2415160]). The patient cohorts are largely generalisable and can be extrapolated to the UK setting. The STELLAR cohort (FC II-III with background PAH therapy) and the ZENITH cohort (high risk FC III-IV already on maximal double or triple PAH therapy) reflect typical UK patient populations in specialist centres. The only difference of note was a degree of underrepresentation of patients with connective tissue disease-associated PAH, which is a common aetiology of PAH in UK clinical practice. The endpoints were clinically meaningful and relevant (the primary endpoint in STELLAR was a significant and clinically meaningful improvement (40 metres) in 6-minute walk distance at 24 weeks; the primary endpoint in ZENITH was a hard composite endpoint of a 76% risk reduction (HR 0.24) of all-cause death, lung transplantation, or PAH hospitalization, prompting early trial stoppage). Moreover, clinically relevant surrogate secondary endpoints in the STELLAR trial, such as improvements in NT-proBNP, WHO FC, pulmonary vascular resistance, time to non-fatal clinical worsening and quality-of-life metrics, were met. These surrogate outcome measures are well validated in PAH and are predictive of long-term clinical outcome. In summary: - The significant improvements in the primary and secondary endpoints in the STELLAR trial in patients with PAH who are WHO FC II-III can be expected with sotatercept in UK patients with similar background treatment regimes. - The ZENITH trial data supports targeting high-risk patient with PAH early, adding it to background PH therapy in WHO FC III-IV patients to reduce hospitalisation and transplantation risk.

23. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No.
24. How do data on real-world experience compare with the trial data?	I have personal real-world experience of 23 patients with PAH receiving sotatercept in a free-of-charge patient access program that I have led at the Royal Free Hospital. I believe that this is the largest roll out of the program in the 6 PH centres in England. All patients had PAH and were at intermediate-high or high risk of mortality (as defined by ESC/ERS 2022 four-strata PH risk criteria). All patients were already on triple targeted PH therapies (including intravenous Epoprostenol for at least 3 months). I have calculated the REVEAL Lite 2 clinical risk scores for all patients at baseline and 3 months post-treatment from our prospectively acquired datasets (with no missing or imputed data). This score is a real-world, non-invasive simplified risk assessment that we use in clinic and that provides discrimination among patients with PAH at low, intermediate, and high risk of 1-year mortality (Benza et al. Chest 2021;159(1):337-346. [doi: 10.1016/j.chest.2020.08.2069]): - Out of the 17 patients who have so far completed 3 months of sotatercept treatment as part of quadruple targeted PH therapy, 14 (82%) improved their REVEAL Lite 2 score and 3 (18%) had no change in their score. - Overall, sotatercept significantly reduced REVEAL Lite 2 score at 3 months in these 17 patients ($Z = -3.32$, $p < 0.001$ using Wilcoxon signed-rank test), indicating a real, favourable and meaningful treatment effect. These are unpublished data but are forming part of our single centre registry study; I can share the anonymised components of the dataset should the panel wish to scrutinise them.
25. What are the appropriate target doses for PGI2 analogue preparations used in NHS clinical practice? • The company uses the following doses for PGI2 analogue preparations. Are these applicable to what is seen in clinical practice? • Epoprostenol: 50ng/kg/min	These are all applicable to what we use in clinical practice.

Clinical expert statement

• Treprostinil: 30ng/kg/min • Inhaled Iloprost: 5.0 microgms x (6-9 times per day).	
26. NICE considers whether there are any equalities issues at each stage of an evaluation. Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged. Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics. Please state if you think this evaluation could • exclude any people for which this treatment is or will be licensed but who are protected by the equality legislation • lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population • lead to recommendations that have an adverse impact on disabled people. Please consider whether these issues are different from issues with current care and why. More information on how NICE deals with equalities issues can be found in the NICE equality scheme.	There are no obvious equality issues that I can foresee with respect to this evaluation in terms of protected characteristics. Patient accessibility is a key issue for the minimum of 5 extra clinic appointments required during the initial 3-month monitoring phase. These appointments during the initial 3-months are required to ensure that FBC parameters are within safe limits, to check for side-effects and adverse events, and to train patients in how to reconstitute and safely inject the medication for the longer term. This will involve patients travelling to their PH centre (rather than shared care centres) +/- staying nearby in local accommodation. Patient accessibility to sotatercept will, therefore, be affected by factors including their symptom burden, financial mobility and/or time commitments (e.g. due to time off work). Whilst this is usually only during the first 3 months of treatment, it may be a significant barrier for some patients. This could be offset, at least in part, through the provision of hospital transport for some patients.

Clinical expert statement

Sotatercept for treating pulmonary arterial hypertension [ID6163]

15 of 18

[Find more general information about the Equality Act and equalities issues here.](#)

Part 2: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

- (1) Unlike traditional targeted PH therapies that work solely by pulmonary vasodilation, sotatercept is a first-in-class activin-signalling inhibitor that addresses the underlying disease driver, potentially reversing pulmonary arterial remodelling.
- (2) The characteristics of patient cohorts in the main clinical trials of sotatercept, namely STELLAR and ZENITH, are largely generalisable and can be extrapolated to the UK setting.
- (3) The significant improvements in the primary and secondary endpoints in the STELLAR trial in patients with PAH who are WHO FC II-III can be expected with sotatercept in UK patients with similar background treatment regimes.
- (4) The ZENITH trial supports targeting high-risk patient with PAH early, adding it to background PH therapy in WHO FC III-IV patients with PAH to reduce hospitalisation and transplantation risk.
- (5) My personal 'real-world' experience of sotatercept as part of a patient access program (in 23 patients with PAH at intermediate-high or high risk of mortality despite background triple PH therapy including parenteral prostanoid) is that this technology is a step-change in the management of PAH, with significant improvements in symptoms and clinical risk profile.

Thank you for your time.

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.

Clinical expert statement

Sotatercept for treating pulmonary arterial hypertension [ID6163]

17 of 18

For more information about how we process your personal data please see our [privacy notice](#).

Single Technology Appraisal

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Clinical expert statement

Information on completing this form

In [part 1](#) we are asking for your views on this technology. The text boxes will expand as you type.

In [part 2](#) we are asking you to provide 5 summary sentences on the main points contained in this document.

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. 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.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. 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.

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 as '**confidential [CON]**' in turquoise, and all information submitted as '**depersonalised data [DPD]**' in pink. If confidential information is submitted, please also

Clinical expert statement

Sotatercept for treating pulmonary arterial hypertension [ID6163]

1 of 12

send a second version of your comments with that information redacted. See [Health technology evaluations: interim methods and process guide for the proportionate approach to technology appraisals](#) (section 3.2) for more information.

The deadline for your response is **5pm on Monday 23 June 2025**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received, 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 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.

Part 1: Treating pulmonary arterial hypertension and current treatment options

Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality

1. Your name	Dr Iain Armstrong
2. Name of organisation	Pulmonary Hypertension Association UK (PHA UK)
3. Job title or position	Chairman and co-founder
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 pulmonary arterial hypertension? ☐ A specialist in the clinical evidence base for pulmonary arterial hypertension or technology? ☒ Other (please specify): Nurse Consultant in PH (30 years) and Chair of the patient organisation Pulmonary Hypertension Association UK (25 years)
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 did not submit one, I do not 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. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	None
8. What is the main aim of treatment for pulmonary arterial hypertension?	To improve symptoms and function, prevent deterioration and slow disease progression, and improve quality of life.

Clinical expert statement

Sotatercept for treating pulmonary arterial hypertension [ID6163]

3 of 12

(For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	
9. 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)	I consider all the following to be clinically significant treatment responses: A reduction in symptoms and improved function (particularly being able to engage in day-to-day life in a more fulfilling capacity). Overall improvement in quality of life, by either personal testimony or validated patient reported outcome measures (such as emPHAsis-10). Reduction in deterioration and / or hospital admissions.
10. In your view, is there an unmet need for patients and healthcare professionals in pulmonary arterial hypertension?	Yes. We have been treating PAH in the UK for over 30 years, yet the majority of patients remain in WHO functional class 3 or 2. Our recent survey work at the PHA UK, involving almost 300 people living with PH, demonstrated a very high burden of daily symptoms. The three symptoms that respondents identified as affecting them the most are breathlessness during daily living activities, fatigue, and poor sleep. The results are as yet unpublished (the research was conducted very recently) but they clearly demonstrate that even with the best treatments in the world, the burden of symptoms in PAH remains high and there is still much to do.
11. How is pulmonary arterial hypertension currently treated in the NHS? • Are any clinical guidelines used in the treatment of the condition, and if so, which? • Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals	As Chairman of the PHA UK, I believe that in regard to this technology, it is more appropriate this question is answered by the clinical teams and those involved in commissioning services.

Clinical expert statement

Sotatercept for treating pulmonary arterial hypertension [ID6163]

4 of 12

across the NHS? (Please state if your experience is from outside England.) What impact would the technology have on the current pathway of care?	
12. Specific treatment pathway questions - How are patients in the NHS treated when they escalate from intermediate-low to intermediate-high or high-risk states? - In the company model, patients who progress from the intermediate-low to intermediate-high risk state on sotatercept, are assumed to remain on sotatercept and not require initiation of PGI2 analogues, despite progressing on treatment. Is this likely in NHS practice? If not what proportion would discontinue sotatercept? - In the company model, patients who start PGI2 treatment in the high-risk state (following progression on sotatercept) would stop receiving sotatercept. Is this likely in NHS practice? If not what proportion would continue sotatercept?	As Chairman of the PHA UK, I believe that in regard to this technology, it is more appropriate this question is answered by the clinical teams and those involved in commissioning services
13. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice? - How does healthcare resource use differ between the technology and current care? - In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic)	There would be a human resource consideration as the initiation of this therapy requires hospital visits every three weeks for the first 4-5 injections. I perceive this will have an impact on clinic slots and potentially hospital transportation. It may potentially present difficulties for those whose symptoms make travel one of their greatest challenges. There will be an implication for patients and their loved ones as this is a long-term self-administered therapy requiring injections every three weeks for the duration of the pathway.

Clinical expert statement

Sotatercept for treating pulmonary arterial hypertension [ID6163]

5 of 12

What investment is needed to introduce the technology? (for example, for facilities, equipment, or training)	
14. Do you expect the technology to provide clinically meaningful benefits compared with current care? - Do you expect the technology to increase length of life more than current care? - Do you expect the technology to increase health-related quality of life more than current care?	I feel unable to pass comment on increased life expectancy. However I can say with clinical and academic confidence that the early access programme taking place in the UK at present is showing emerging evidence of the following: A significant proportion (specific numbers unknown at present) of patients on triple therapy (including very invasive IV therapy) are expressing improvement in their quality of life and in their ability to engage more meaningfully and fully in day-to-day life. I am the co-creator of emPAHsis-10, the disease-specific patient reported outcome measure which has been used to measure quality of life in PAH across the world since 2011. I believe sotatercept will show sustained and meaningful improvement in quality of life for people with PAH.
15. Is this treatment more beneficial to any specific patient groups? Is this treatment more beneficial or more appropriate for low/intermediate/high risk patients and why?	As Chairman of the PHA UK, I believe that in regard to this technology, it is more appropriate this question is answered by the clinical teams and those involved in commissioning services
16. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	I am unsure of what this question is asking specifically (with reference to the 'general population').
17. 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?	The answer to this is the same as to question 13: There would be a human resource consideration as the initiative of this therapy requires hospital visits every three weeks for the first 4-5 injections. I perceive this will have an impact on clinic slots, potentially hospital transportation, and

Clinical expert statement

(For example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed)	potentially present difficulties for those whose symptoms make travel one of their greatest challenges. There will be an implication for patients and their loved ones as this is a long-term self-administered therapy requiring injections every 3 weeks for the duration of the pathway.
18. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	As Chairman of the PHA UK, I believe that in regard to this technology, it is more appropriate this question is answered by the clinical teams and those involved in commissioning services
19. 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? Do the instruments that measure quality of life fully capture all the benefits of the technology or have some been missed? For example, the treatment regimen may be more easily administered (such as an oral tablet or home treatment) than current standard of care	I can say with clinical and academic confidence that the early access programme taking place in the UK at present is showing emerging evidence of the following: A significant proportion (specific numbers unknown at present) of patients on triple therapy (including very invasive IV therapy) are expressing improvement in their quality of life and in their ability to engage more meaningfully and fully in day-to-day life. I believe sotatercept will show sustained and meaningful improvement in quality of life for people with PAH. In addition, all clinical trials in PAH so far report quality of life as a secondary endpoint (if it is reported at all). This is a frustrating issue as our own research over the last 17 years shows patients consistently place quality of life as their primary goal of treatment. This is placed above increased life expectancy and treatment burden. So, in summary, the instruments that measure quality of life do not fully capture all the benefits of the technology.

Clinical expert statement

	We do feel this question has been covered in previous answers (specifically question 14).
20. 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? - Is the technology a 'step-change' in the management of the condition? - Does the use of the technology address any particular unmet need of the patient population?	The unmet need right now in 2025, as stated further up this form, is the lack of a treatment pathway that fully addresses the high symptom burden that confines PAH patients to a life in WHO functional classes 2 and 3. Therefore, sotatercept does have the potential to be a step-change in the management of PAH.
21. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	I am aware this is a therapy that may cause increased episodes of bleeding such as significant nosebleeds, telangiectasia, increased haemoglobin levels and blood platelet counts, dizziness, fainting, and headaches. These are very common side effects in most treatment pathways in PAH. I do not believe they are significant and problematic. They will not affect quality of life as this patient population is appropriately managed and supported within specialist care services in the UK.
22. Do the clinical trials on the technology reflect current UK clinical practice? - 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?	Our PHA UK research over the last 17 years shows patients consistently place quality of life as their primary goal of treatment. This is placed above increased life expectancy and treatment burden. However, all clinical trials in PAH so far (not just in sotatercept) report quality of life as a secondary endpoint (if it is reported at all).

Clinical expert statement

Sotatercept for treating pulmonary arterial hypertension [ID6163]

8 of 12

• 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?	
23. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No
24. How do data on real-world experience compare with the trial data?	This real-world data is only just emerging from the early access programme in the UK. However it does seem to be reflecting, so far, what the clinical trials have reported.
25. What are the appropriate target doses for PGI2 analogue preparations used in NHS clinical practice? • The company uses the following doses for PGI2 analogue preparations. Are these applicable to what is seen in clinical practice? • Epoprostenol: 50ng/kg/min • Treprostinil: 30ng/kg/min • Inhaled Iloprost: 5.0 microgms x (6-9 times per day).	As Chairman of the PHA UK, I believe that in regard to this technology, it is more appropriate this question is answered by the clinical teams and those involved in commissioning services.
26. NICE considers whether there are any equality issues at each stage of an evaluation. Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged. Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil	I do not believe there are any equality issues with this evaluation, because this therapy would be delivered through specialist care. This care is delivered by teams with decades of experience treating people with additional needs - for example those with connective tissue disease causing dexterity issues that may affect their ability to self-administer the technology.

Clinical expert statement

Sotatercept for treating pulmonary arterial hypertension [ID6163]

9 of 12

partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics.

Please state if you think this evaluation could

- exclude any people for which this treatment is or will be licensed but who are protected by the equality legislation
- lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population
- lead to recommendations that have an adverse impact on disabled people.

Please consider whether these issues are different from issues with current care and why.

More information on how NICE deals with equalities issues can be found in the [NICE equality scheme](#).

[Find more general information about the Equality Act and equalities issues here.](#)

Part 2: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

This is a first-in-class drug in PAH and the clinical trial results are very encouraging.

The UK early access programme is showing, with real-world data, that the clinical trial outcomes are being reflected in the more complex clinical environment.

Once established, I believe this will be an easily administered therapy which when managed within specialist care services will prevent any significant untoward side effects being undetected.

The real-world data (from the early access programme) is showing that patients who are on multiple PAH therapies are showing an improvement with sotatercept after many years on the present treatment pathways.

The UK early access programme, and lived experiences from across the globe where this treatment is already in place, is showing some significant improvement in quality of life – which patients in the UK consistently place as their most important treatment outcome (evidenced by our research at the PHA UK over the last 17 years).

Thank you for your time.

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.

Clinical expert statement

Sotatercept for treating pulmonary arterial hypertension [ID6163]

11 of 12

For more information about how we process your personal data please see our [privacy notice](#).

Sotatercept for treating pulmonary arterial hypertension [ID6163]

Produced by Aberdeen HTA Group

Authors Clare Robertson¹
Dwayne Boyers²
Thenmalar Vadiveloo¹
David Cooper¹
Melanie Brewis³
Colm McCabe⁴
Paul Manson¹
Miriam Brazzelli¹
Graham Scotland²

1 Health Services Research Unit, University of Aberdeen, UK

2 Health Economics Research Unit, University of Aberdeen, UK

3 Consultant Respiratory Physician, Scottish Pulmonary Vascular Unit

4 Consultant Respiratory Physician, National Pulmonary Hypertension Service
Royal Brompton Hospital, part of GSTT Foundation Trust

Correspondence to Graham Scotland
Health Economics Research Unit, University of Aberdeen
2nd Floor, Health Sciences Building, Foresterhill
Aberdeen, AB25 2ZD
g.scotland@abdn.ac.uk

Date completed 21 May 2025

Contains 

Copyright belongs to the Aberdeen HTA Group, University of Aberdeen unless otherwise stated.

Source of funding: This report was commissioned by the NIHR Evidence Synthesis Programme as project number **NIHR166802**.

Declared competing interests of the authors

C McCabe declares speaker fees and educational grant support, Janssen Pharmaceuticals & Advisory board participation, MSD. M Brewis declares participation as a sub investigator on commercial studies for PH therapies that are referred to in the report, including GRIPHON, AMBITION, SOTERIA, TRITON.

Acknowledgements

Copyright is retained by Merck Sharp & Dohme (UK) Limited for Tables 3, 7, 9, 12, 17 and 22, and for Figures 1, 2, 3, 6 and 7 and for text referenced on pages 8-9.

Rider on responsibility for report

The view expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Robertson C, Boyers D, Vadiveloo T, Cooper D, Brewis M, McCabe C, Manson P, Brazzelli B, Scotland G. Sotatercept for treating pulmonary arterial hypertension [ID6163]. NICE Single Technology Appraisal, Aberdeen HTA Group, 2025.

Contribution of authors

CR reviewed and critiqued the clinical effectiveness evidence presented in the company submission and drafted the background section; TV and DC checked and critiqued the statistical analyses presented in the company submission; Miriam Brazzelli oversaw and commented on the critique of clinical effectiveness; DB and GS reviewed and critiqued the cost-effectiveness evidence and economic model presented in the company submission; Melannie Brewis and CB provided expert clinical guidance and comments on the draft report; PM checked and critiqued the company's search strategies. GS coordinated all aspects of this appraisal and is the guarantor of this report. All authors contributed to the writing of this report and approved its final version.

Table of Contents

Table of Contents	iii
List of Tables	vi
List of Figures.....	vii
List of abbreviations	1
1. Executive summary.....	1
1.1 Overview of the EAG's key issues	2
1.2 Overview of key model outcomes	4
1.3 The decision problem: summary of the EAG's key issues	5
1.4 The clinical effectiveness evidence: summary of the EAG's key issues	6
1.5 The cost-effectiveness evidence: summary of the EAG's key issues	7
1.6 Summary of EAG's preferred assumptions and resulting ICER.....	14
2 INTRODUCTION AND BACKGROUND	16
2.1 Introduction.....	16
2.2 Background	16
2.3 Critique of company's definition of decision problem	20
3 CLINICAL EFFECTIVENESS.....	31
3.1 Critique of the methods of review(s)	31
3.2 Critique of trials of the technology of interest, the company's analysis and interpretation (and any standard meta-analyses of these).....	33
3.2.2 <i>Primary and secondary efficacy endpoints</i>	36

3.2.3	<i>Subgroup analyses</i>	37
3.2.4	<i>Adverse events</i>	37
3.3	Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison	42
3.4	Critique of the indirect comparison and/or multiple treatment comparison....	47
3.5	Critique of the within trial analysis.....	48
3.6	Additional work on clinical effectiveness undertaken by the EAG.....	48
3.7	Conclusions of the clinical effectiveness section.....	49
4	COST EFFECTIVENESS.....	50
4.1	EAG comment on company’s review of cost-effectiveness evidence.....	50
4.2	Summary and critique of the company’s submitted economic evaluation by the EAG	50
4.2.1	<i>NICE reference case checklist</i>	50
4.2.2	<i>Model structure</i>	52
4.2.3	<i>Population</i>	55
4.2.4	<i>Interventions and comparators</i>	56
4.2.5	<i>Perspective, time horizon and discounting</i>	56
4.2.6	<i>Treatment effectiveness and extrapolation</i>	57
4.2.7	<i>Health related quality of life</i>	66
4.2.8	<i>Resources and costs</i>	73
5	COST EFFECTIVENESS RESULTS	85
5.1	Company’s cost effectiveness results.....	85

5.2	Company’s sensitivity analyses.....	86
5.3	Model validation and face validity check.....	86
6	EVIDENCE REVIEW GROUP’S ADDITIONAL ANALYSES	91
6.1	Exploratory and sensitivity analyses undertaken by the EAG.....	91
6.2	Impact on the ICER of additional clinical and economic analyses undertaken by the EAG	97
6.3	EAG’s preferred assumptions	102
6.4	Conclusions of the cost effectiveness section	112
8	References	113

List of Tables

Table 1 Summary of the key issues identified by the EAG	3
Table 2 Summary of the EAG preferred modelling assumptions.....	14
Table 3 2022 ESC/ERS simplified four-strata risk assessment tool for guidance of therapeutic decision-making.....	17
Table 4 Summary of the company's decision problem	21
Table 5 EAG's appraisal of the literature review methods presented in the CS.....	31
Table 6 Quality assessment of the company's systematic literature review of clinical effectiveness evidence	33
Table 7 Summaries of the STELLAR and SOTERIA methodologies	34
Table 8 Summary of key adverse events reported in STELLAR	39
Table 9 Summary demographics and baseline characteristics of the trials included in the indirect treatment comparison.....	43
Table 10 NICE reference case checklist	50
Table 11 Summary of company and EAG preferred HSUVs.....	72
Table 12 Comparison of alternative estimates of patient weight used in treatment acquisition cost calculations [reproduced from Tables 17 and 18 of the company response to the clarification letter].....	75
Table 13 Comparison of company and EAG preferred sotatercept treatment administration costs.....	77
Table 14 Comparison of company and EAG preferred proportions of patients receiving PGI2 treatments.....	80
Table 15 Comparison of company and EAG preferred PGI2 costs.....	81
Table 16 Summary of company and EAG preferred costing assumptions.....	83
Table 17 Deterministic base case results (Source: Table 21 of the company response to the clarification letter).....	86
Table 18 EAG scenarios around the company corrected base case (explanation).....	92
Table 19 EAG scenarios around the company base case (results).....	98
Table 20 Results using EAG's preferred model assumptions with and without severity weighting (x1.2 where appropriate).....	105
Table 21 Alternative EAG Deterministic base case results	107
Table 22 Probabilistic EAG base case (Mid) results	108
Table 23 Scenario analysis around the EAG base case (Mid)	110

List of Figures

Figure 1 Pulmonary arterial hypertension treatment algorithm based on the European society of cardiology/European respiratory society guidelines.....	18
Figure 2 Pulmonary arterial hypertension outline of overall treatment recommendations/treatment algorithm by the World Symposium on Pulmonary Hypertension	19
Figure 3 Markov cohort traces for the (a) sotatercept and (b) selexipag arm of the company model under company base case assumptions (Source: Company model).....	88
Figure 4 Alive in progression free states (Model output versus GRIPHON Primary Outcome for selexipag)	90
Figure 5 Percentage of patients remaining alive in low-risk or intermediate-low risk states for the company base case and alternative EAG base case options	107
Figure 6 Cost-effectiveness scatter plot (1000 iterations) – EAG base case (Mid) with severity modifier applied	109
Figure 7 Cost-effectiveness acceptability curves (1000 iterations) – EAG base case (Mid) with severity modifier applied	109

List of abbreviations

ActRIIA	Activin receptor type IIA
CRD	Centre for Reviews and Dissemination
CS	Company submission
CI	Confidence interval
EAG	External Assessment Group
ERA	Endothelin receptor antagonists
ERS	European Respiratory Society
ESC	European Society of Cardiology
FC	Functional class
HR	Hazard ratio
IgG	Immunoglobulin G
NT-proBNP	N-terminal pro-brain natriuretic peptide
OLE	Open label extension
PAH	Pulmonary arterial hypertension
PAH-SYMPACT	pulmonary arterial hypertension-symptoms and impact questionnaire
PDE5i	Phosphodiesterase-5 inhibitor
RCT	Randomised controlled trial
RoB	Risk of bias
SD	Standard deviation
SLR	Systematic literature review
TGF-β	Transforming growth factor beta
WHO	World Health Organisation

1. Executive summary

This summary provides a brief overview of the key issues identified by the external assessment group (EAG) as being potentially important for decision making. It also includes

the EAG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 provides an overview of key model outcomes and the modelling assumptions that have the greatest effect on the ICER. Sections 1.3 to 1.6 explain the key issues in more detail. Background information on the condition, technology and evidence and information on non-key issues are in the main EAG report.

All issues identified represent the EAG's view, not the opinion of NICE.

1.1 Overview of the EAG's key issues

The focus of the submission received from MSD is sotatercept for the treatment of pulmonary arterial hypertension (PAH). The company's positioning of sotatercept in the care pathway is as an alternative therapy to standard care, comprising phosphodiesterase 5 inhibitor, endothelin receptor antagonists and selexipag, for PAH patients with intermediate-low risk status according to the European Society of Cardiology/European Respiratory Society (ESC/ERS) 4-strata risk rating system.

In the CS, the main clinical effectiveness evidence for sotatercept is obtained from the multicentre, double-blind, placebo controlled, phase 3 randomised controlled trial, STELLAR(n=323) and the ongoing open label, long-term extension study, SOTERIA (n=426). An indirect treatment comparison with selexipag is made using data from the GRIPHON and TRACE trials.

Table 1 Summary of the key issues identified by the EAG

ID[6163]	Summary of issue	Report sections
1	Limitations of the evidence for the comparative efficacy of sotatercept versus selexipag on ESC/ERS risk strata deterioration/improvement, relying on a post-hoc subgroup analysis of STELLAR, comparing ERA+PDE5i+sotatercept with ERA+PDE5i+seleixpag+placebo.	3.6 4.2.6
2	Use of relative risks from the post-hoc subgroup analysis of STELLAR, comparing ERA+PDE5i+sotatercept with ERA+PDE5i+selexipag+placebo, to inform short term transition probabilities for selexipag is likely to be biased.	4.2.6
3	Ongoing application of a short-term (24 week) relative risk of deterioration, to generate long-term transition probabilities for selexipag, results in an overly pessimistic extrapolation of risk status.	4.2.6
4	The simplifying assumption of allowing no clinical improvement following initiation of PGI ₂ analogues, lacks plausibility and may bias against selexipag.	4.2.2
5	Timing and approach to PGI ₂ initiation in the sotatercept arm of the model, is inconsistent with clinical expert feedback received by the EAG.	4.2.2 4.2.6
6	The proportion of patients who receive PGI ₂ analogues, when indicated according to the modelling assumptions is high	4.2.8
7	The weight-based doses applied to the IV PCA preparations included in the model are high compared to European guidelines and clinical expert feedback to the EAG.	4.2.8

The key differences between the company and EAG assumptions are that the EAG prefer to: inform comparator (selexipag) transition probabilities of ESC/ERS risk state improvement/deterioration using relative risks for WHO-FC improvement/deterioration (as proxies) derived from the company's ITC, rather than relative risks based on a post-hoc subgroup analysis of STELLAR; remove a health state utility value decrement applied to those on IV prostacyclin analogues (PCA); assume infused PCA therapies are added to

sotatercept for a proportion of patients who progress to the intermediate-high risk state on sotatercept; assume that 85% rather than 100% of patients receive infused PCA following progression to the high risk state or intermediate-high risk state on selexipag; apply dosing assumptions for epoprostenol and Treprostinil (IV PCA preparations) that are aligned with target maintenance doses in ESC/ERS guidelines.

1.2 Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (overall survival) and quality of life in a quality-adjusted life year (QALY). An ICER is the ratio of the extra cost for every QALY gained.

The company used a Markov model with health states reflecting the ESC/ERS 4-strata risk status (low, intermediate-low, intermediate-high and high), which reflect increasing mortality risk based on several clinical measures.

Overall, the technology, sotatercept, is modelled to affect QALYs by:

- 1) Effects of treatment on probabilities of improvement/deterioration between ESC / ERS risk states, with increasing degrees of risk severity associated with decreasing health state utility values (HSUVs) as derived from the STELLAR trial data.
- 2) Prolonged survival, mediated through the effects of treatment on ESC/ERS risk status.
- 3) HSUV decrements applied to patients requiring IV PCA treatment when they progress to higher risks states.
- 4) HSUV decrements applied to patients experiencing PAH related hospitalization or transplant events, which increase in probability as patients progress to higher ESC/ERS risk states. A direct effect of sotatercept is also applied to the hospitalization risk in the low and intermediate-low risks states of the model.
- 5) Utility benefits applied to carers of patients who remain in lower risk health states relative to the utility experienced by carers of patients in the high-risk health state.

Overall, the technology is modelled to affect costs by:

- 1) Affecting treatment acquisition and initiation costs compared to the comparator treatment selexipag

- 2) Slowing progression to higher risk states, in which treatment with infused prostacyclin analogues is initiated
- 3) Effects of treatment on hospitalisation costs, mediated through slowed progression to higher risk states, and a direct effect applied in the low and intermediate-low risk states.
- 4) Effects on lung-transplantation, which occur with a low probability in the intermediate-high and high risks states.
- 5) Through prolonging survival compared to the comparator treatment.

The modelling assumptions that have the greatest effect on the ICER are:

- The relative effect of treatment with sotatercept versus selexipag on long-term transitions probabilities between the ESC/ERS risk strata health states
- The proportion of patients who receive infused PCA treatments upon progression to the higher risk states, and the point at which such treatment is initiated in each arm of the model
- Assumptions about whether IV PCA is administered in addition to sotatercept or instead of it when it is initiated in higher risk states.
- The assumed weight-based doses of the IV PCA preparations that are applied in the model
- The HSUV decrement applied to IV PCA use.
- The assumed weight-based dose of sotatercept that is applied.

1.3 The decision problem: summary of the EAG's key issues

Overall, the company's decision problem aligns with the NICE final scope. However, the EAG notes that the company defined a narrower population than the NICE scope, which includes all adults with pulmonary arterial hypertension. The company focused specifically on adults with PAH, (WHO Group 1), who are at intermediate-low risk according to the

ESC/ERS four-strata risk rating system. The EAG clinical experts agree that intermediate-low risk PAH patients are most likely to receive sotatercept in the UK, but they also note that there may be some overlap with intermediate-high risk patients who could also benefit from treatment with sotatercept. The EAG also notes that PDE5i + ERA + selexipag is a relevant comparator for this appraisal, however, limiting the appraisal to this combination of therapy excludes those intermediate-low risk patients who have either not responded to selexipag or who have responded to IV prostacyclin therapy.

1.4 The clinical effectiveness evidence: summary of the EAG's key issues

Generally, the EAG are happy with the indirect treatment comparisons presented by the Company. The EAG however does have a concern with the method used by the Company to compare sotatercept and selexipag for ESC/ERS risk status. This has been done within the STELLAR study by the Company, comparing the subgroups of participants randomised to sotatercept and previously treated with PDE5i + ERA to participants randomised to placebo and previously treated with PDE5i + ERA + selexipag. In this second group, selexipag is not the active intervention. Participants are already stabilised on their prior treatment combination and the effect of receiving selexipag will therefore not be seen in this group. This is not the case for the group of participants randomised to sotatercept where the full effect of the intervention will be seen.

Issue 1 Within-trial analyses of PDE5i + ERA + sotatercept versus PDE5i + ERA + selexipag

Report section	3.5
Description of issue and why the EAG has identified it as important	To assess the outcome of change from baseline in risk status at week 24, the company has conducted a post-hoc within-trial analysis of the STELLAR study. This method does not allow for a direct comparison of the treatment effects of sotatercept and selexipag as participants in the placebo arm were receiving selexipag prior to randomisation. Therefore, any potential treatment effect of selexipag will not be observable.
What alternative approach has the EAG suggested?	The EAG suggest that the ITC with the GRIPHON trial of selexipag offers potentially less biased estimates of relative risks of improvement/deterioration, albeit for the World Health Organisation-Functional Class (WHO FC) rather than ESC/ERS risks stratification.
What is the expected effect on the cost-effectiveness estimates?	Applying more conservative relative risks of risks status improvement deterioration in the economic model, increases the ICER for sotatercept. See Issues 2 and 3 below.
What additional evidence or analyses might help to resolve this key issue?	This issue cannot be resolved unless the company can get hold of IPD from a registry or the GRIPHON trial, and risk status could be estimated retrospectively for comparison with sotatercept.

1.5 The cost-effectiveness evidence: summary of the EAG's key issues

The EAGs main issues with the company's cost-effectiveness evidence are summarised in boxes 2 to 7 below.

Issue 2 Use of relative risks from the post-hoc subgroup analysis of STELLAR to inform transition probabilities for selexipag

Report section	4.2.6
Description of issue and why the EAG has identified it as important	The company use relative risks from the post-hoc subgroup analysis of STELLAR, comparing ERA+PDE5i+sotatercept with ERA+PDE5i+seleixpag+placebo, to inform transition probabilities for selexipag relative to those for sotatercept (estimated from STELLAR data). For the reasons outlined in issue one (above), the EAG believe these relative risks to be at high risk of bias in favour of sotatercept.
What alternative approach has the EAG suggested?	Whilst recognising the limitations, the EAG believe the company's scenario analysis using relative risks for WHO Functional Class (WHO FC) improvement/deterioration, derived from an ITC with the GRIPHON trial of selexipag versus placebo, may offer a less biased effects for use in the model - as proxies for relative effect on ESC/ERS risks strata improvement/deterioration.
What is the expected effect on the cost-effectiveness estimates?	Applying these more conservative relative treatment effects, to derive short and long-term transition probabilities for selexipag in the model, results is a substantial upward shift in the ICER.
What additional evidence or analyses might help to resolve this key issue?	The EAG acknowledge the challenges around generating evidence for the comparative efficacy of sotatercept versus selexipag on ESC/ERS risk improvement/deterioration. However, further expert opinion on the face validity of the modelled projections of progression to intermediate-high and high risks states may be beneficial.

Issue 3 Ongoing application of a short-term (24 week) relative risk of deterioration, to generate long-term transition probabilities for selezipag

Report section	4.2.6
Description of issue and why the EAG has identified it as important	Compounding the potential for bias raised in issue 2 (above), the company continue to apply these relative risks of ESC/ERS risk strata improvement/deterioration, to derive 12-week transition probabilities for selezipag over the entire time horizon of the model. This results in rapid risk status deterioration for the cohort treated with selezipag, which the EAG believe lacks clinical face validity (i.e. no one (<1%) remaining in the intermediate-low or low risks states by approximately ■ months compared to ■ in the sotatercept arm.
What alternative approach has the EAG suggested?	Like for issue 2, the EAG believe that applying the more conservative relative/risks of WHO FC improvement/deterioration, derived from the ITC with the GRIPHON trial, produces more plausible output. However, the repeated application of these full relative risks to 12-week transition probabilities still produces pessimistic projections of risk status deterioration on selezipag - based on clinical expert feedback received by the EAG. The EAG therefore propose a series of scenarios that reduce the magnitude of these effects when applied to long-term transition probabilities - from 100% to 25% of relative treatment effects.
What is the expected effect on the cost-effectiveness estimates?	Applying these smaller relative treatment effects, to derive long-term transition probabilities for selezipag, results in substantial upward shifts in the ICER.
What additional evidence or analyses might help to resolve this key issue?	The EAG acknowledge the challenges around generating evidence for the comparative efficacy of sotatercept versus selezipag on ESC/ERS risk improvement/deterioration. Expert opinion on the face validity of the modelled projections of progression to intermediate-high and high risks states may be beneficial.

Issue 4 The simplifying assumption of not allowing for clinical improvement following initiation of PGI2 analogues.

Report section	4.2.2 and 4.2.6
Description of issue and why the EAG has identified it as important	Compounding potential biases in the transition probabilities for patients on selexipag, which results in rapid progression in ESC/ERS risk status in the selexipag arm of the model, the company apply a simplifying structural assumption of not allowing any subsequent improvement in risk status following deterioration. This is despite all those who progress to the intermediate-high or high risks states on selexipag, discontinuing selexipag treatment and being escalated to treatment with PGI ₂ analogues (primarily IV prostacyclin). Expert feedback to the EAG suggests the simplifying assumption lacks clinical validity.
What alternative approach has the EAG suggested?	At the clarification stage, the EAG asked the company to explore scenarios allowing for a one-off chance of improvement in risk status following initiation of PGI ₂ analogues in the selexipag arm of the model. This was originally what was assumed for patients who commence PGI ₂ analogues in the high risk state following progression on sotatercept. However, the company noted the substantial structural changes and additional data required to account for this event history in their Markov model. Instead of allowing for the one-off chance of risk status improvement following initiation of PGI ₂ analogues in the selexipag arm, they switched off this possibility in the sotatercept arm. The EAG have not be able to modify the structure to allow for risk status improvement following deterioration within the timelines of the appraisal.
What is the expected effect on the cost-effectiveness estimates?	With a large proportion in the selexipag arm progressing rapidly to the intermediate-high and high risks states, where they receive costly treatment with PGI ₂ analogues, and where a substantial utility decrement is applied to those on IV PCA, this simplifying assumption may bias against selexipag. The impact on the ICER of sotatercept versus selexipag is, however, unpredictable, as allowing for improvement may disproportionately increase costs and QALYs in the selexipag arm relative to the sotatercept arm. To counter the potential biases EAG prefer at least not to apply the utility decrement applied to IV PCA in the intermediate-high and high risk states of the model.
What additional evidence or analyses might help to resolve this key issue?	The approach is less than ideal, and so further work to expand the model structure to allow for improvement and subsequent maintenance of risk status following PGI ₂ initiation in both arms may be beneficial.

Issue 5 Timing and approach to PGI2 initiation in the sotatercept arm of the model, is inconsistent with clinical expert feedback received by the EAG

Report section	4.2.2 and 4.2.6
Description of issue and why the EAG has identified it as important	Patients who progress from the intermediate-low to intermediate-high risk state on sotatercept, are assumed to remain on sotatercept and not require initiation of PGI2 analogues, despite progressing on treatment. Patients who commence PGI2 analogues in the high-risk state following progression on sotatercept, are assumed to commence PGI2 analogue treatment in place of sotatercept rather than in addition to it. Clinical expert feedback to the EAG suggests that for patients who progress on sotatercept from the intermediate-low to the intermediate high risk state, it is likely that a sizable proportion would be initiated on PGI2 analogues on top of sotatercept at this earlier stage, and that patients receiving PGI2 analogues following progression to the high-risk state would also remain on sotatercept.
What alternative approach has the EAG suggested?	The EAG suggest including PGI2 costs, in addition to sotatercept, for a proportion of patients who progress to the intermediate-high risk state. Whilst the STELLAR trial does not allow us to determine the efficacy of adding PGI2 analogues for those who progress on sotatercept from intermediate-low to intermediate-high risk, the data used to inform onward transitions includes patients who started sotatercept in the intermediate high-risk state on top of triple background therapy (including PGI2 analogues). The EAG approach assumes that adding PGI2 analogues for patients who progress to intermediate-high risk on ERA+PDE5i+sotatercept, has similar efficacy to adding sotatercept for those with intermediate-high risk status on triple background therapy (including PGI2 analogues).
What is the expected effect on the cost-effectiveness estimates?	Applying a PGI2 analogue costs to a proportion of patients on sotatercept in the intermediate-high risk state results in a considerable upward shift in the ICER for sotatercept. Scenarios that retain people on sotatercept in addition to PGI2 analogues in the high risk state, also substantially increase the ICER, but the EAG are less confident in these scenarios as there is no way to account for the increased efficacy that this maximal treatment regimen may have compared to the less intensive regimen received by patients in the high risk state in the selixipag arm.
What additional evidence or analyses might help to resolve this key issue?	Further clinical expert opinion, or real-world evidence, to inform the likely approach to treatment escalation for patients who progress from the intermediate-low to intermediate-high or high risk states on sotatercept would be beneficial.

Issue 6 The proportion of patients who receive PGI2 analogues, when indicated according to the modelling assumptions, is too high at 100%

Report section	4.2.2 and 4.2.8
Description of issue and why the EAG has identified it as important	When infused PGI2 analogues are initiated in the intermediate-high and high risks states following progression, the company base case assumes that all patients (100%) receive these treatments. Clinical expert feedback to the EAG suggests that these treatments may not be suitable for all patients and of those offered, not all accept it.
What alternative approach has the EAG suggested?	The EAG assume that 85% of patients will receive infused PGI2 analogues upon progression to the intermediate-high risk state in the selexipag arm and the high-risk state in the sotatercept arm of the model. The remainder are assumed to remain on their current treatment regimen. Whilst this lower proportion may not be in keeping with the efficacy data used to inform transition probabilities for PGI2 analogues, they further counter the potential bias caused by limiting the efficacy of these treatments in the model by not allowing for improvement in risk status following their initiation.
What is the expected effect on the cost-effectiveness estimates?	Reducing the proportion who receive PGI2 analogues upon progression to higher risk states lowers costs in both arms, but more so in the selexipag arm, resulting in a moderate increase in the ICER for sotatercept compared to the company base case. However, the reverse holds when the EAGs preferred costing assumptions are applied to IV PCA preparations.
What additional evidence or analyses might help to resolve this key issue?	Clinical expert opinion or real-world evidence to support the proportions receiving PGI2 analogues following progression to intermediate-high and high risks states.

Issue 7 The weight-based doses applied to the IV PCA preparations included in the model are high compared to European guidelines and clinical expert feedback to the EAG.

Report section	4.2.8
Description of issue and why the EAG has identified it as important	The IV PCA preparations (epoprostenol and Treprostinil) applied in the economic model are administered as weight-based doses by continuous IV infusion (ng/kg/min). The company use target doses for these derived from a prior appraisal of selexipag considered by CADTH, setting the target dose of epoprostenol at 50ng/kg/min and at 30ng/kg/min for Treprostinil. Clinical expert opinion to the EAG suggests these doses differ substantially (particularly for epoprostenol) from what is used routinely in UK NHS practice. This suggests over costing of epoprostenol, the most frequently used preparation in the UK NHS, and potential bias in favour of sotatercept.
What alternative approach has the EAG suggested?	The EAG prefers to apply dosing for PGI2 analogues as per the ESC/ERS guidelines, which aligns very closely with the clinical expert feedback it received; Epoprostenol: 23ng/kg/min (midpoint of 16-30ng/kg/min); Treprostinil: 42.5ng/kg/min (midpoint of 25-60 ng/kg/min).
What is the expected effect on the cost-effectiveness estimates?	Changing the doses of PGI2 analogues lowers costs more in the selexipag arm than in the sotatercept arm, resulting in a substantial increase in the ICER for sotatercept
What additional evidence or analyses might help to resolve this key issue?	Real world evidence for average doses of IV PCA applied in the UK NHS, or further validation of the doses applied in the model based on clinical expert opinion.

1.6 Summary of EAG's preferred assumptions and resulting ICER

Table 2 below outlines the impact on the company's base case of each of EAG's changes or preferred modelling assumptions. The final rows in the table incorporate all changes.

However, recognising the uncertainty in the long-term comparative efficacy on sotatercept versus selezipag on ESC/ERS risk strata change, the EAG have provided a two alternative base case in addition to its central base case:

1. A lower ICER alternative, whereby the full relative risk reduction for deterioration derived from the ITC between STELLAR and GRIPHON for WHO FC, is applied as the proxy for the long-term effect of sotatercept versus selezipag on ESC /ERS risk status deterioration
2. A higher ICER alternative, whereby the 25% of the relative risk reduction for deterioration derived from the ITC between STELLAR and GRIPHON for WHO FC, is applied as the proxy for the long-term effect of sotatercept versus selezipag on ESC /ERS risk status deterioration

Table 2 Summary of the EAG preferred modelling assumptions

EAG Preferred assumption	Incremental costs (£)	Incremental QALYs	Cumulative ICER (£/QALY)	Cumulative ICER with severity modifier (£/QALY)
Company base case	██████	2.53	██████	██████
Hospitalisation QALY loss calculations adjusted to the 12 week cycle length	██████	2.48	██████	██████
Selezipag short term (24 week) transition probabilities informed by application of WHO FC relative risks from the ITC with GRIPHON	██████	2.43	██████	██████
Selezipag long-term transition probabilities derived by applying half of the █████ relative risk reduction of disease progression for sotatercept versus selezipag at 24 weeks, based on the ITC with GRIPHON (= █████ relative risk reduction)	██████	1.89	██████	██████

Remove independent effect of sotatercept versus selexipag on hospitalisation in the low and intermediate-low risk state	██████	2.49	██████	██████
Remove IV PCA administration disutility (to compensate for the lack of clinical benefit included in the model for these treatments)	██████	1.59	██████	██████
Reduce utility decrement associated with hospitalisation from 0.105 to 0.071, applied for the duration of the cycle in which events occur	██████	2.51	██████	██████
Assume infused PCA therapies added to sotatercept for 40% upon progression to the intermediate-high risk state on sotatercept	██████	2.33	██████	██████
Assume 85% rather than 100% of patients receive PGI2 analogues following progression on sotatercept or selexipag to high risk and intermediate high risks state respectively.	██████	2.39	██████	██████
Dosing assumptions for epoprostenol and Treprostinil based on target maintenance doses in ESC / ERS guidelines: 23ng/Kg/min and 42.5ng/Kg/min, respectively	██████	2.53	██████	██████
Assume respiratory consultant led OP appointments to initiate treatment on sotatercept and selexipag but retain company assumptions about number of contacts	██████	2.53	██████	██████
EAG base case (central)	██████	0.87	██████	██████
EAG base case (low)	██████	1.24	██████	██████
EAG base case (high)	██████	0.68	██████	██████

ICER, incremental cost-effectiveness ratio; EAG, External Assessment Group; WHO FC, World Health Organisation Functional Class; ITC, indirect treatment comparison; IV PCA, intravenous prostacyclin analogues; ESC, European Society of Cardiology; ERS, European Respiratory Society.

Modelling errors or inconsistencies identified and amended by the EAG are described in section 5.3 of this report ([Section 5.3](#)). For further details of the exploratory and sensitivity analyses done by the EAG, see chapter 6 ([EAGs additional analysis](#)).

2 INTRODUCTION AND BACKGROUND

2.1 Introduction

The relevant health condition for the submission received from Merck Sharp & Dohme (UK) Limited is pulmonary arterial hypertension (PAH). The company's description of the health condition in terms of prevalence, symptoms and complications appears accurate and in line with the decision problem. The relevant intervention for this submission is sotatercept (brand name: Winrevair).

2.2 Background

The company submission (CS) describes PAH as a chronic, highly progressive orphan disease that is characterised by high blood pressure (hypertension) in the blood vessels in the lungs.¹ This reduces blood flow and causes the right ventricle of the heart to work harder, leading to a weakening of the heart muscle, which can eventually progress to right heart failure and, ultimately, death.^{1, 2}

PAH is a rare condition occurring in 15 to 50 persons per million within the United States and Europe and affects people of all ages, but it is more common in women aged between 30 to 60 years.^{3, 4} It is primarily caused by imbalance of anti-proliferative and pro-proliferative signals that leads to pulmonary vascular hyperproliferation.⁵ Symptoms include shortness of breath (dyspnoea), fatigue, chest pain, dizziness, and fainting (syncope). PAH is associated with a significant, debilitating effect on the quality of life of patients and their carers and negatively impacts psychological health, employment and work productivity, and personal finances.⁶⁻⁹

There is no current cure for PAH and prognosis is poor; median survival is 6-7 years after initial diagnosis.^{4, 10, 11} The company outlines the treatment pathway for PAH in section 1.3.2 of the CS. The company states that the availability of treatment options are defined by NHS commissioning policies, however, feedback from UK clinical experts and NHS England specifies that the current treatment pathway reflects the recommendations of the European Society of Cardiology (ESC) and European Respiratory Society (ERS) 2022 guidelines and is based on a patient's risk status assessed using three categories (low, intermediate and high) at diagnosis and four categories from first assessment on pulmonary hypertension therapy (low,

intermediate-low, intermediate-high, or high risk). Treatment goals aim to achieve and maintain low risk status. The company outlines the scoring system for determining risk status in Table 4 of the CS and this is reproduced by the EAG as Table 3. *The EAG clinical experts agree with the company's position that UK clinical practice has moved away from the WHO functional class (FC) status classification and that current clinical prescribing is based on ESC/ERS risk status.*

Table 3 2022 ESC/ERS simplified four-strata risk assessment tool for guidance of therapeutic decision-making

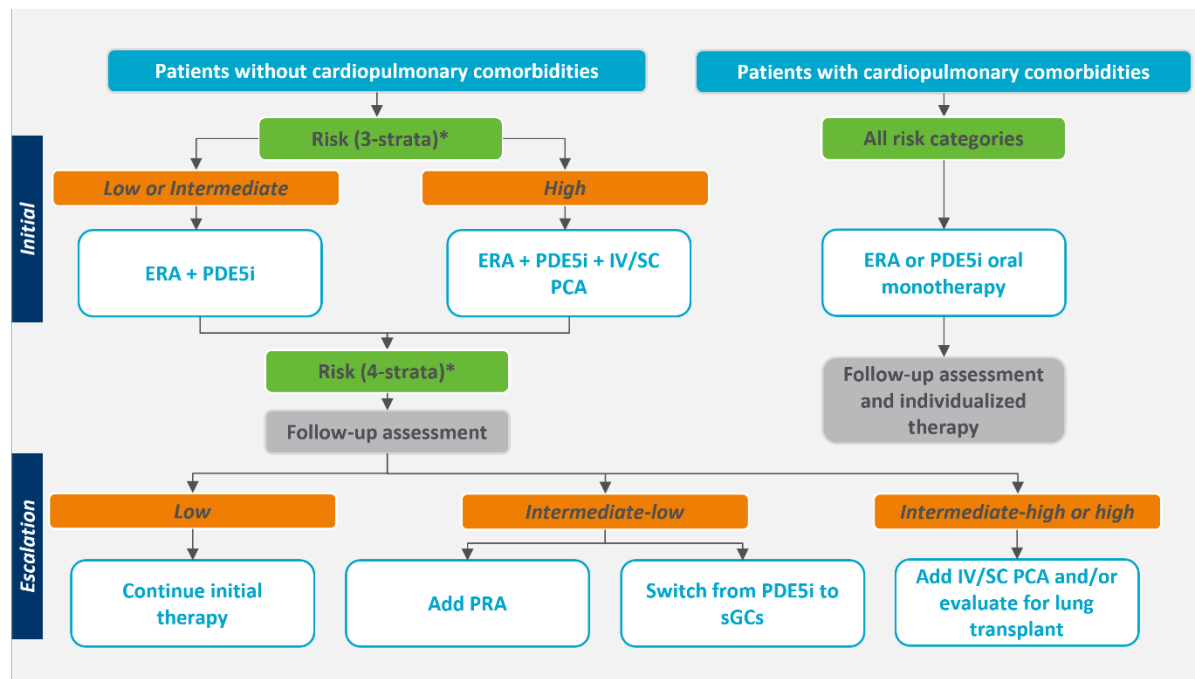
Determinants of prognosis	Low risk	Intermediate-low risk	Intermediate-high risk	High risk
Points assigned	1	2	3	4
WHO FC	I or II ^a	-	III	IV
6MWD, m	>440	320–440	165–319	<165
BNP or NT-proBNP, ng/L	<50 <300	50–199 300–649	200–800 650–1100	>800 >1100
Risk is calculated by dividing the sum of all grades by the number of variables and rounding to the next integer.				

6MWD: 6-minute walking distance; BNP: brain natriuretic peptide; ESC: European Society of Cardiology; ERS: European Respiratory Society; FC: functional class; NT-proBNP: N-terminal pro-brain natriuretic peptide; WHO: World Health Organization

^a WHO FC I and II are assigned 1 point as both are associated with good long-term survival.

The therapeutic options according to risk status are presented by the company in Figures 1 and 2 of the CS and these are reproduced by the EAG below.

Figure 1 Pulmonary arterial hypertension treatment algorithm based on the European society of cardiology/European respiratory society guidelines

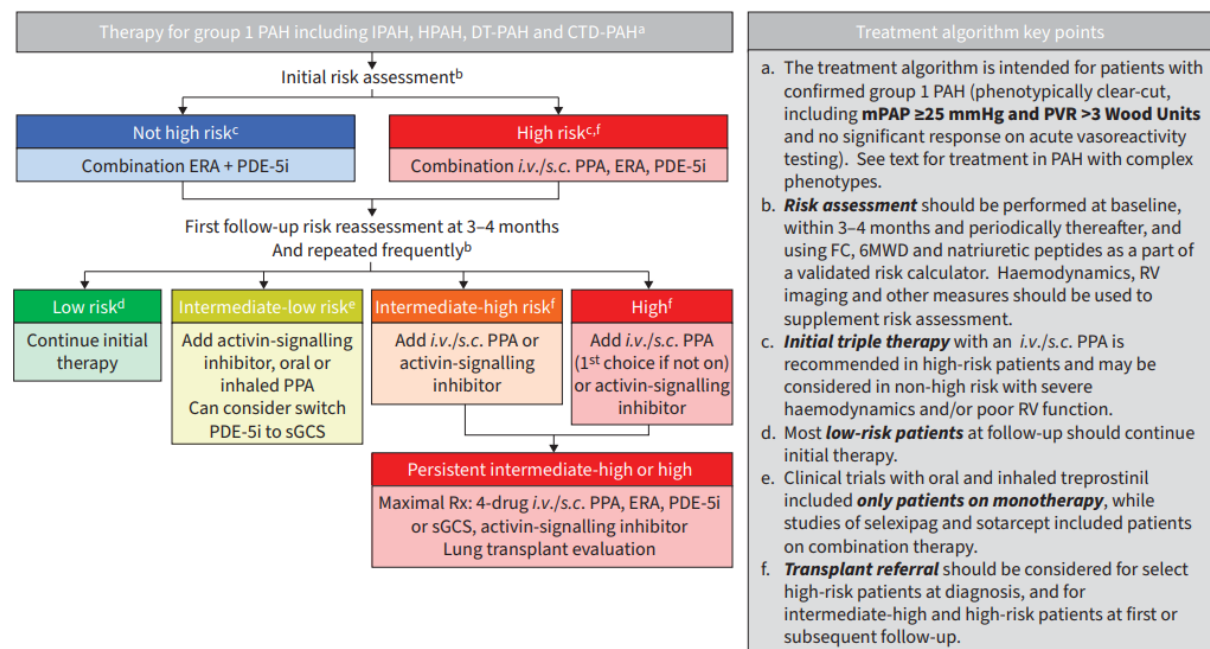


ERA: endothelin receptor antagonist; IV: intravenous; PCA: prostacyclin analogue; PDE5i: phosphodiesterase type-5 inhibitor; PRA: prostacyclin receptor antagonist; SC: subcutaneous; sGCs: soluble guanylate cyclase stimulator

*Risk assessment using 3-strata model to determine risk stratification at diagnosis; four-strata model based on determinants of prognosis.

Source: Humbert et al., 2022¹²

Figure 2 Pulmonary arterial hypertension outline of overall treatment recommendations/treatment algorithm by the World Symposium on Pulmonary Hypertension



6MWD: 6-min walk distance; CTD: connective tissue disease; DT: drug and toxin; ERA: endothelin-1 receptor antagonist; FC: functional class; HPAH: hereditary PAH; IPAH: idiopathic PAH; *i.v.*: intravenous; mPAP: mean pulmonary artery pressure; PAH: pulmonary arterial hypertension; PDE-5i: phosphodiesterase-5 inhibitor; PPA: prostacyclin pathway agent; PVR: pulmonary vascular resistance; RV: right ventricle; Rx: prescription; *s.c.*: subcutaneous; sGCS: soluble guanylyl cyclase stimulator

Source: Chin et al., 2024¹³

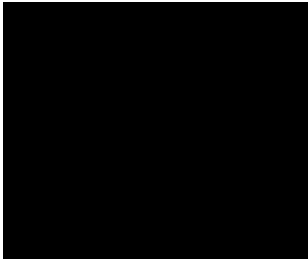
The company anticipates that sotatercept will be indicated for patients who are intermediate-low risk according to the ESC/ESR 4-risk strata as an alternative to the current standard of care comprising phosphodiesterase-5 inhibitor (PDE5i) + endothelin-1 receptor antagonist (ERA) + selexipag. Treatment with sotatercept will be initiated and managed in specialist centres and administered at home. The company outlines their rationale for the positioning of sotatercept in the treatment pathway in Table 5 of the CS. *The EAG clinical experts agree with the company's description of the current treatment landscape for PAH in the UK and the company's positioning of sotatercept. They note that intermediate-high risk patients could benefit from having an alternative treatment option to intravenous therapies that are initiated in the hospital setting followed by administration at home but accept that the available clinical effectiveness evidence presented in the CS does not provide comparative data for sotatercept in this patient population and that some patients are not suitable for IV therapy.*

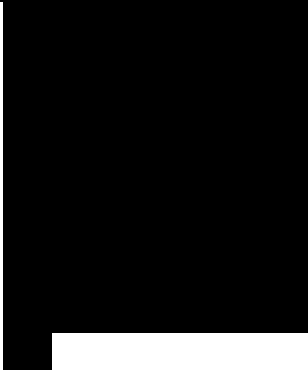
2.3 Critique of company's definition of decision problem

A summary of the company's decision problem in relation to the NICE final scope is presented in Table 4 below. A critique of the adherence of the company's economic modelling to the NICE reference case is presented in Chapter 4.

Table 4 Summary of the company's decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	EAG comment
Population	Adults with pulmonary arterial hypertension	Adults with PAH, (WHO Group 1), who are at intermediate-low risk according to the ESC/ERS four-strata risk rating system.	Given the available evidence and advice from clinicians, this submission considers the use of sotatercept in patients with PAH who have an intermediate-low risk status: • Patients with low risk status as per ESC/ERS guidelines would not require the addition of sotatercept as they would already be in the lowest risk category, which is the treatment goal in PAH management. • Patients who progress to intermediate-high risk status are currently treated	The company's target population is more specific than that of the final scope and aligns with the company's proposed positioning of sotatercept in the care pathway. The EAG clinical experts agree that intermediate-low risk PAH patients are most likely to receive sotatercept in the UK, although they note that there may be some overlap with intermediate-high risk patients who could also benefit from treatment with sotatercept.

			with PCA in combination with PDE5i + ERA. The sotatercept-based regimen in these patients is sotatercept in combination with PDE5i + ERA instead of PCA. Robust comparative evidence between these (from either direct trial evidence or indirect comparison) is not available, so this population has been excluded from the decision problem addressed in the company submission. 	

				
Intervention	Sotatercept in combination with established clinical management	Sotatercept in combination with established clinical management	N/A	The intervention described in the CS matches that described in the NICE final scope. Sotatercept is a first-in-class fusion protein consisting of the Fc domain of human immunoglobulin G (IgG) linked to the extracellular domain of human activin receptor type IIA (ActRIIA), which acts as a ligand trap for selected transforming growth factor beta (TGF-β) superfamily members. Inhibition of these ligands by sotatercept is proposed to help rebalance pulmonary vascular

				homeostasis toward growth-inhibiting and proapoptotic signalling. Sotatercept is administered subcutaneously at a starting dose of 0.3 mg/kg with a target dose of 0.7 mg/kg administered every 21 days. This submission is based on a subset of the company's full marketing authorisation which is: <i>"Winrevair, in combination with other PAH therapies, is indicated for the treatment of PAH in adult patients with WHO FC II to III, to improve exercise capacity."</i> The marketing authorisation for sotatercept in the indication detailed in this submission was approved by the MHRA on 27th December 2024.
Comparator(s)	Established clinical management without sotatercept (given in	Triple combination therapy of PDE5i + ERA + selexipag	The relevant comparators should reflect current clinical practice.	The EAG clinical experts agree that PDE5i + ERA + selexipag is a relevant comparator for this

	combination as dual or triple therapy): • endothelin receptor antagonists: ambrisentan, bosentan and macitentan • phosphodiesterase 5 inhibitors: sildenafil and tadalafil • prostaglandins/prostanoids: epoprostenol, selexipag, intravenous/inhaled iloprost and Treprostinil • soluble guanylate cyclase stimulators: riociguat.		Current reimbursed dual- and triple therapy regimens are specified in the relevant NHS England commissioning policies below (further details provided in Appendix J): • Targeted Therapies for use in Pulmonary Hypertension in Adults, July 2015, A11/P/c. • Riociguat for pulmonary arterial hypertension, February 2017, 16055/P. • Selexipag for treating pulmonary arterial hypertension (adults), December 2018, 170104P. However: • Riociguat-based regimens (usually in combination with ERA) are not a relevant	appraisal, however, limiting the appraisal to this combination of therapy excludes those intermediate-low risk patients who have either not responded to selexipag or who have responded to IV prostacyclin therapy.

			comparator for sotatercept, based on clinical advice that the target patient groups for these treatments do not overlap. Riociguat is currently administered for patients who require further pharmacological targeting of the NO pathway. This position exists for a limited subset of patients prior to the initiation of a third pharmacological therapy. The most recent National PAH Audit reports fewer than 100 PAH patients (i.e., not including CTEPH patients) currently treated with riociguat¹⁴. • Monotherapy/dual therapy with PDE5i +/- ERA are	

			used to manage patients with ESC/ERS low risk status. Patients with low risk status would not require the addition of sotatercept (or any other intervention) as they have already achieved the lowest risk category, which is the treatment goal of PAH management and so these would not be relevant comparators to sotatercept. Both treatment guidelines and clinical opinion indicate this. • Triple combination therapy including PCA are used in the treatment of patients with ESC/ERS intermediate-high or high risk status	

			i.e., patients who may have failed to respond to, or had subsequent progression with selexipag. Therefore, triple combination therapy of PDE5i + ERA + selexipag, which is used to treat patients with ESC/ERS intermediate-low risk status, is the relevant comparator for the proposed positioning.	
Outcomes	The outcome measures to be considered include: • time to clinical worsening • hospitalisations • overall survival • transplant-free survival • exercise capacity (for example 6-minute walking distance) • haemodynamic assessment (e.g., cardiac index, cardiac output, right atrial pressure,	• 6MWD • PVR • NT-proBNP • MCI • WHO FC • Simplified French Risk Score • Time to death or the first occurrence of a clinical worsening event • HRQoL	As provided by pivotal clinical trial, STELLAR	The EAG clinical experts believe that the outcomes considered in the CS are comprehensive for this appraisal.

	pulmonary arterial pressure and pulmonary vascular resistance) • 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. 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. Costs will be considered from an NHS and Personal Social Services perspective.	Cost-effectiveness analysis	N/A	The company case is generally well aligned with the NICE reference case, but with some deviations related to health related quality of life inputs.
Subgroups	If the evidence allows, the following subgroups will be considered: • WHO functional class 2 or 3	• ESC/ERS four-strata risk status • WHO functional class 2 or 3	Subgroups relevant to the current treatment pathway in PAH in UK clinical practice.	The EAG are satisfied that the considered subgroups are comprehensive and relevant for this appraisal.

	• ESC/ERS four-strata risk status • background PAH therapy (dual or triple therapy) • clinical subgroups within group 1 PAH ○ idiopathic ○ heritable ○ associated with drugs/toxins ○ associated with connective tissue disease ○ associated with congenital heart disease.	• background PAH therapy (dual or triple therapy)		
Special considerations including issues related to equity or equality	N/A	N/A	N/A	The EAG agrees with the company that no equity or equality considerations are anticipated.

6MWD: 6-minute walk distance; CTEPH: chronic thromboembolic pulmonary hypertension; ERA: endothelin receptor antagonist; ERS: European Respiratory Society; ESC: European Society of Cardiology; FC: functional class; HRQoL: health-related quality-of-life; MCI: multicomponent improvement; N/A: not applicable; NHS: National Health Service; NO: nitrous oxide; NT-proBNP: N-terminal pro-brain natriuretic peptide; PAH: pulmonary arterial hypertension; PCA: prostacyclin analogue; PDE5i: phosphodiesterase type 5 inhibitor; PVR: pulmonary vascular resistance; UK: United Kingdom; WHO: World Health Organization

3 CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

Full details of the methods used by the company to identify and select the clinical evidence relevant to this appraisal are reported in Appendix B of the CS. The EAG's appraisal of the company's systematic literature review (SLR) methodology is summarised in Table 5.

Table 5 EAG's appraisal of the literature review methods presented in the CS

Review process ERG	ERG response	Comments
Were appropriate searches (e.g., search terms, search dates) performed to identify all relevant clinical and safety studies?	YES	The CS provides full details of the searches used to identify the studies for the clinical effectiveness review. The search strategies include relevant controlled vocabulary and text terms with appropriate use of Boolean operators and are fully reproducible. Details are provided in Appendix B of the CS. Searches were not restricted by any eligibility criteria, so all results were discovered and only those relevant to the scope were selected.
Were appropriate bibliographic databases/sources searched?	YES	Sources included Embase, Medline, and CENTRAL for primary research. Relevant conference proceedings and trial registers were also searched. Bibliographies of recent SLRs were examined to identify relevant studies not captured by the literature searches. Full details are provided in Appendix B of the CS.
Were eligibility criteria consistent with the decision problem outlined in the NICE final scope?	YES	The eligibility criteria outlined in Appendix B, section B1.1.1 are consistent with the decision problem outlined in the NICE final scope.

Was study selection conducted by two or more reviewers independently?	YES	Appendix B, section B1.1.1 states both title and abstract and full text selection was conducted <i>“by two independent researchers to determine eligibility according to the inclusion and exclusion criteria.”</i>
Was data extraction conducted by two or more reviewers independently?	PARTIALLY	Appendix B, section B1.1.1 states <i>“The data extraction was quality checked by a second researcher not involved with the extraction; this quality check included verification of all extracted data against the source document.”</i>
Were appropriate criteria used to assess the risk of bias of identified studies?	PARTIALLY	Randomised controlled trials (RCTs) were appropriately assessed using the Cochrane Risk of Bias (RoB) Assessment Tool 2.0. The EAG notes that the company did not include open-label extension (OLE) studies in their critical appraisal where the relevant variables for the assessment were consistent between the primary and extension trial. The company also appears to have used the Cochrane RoB tool to assess the risk of bias of the OLE PATENT-2, SUPER-2 and SOTERIA studies. Participants were not randomised in the OLE studies, therefore a tool for assessing risk of bias in non-randomised studies would have provided more appropriate risk of bias criteria.
Was the risk of bias assessment conducted by two or more reviewers independently?	YES	The risk of bias assessment was conducted by two independent reviewers, with any disagreements resolved by a third senior reviewer.
Was identified evidence synthesised using appropriate methods?	PARTIALLY	The EAG have concerns with the analysis of change in risk status from baseline but are satisfied with the analysis for all other outcomes.

The EAG conducted a quality assessment of the methods used by the company for the SLR of clinical evidence based on the Centre for Reviews and Dissemination (CRD) criteria. The results are presented in Table 6.

Table 6 Quality assessment of the company's systematic literature review of clinical effectiveness evidence

CRD quality item	Yes/No/Unclear
1. Are any inclusion/exclusion criteria reported relating to the primary studies, which address the review question?	Yes
2. Is there evidence of a substantial effort to search for all of the relevant research?	Yes
3. Is the validity of included studies adequately assessed?	Yes
4. Are sufficient details of the individual studies presented?	Yes
5. Are the primary studies summarised appropriately?	Yes

3.2 Critique of trials of the technology of interest, the company's analysis and interpretation (and any standard meta-analyses of these)

Details of the key clinical effectiveness evidence are presented in Section 2 of the CS. The company's main evidence for sotatercept is obtained from the STELLAR (NCT04576988) and SOTERIA (NCT04796337) studies. STELLAR is a multicentre, double-blind, phase 3 RCT which randomly assigned adults with WHO functional class II or III to receive sotatercept (n=163) or placebo (n=160) in addition to their background therapy. Sotatercept or placebo was administered as a subcutaneous (SC) injection every 21 days. The starting dose of sotatercept was 0.3 mg/kg at visit 1 and escalated to the target dose of 0.7 mg/kg at visit 2 and kept at this dose throughout the treatment period (unless haemoglobin or platelet changes warranted dose decrease).¹⁵ The company presents the study findings from STELLAR from the most recent 06-DEC-2022 data cutoff date.

SOTERIA (n=426) is a long-term, OLE study of adults with PAH who have completed prior sotatercept trials (SPECTRA, PULSAR, ZENITH, HYPERION, STELLAR). Participants were permitted to roll over into SOTERIA if they experienced a clinical worsening event or completed the 24-week treatment period in

the STELLAR trial. The company provides a summary of the studies from which participants may enter SOTERIA in Table 10 of the CS. The patient population of SOTERIA is broader than that of STELLAR, however, the company provides data for those participants in the SOTERIA study who were originally in the sotatercept treatment arm of the STELLAR study, *therefore, the EAG have no concerns about the inclusion of data from SOTERIA for this appraisal.* Results from SOTERIA are presented by the company from the 08-NOV-2023 cumulative data cutoff date.

The company's SLR identified one further eligible phase 2 RCT, PULSAR (NCT03496207); however, the company states in Section 2.9 of the CS that evidence from STELLAR and SOTERIA supersedes PULSAR and that data from STELLAR and PULSAR are not compatible for inclusion in a meta-analysis due to differences in the dosing of sotatercept, the trial end points, and the statistical methods used in the two trials. *The EAG accepts the company's argument that STELLAR and PULSAR are not suitable for inclusion in a meta-analysis.*

Summaries of the methodologies of STELLAR and SOTERIA are presented in Table 7, Section 2.3 of the CS and are reproduced by the EAG as Table 7x. *It is the EAG's opinion that STELLAR and SOTERIA are of good methodological quality.*

Table 7 Summaries of the STELLAR and SOTERIA methodologies

Study name	STELLAR	SOTERIA
Trial design	Phase 3, multicentre, double-blind, randomised, placebo-controlled clinical trial	Phase 3, open-label, long-term follow-up, interventional study
Eligibility criteria for participants	Eligible patients had a confirmed diagnosis of pulmonary arterial hypertension (idiopathic, heritable, drug-induced, connective tissue disease–associated, or after shunt correction) in WHO functional class II or III, excluding subtypes associated with portopulmonary disease, schistosomiasis, human immunodeficiency virus	Participants eligible to enrol in the study will have participated in and completed the relevant study requirements of the parent PAH sotatercept clinical studies.

Study name	STELLAR	SOTERIA
	infection, or veno-occlusive disease. All the patients were receiving stable background therapy for pulmonary arterial hypertension for at least 90 days before enrolment and continued receiving background therapy throughout the trial.	
Settings and locations where the data were collected	Multinational trial with locations in: Argentina, Australia, Belgium, Brazil, Canada, Czechia, France, Germany, Israel, Italy, Korea, Republic of, Mexico, Netherlands, New Zealand, Poland, Serbia, Spain, Sweden, Switzerland, United Kingdom, United States	Multinational trial with locations in: Argentina, Australia, Austria, Belgium, Brazil, Canada, Colombia, Croatia, Czechia, Denmark, France, Germany, Greece, Israel, Italy, Korea, Republic of, Mexico, Netherlands, New Zealand, Poland, Portugal, Serbia, Spain, Sweden, Switzerland, Taiwan, United Kingdom, United States
Trial drugs	• Sotatercept at a starting dose of 0.3 mg/kg with a target dose of 0.7 mg/kg administered SC every 21 days plus background PAH therapy • Placebo administered SC every 21 days plus background PAH therapy	• Sotatercept SC injection at a dose of 0.3 to 0.7 mg/kg.
Primary outcomes	• 6MWD	• Long-term safety and tolerability
Other outcomes used in the economic model/specified in the scope	• MCI • PVR • NT-proBNP • WHO FC	• 6MWD • PVR • NT-proBNP • WHO FC
Pre-planned subgroups	• Sex (male and female) • PAH etiological subgroups (iPAH, hPAH, drug/toxin-induced PAH, connective tissue disease, congenital heart disease with s/p shunt repair)	None

Study name	STELLAR	SOTERIA
	• Monotherapy vs. double vs. triple combination therapy at baseline • Prostacyclin infusion therapy vs. non-prostacyclin infusion at baseline • Baseline WHO functional class (II or III) • Baseline PVR (≤ 800 or > 800 dynes*sec/cm⁵)	

6MWD: 6-minute walking distance; FC: functional class; hPAH: heritable PAH; iPAH: idiopathic PAH; MCI: multicomponent improvement; NT-proBNP: N-terminal pro-brain natriuretic peptide; PAH: pulmonary arterial hypertension; PVR: pulmonary vascular resistance; SC: subcutaneously; s/p: systemic to pulmonary; WHO: World Health Organization.

Source: Hooper et al 2023;¹⁵ Preston et al 2025¹⁶

The baseline characteristics of STELLAR participants are broadly similar across the study treatment groups. *The EAG clinical advisors are of the opinion that the participants enrolled in the STELLAR and SOTERIA studies are broadly representative of patients with PAH seen in NHS clinical practice. However, they noted that the study participants have a longer disease duration, better cardiac index, and more stable disease than patients who would be eligible for sotatercept treatment in clinical practice.*

3.2.2 Primary and secondary efficacy endpoints

The clinical effectiveness data for STELLAR and SOTERIA are provided in Sections 2.61 and 2.6.2 of the CS respectively.

Primary efficacy endpoints

The primary endpoint of STELLAR was the change from baseline at week 24 in the 6-minute walk distance (6MWD). At Week 24, there was a significantly greater proportion of patients with an improvement in 6MWD of at least 30 metres in the patients in the sotatercept arms than in the placebo arm (54% compared to 22%). The median treatment difference (Hodges-Lehmann location shift) between the sotatercept and placebo groups was 40.4 metres (asymptomatic standard error, 6.70; 95% CI 27.28, 53.53; $p < 0.001$), and the observed mean change from baseline was higher in

the sotatercept group (40.3 metres) compared with the placebo group (-0.6 metres), a trend that continued up to Week 84.

Secondary efficacy endpoints

The secondary endpoints considered in STELLAR are presented in Section 2.6.1 of the CS and include: multicomponent improvement; pulmonary vascular resistance; N-terminal pro-brain natriuretic peptide (NT-proBNP); WHO FC improvement, time to death or first occurrence of a clinical worsening event; French Risk score; change from baseline to Week 24 according to the ESC/ERS (COMPERA 2.0) four-strata risk scoring system; change from baseline in the physical impacts, cardiopulmonary symptoms, and cognitive/emotional impacts domain scores of the PAH-SYMPACT (pulmonary arterial hypertension-symptoms and impact questionnaire); EQ-5D-3L and EQ-5D-5L VAS. All secondary outcomes favoured treatment with sotatercept, although the difference between the cognitive/emotional impacts domain score at week 24 was not statistically significant between treatment groups and there was no notable difference in the change from baseline in EQ-5D-3L index score between the sotatercept and placebo groups at Week 24.

3.2.3 Subgroup analyses

The company reports the results of their subgroup analyses in Table 59, Section 2.8 and Appendix C of the CS. The following subgroups were considered as part of the STELLAR trial: sex; PAH diagnostic group; background PAH therapy at baseline; prostacyclin infusion therapy at baseline; WHO functional class at baseline; PVR at baseline; and risk status at baseline. The company also presents subgroup results for STELLAR participants who had ESC/ERS intermediate-low risk at baseline in Table 60 of the CS. This was not a pre-defined subgroup analysis of STELLAR.

Results of the subgroup analyses favoured treatment with sotatercept.

3.2.4 Adverse events

The company reports adverse events (AE) data for STELLAR and SOTERIA in section 2.11 of the CS.

STELLAR

Severe treatment-emergent adverse events (TEAEs), serious TEAEs, serious TEAEs leading to treatment discontinuation, serious TEAEs leading to death, TEAEs leading to treatment discontinuation and TEAEs leading to death occurred less commonly in the sotatercept group than in the placebo group. The incidences of specific TEAEs were generally similar in both groups. Frequently reported TEAEs (incidence $\geq 5\%$) with a higher incidence in the sotatercept group than in the placebo group included: epistaxis, telangiectasia, dizziness, nasal congestion, thrombocytopenia, and increased haemoglobin. Less frequently reported TEAEs (incidence $< 5\%$) with a higher incidence in the sotatercept group than in the placebo group were abdominal pain upper, chills, gingival bleeding, injection site erythema, bursitis, polycythemia, cataract and blurred vision. The EAG presents a summary of key adverse events reported in STELLAR in Table 8.

Table 8 Summary of key adverse events reported in STELLAR

Number of participants with any	Sotatercept (N=163) n (%)	Placebo (N=160) n (%)	All patients (N=323) n (%)	Sotatercept vs placebo comparison	
				Point estimate (%) ^a	95% CI ^a
TEAEs	151 (92.6)	149 (93.1)	300 (92.9)	-0.5	(-6.4, 5.4)
TEAEs related to treatment	83 (50.9)	45 (28.1)	128 (39.6)	22.8	(12.2, 32.9)
TEAE of interest	97 (59.5)	71 (44.4)	168 (52.0)	15.1	(4.2, 25.7)
TEAE of special interest	27 (16.6)	7 (4.4)	34 (10.5)	12.2	(5.8, 19.2)
TEAEs leading to treatment discontinuation	6 (3.7)	11 (6.9)	17 (5.3)	-3.2	(-8.7, 1.8)
TEAEs leading to study discontinuation	6 (3.7)	9 (5.6)	15 (4.6)	-1.9	(-7.1, 2.9)
TEAEs leading to death	2 (1.2)	7 (4.4)	9 (2.8)	-3.1	(-7.7, 0.5)
Severe TEAEs	24 (14.7)	33 (20.6)	57 (17.6)	-5.9	(-14.3, 2.5)
Serious TEAEs	40 (24.5)	47 (29.4)	87 (26.9)	-4.8	(-14.5, 4.9)
Serious TEAEs related to treatment	3 (1.8)	2 (1.3)	5 (1.5)	0.6	NR
Serious TEAEs leading to treatment discontinuation	4 (2.5)	9 (5.6)	13 (4.0)	-3.2	(-8.2, 1.3)
Serious TEAEs leading to study discontinuation	4 (2.5)	8 (5.0)	12 (3.7)	-2.5	(-7.4, 1.8)
Serious TEAEs leading to death	2 (1.2)	7 (4.4)	9 (2.8)	-3.1	(-7.7, 0.5)
Frequently reported (incidence ≥5%) with higher incidence in the sotatercept group					
Epistaxis	36 (22.09)	3 (1.88)	39 (12.07)	20.21	(13.88, 27.41)
Telangiectasia	27 (16.56)	7 (4.38)	34 (10.53)	12.19	(5.78, 19.18)
Dizziness	24 (14.72)	10 (6.25)	34 (10.53)	8.47	(1.85, 15.46)

Number of participants with any	Sotatercept (N=163) n (%)	Placebo (N=160) n (%)	All patients (N=323) n (%)	Sotatercept vs placebo comparison	
				Point estimate (%) ^a	95% CI ^a
Nasal congestion	10 (6.13)	0 (0.00)	10 (3.10)	6.13	(3.36, 10.93)
Thrombocytopenia	16 (9.82)	3 (1.88)	19 (5.88)	7.94	(3.12, 13.72)
Increased haemoglobin	10 (6.13)	0 (0.00)	10 (3.10)	6.13	(3.36, 10.93)
Less frequently reported TEAEs (incidence <5%) with a higher incidence in the sotatercept group					
Upper abdominal pain	7 (4.29)	1 (0.63)	8 (2.48)	3.67	(0.36, 8.06)
Chills	5 (3.07)	0 (0.00)	5 (1.55)	3.07	(0.68, 6.99)
Gingival bleeding	7 (4.29)	1 (0.63)	8 (2.48)	3.67	(0.36, 8.06)
Injection site erythema	5 (3.07)	0 (0.00)	5 (1.55)	3.07	(0.68, 6.99)
Bursitis	5 (3.07)	0 (0.00)	5 (1.55)	3.07	(0.68, 6.99)
Polycythaemia	4 (2.45)	0 (0.00)	4 (1.24)	2.45	(0.07, 6.15)
Cataract	4 (2.45)	0 (0.00)	4 (1.24)	2.45	(0.07, 6.15)
Blurred vision	4 (2.45)	0 (0.00)	4 (1.24)	2.45	(0.07, 6.15)

CI, confidence interval; TEAE, treatment emergent adverse event; vs, versus

SOTERIA

The company reports safety data on long-term treatment with sotatercept of all patients included in SOTERIA from the 08-NOV-2023 data cutoff date of the study (representing 922.2 patient-years of exposure) in Tables 67 and 68 of the CS. Fifteen (3.5%) people experienced TEAEs that led to treatment discontinuation, while 11 (2.6%) people had serious TEAEs that led to treatment discontinuation, and 12 (2.8%) people had serious TEAEs that lead to death.

AEs of special interest included:

- Telangiectasia events: No participants with serious events; only 1 event led to discontinuation.
- Thrombocytopenia events: <1% of participants with serious events.
- Bleeding events: Most were non-serious epistaxis events. 5.2% of participants had serious events and <1% led to treatment discontinuation.

The EAG has inspected the safety data and has no concerns about the rates of reported TEAEs or Serious TEAEs in STELLAR and SOTERIA.

The EAG notes that interim results from the ZENITH trial have been published during this appraisal (cut-off date, 26th July 2024).¹⁷ Whilst the patient population in ZENITH differs from the population considered in the company's decision problem for this appraisal, the higher-risk ZENITH population is relevant to the population considered in the original NICE scope. The EAG, therefore, provides a brief description of the ZENITH trial. Participants in ZENITH were PAH patients with WHO functional class III or IV and a high 1-year risk of death (Registry to Evaluate Early and Long-Term Pulmonary Arterial Hypertension Disease Management Lite 2 risk score, ≥ 9) and were receiving the maximum tolerated dose of background therapy. Participants were randomised to receive add-on sotatercept (n=86) starting at 0.3 mg per kilogram of body weight and escalating to a target dose of 0.7 mg per kilogram, or placebo (n=86) every 3 weeks. Compared with placebo, sotatercept treatment was associated with a lower risk of a composite endpoint of death from any cause, lung transplantation, or hospitalisation lasting at least 24 hours for worsening PAH (hazard ratio 0.24, 95% CI 0.13 to 0.43). The most common adverse events that

occurred more frequently with sotatercept than placebo were epistaxis (44.2% vs. 9.3%), telangiectasia (25.6% vs. 3.5%), vomiting (12.8% vs. 5.8%), gingival bleeding (10.5% vs. 2.3%), and back pain (10.5% vs. 4.7%).

3.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

The company describes their indirect treatment comparison (ITC) of clinical effectiveness outcomes of treatment with sotatercept versus selexipag in Section 2.10 of the CS. The company's SLR identified four studies that reported on the clinical effectiveness of selexipag-based treatment regimens: the Phase 3 GRIPHON study (NCT01106014), Phase 4 TRACE study (NCT03078907), Phase 3 TRITON study (NCT02558231) and Phase 2 NCT00993408 study. However, the company deemed that only data from STELLAR, GRIPHON and TRACE were suitable for inclusion in the ITC. *The EAG agrees with the company's interpretation of the suitability of data for inclusion in the ITC.*

Both GRIPHON and TRACE compared selexipag + background therapy versus background therapy alone. A summary of the main baseline characteristics of STELLAR, GRIPHON and TRACE is presented in Table 9.

Table 9 Summary demographics and baseline characteristics of the trials included in the indirect treatment comparison

Characteristic	STELLAR				GRIPHON				TRACE			
	Sotatercept N=163		Placebo N=160		Selexipag N = 574		Placebo N = 582		Selexipag N = 53		Placebo N = 55	
	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)
Age (years)												
Mean	47.6		48.3		48.2		47.9		49.0		49.8	
SD	14.09		15.50		15.19		15.55		14.8		13.6	
Median	47.0		49.0		NR		NR		NR		NR	
Range	18.0 to 81.0		18.0 to 82.0		NR		NR		NR		NR	
Sex, n (%)												
Male	34	(20.9)	33	(20.6)	117	20.4	116	19.9	18	34.0	13	23.6
Female	129	(79.1)	127	(79.4)	457	79.6	466	80.1	35	66.0	42	76.4
Length of time since PAH diagnosis (years)												
Mean	9.2		8.3		2.3		2.5		NR		NR	

Characteristic	STELLAR				GRIPHON				TRACE			
	Sotatercept N=163		Placebo N=160		Selexipag N = 574		Placebo N = 582		Selexipag N = 53		Placebo N = 55	
	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)
SD	7.32		6.70		3.49		3.75		NR		NR	
Median	7.5		6.8		NR		NR		38.1		33.9	
Range	0.7 to 38.3		0.1 to 40.2		NR		NR		5.3 to 198.3		4.7 to 264.8	
WHO Diagnostic pulmonary hypertension Group I: PAH, n (%)												
Idiopathic PAH	83	(50.9)	106	(66.3)	312	54.4	337	57.9	NR	NR	NR	NR
Heritable PAH	35	(21.5)	24	(15.0)	13	2.3	13	2.2	NR	NR	NR	NR
Idiopathic/heritable PAH	-	-	-	-	-	-	-	-	40	75.5	42	76.4
Drug/toxin-induced PAH	7	(4.3)	4	(2.5)	17	3.0	10	1.7	1	1.9	1	1.8
PAH associated with connective tissue disease	29	(17.8)	19	(11.9)	167	29.1	167	28.7	8	15.1	10	18.2

Characteristic	STELLAR				GRIPHON				TRACE			
	Sotatercept N=163		Placebo N=160		Selexipag N = 574		Placebo N = 582		Selexipag N = 53		Placebo N = 55	
	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)
PAH associated with simple, congenital systemic to pulmonary shunts at least 1 year following repair	9	(5.5)	7	(4.4)	60	10.5	50	8.6	NR	NR	NR	NR
Associated with HIV infection	0	0	0	0	5	0.9	5	0.9	0	0	1	1.8
WHO functional classification for symptomatic pulmonary hypertension, n (%)^a												
Class I	0	0	0	0	4	0.7	5	0.9	0	0	0	0
Class II	80	(49.1)	78	(48.8)	274	47.7	255	43.8	33	62.3	41	74.5
Class III	83	(50.9)	82	(51.3)	293	51.0	314	54.0	20	37.7	14	25.5
Class IV	0	0	0	0	3	0.5	8	1.4	0	0	0	0
Type of background PAH therapy, n (%)^b												
Mono	9	(5.5)	4	(2.5)	NR	NR	NR	NR	NR	NR	NR	NR

Characteristic	STELLAR				GRIPHON				TRACE			
	Sotatercept N=163		Placebo N=160		Selexipag N = 574		Placebo N = 582		Selexipag N = 53		Placebo N = 55	
	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)
Double	56	(34.4)	56	(35.0)	NR	NR	NR	NR	NR	NR	NR	NR
Triple	98	(60.1)	100	(62.5)	NR	NR	NR	NR	NR	NR	NR	NR
Use of medications for PAH, n (%)												
None	NR	NR	NR	NR	112	19.5	124	21.3	0	0	0	0
Endothelin-receptor antagonists	NR	NR	NR	NR	94	16.4	76	13.1	1	1.9	1	1.8
Phosphodiesterase type 5 inhibitors	NR	NR	NR	NR	189	32.9	185	31.8	0	0	0	0
Endothelin-receptor antagonists plus phosphodiesterase type 5 inhibitors	NR	NR	NR	NR	179	31.2	197	33.8	52	98.1	54	98.2

^a One STELLAR participant, in the sotatercept arm, is counted in the WHO FC II category in this table (in alignment with the CSR Shift from baseline in WHO functional class by Visit entries). To note is that the same participant, in the CSR Disease History Associated with PAH entry is counted in the WHO FC III category. This discrepancy was caused by an incorrect entry in the WHO FC intake form in the IRT and was recorded to match the IRT entry by the site in the study database. For purposes of efficacy and safety analyses, this participant is counted as WHO FC II.

^b One STELLAR participant, in the placebo arm, is counted in the double therapy category in this table (in alignment with the CSR Disease History Associated with PAH entry). To note is that the participant in the CSR Concomitant Standard of Care Medications is counted in the triple therapy category. The discrepancy was caused by an incomplete end date record in the study database for a concomitant PAH therapy. For purposes of efficacy and safety analyses, this participant is counted as being on double therapy. Source: Hoeper et al. 2023¹⁵; MSD 2024 STELLAR CSR¹⁸; Sitbon et al 2015¹⁹; Howard et al 2023²⁰

A summary of the key ITC results is reported in Table 65 of the CS. All comparisons favoured sotatercept.

The company also attempted to conduct an ITC of sotatercept versus selexipag, using individual patient data from STELLAR (on sotatercept) and real-world evidence (on selexipag) available from the TRIPHIC database. However, the company reports that relevant data were available for too few patients from the TRIPHIC research database to allow for a credible and meaningful indirect comparison to be feasible. Similarly, the company were unable to conduct an ITC of sotatercept versus infused prostacyclin due to the lack of an appropriate common comparator from the studies identified by the SLR. *The EAG accepts the company's reasons for not performing these ITCs.*

3.4 Critique of the indirect comparison and/or multiple treatment comparison

The EAG is satisfied with the methods used to perform indirect treatment comparisons. The company's submission provides a comprehensive description of how the data from GRIPHON and TRACE were pooled, as well as a clear explanation of the use of Bucher for indirect treatment comparison.

There are notable differences in the study populations between the STELLAR and GRIPHON trials, particularly regarding the time from diagnosis to randomisation and the composition of background therapies. In the STELLAR study, participants had a longer duration from diagnosis to randomisation compared to those in GRIPHON. According to EAG clinical experts, participants in the STELLAR trial responded to sotatercept despite having a more stable disease trajectory, as indicated by the extended time since diagnosis (average survival in PAH is 6–7 years, and 9 years suggests a more stable cohort than typically seen in clinical practice). Since participants in GRIPHON were enrolled earlier in their disease progression and management, there may be a survival bias favouring in STELLAR due to the longer time since diagnosis.

The EAG clinical experts also highlighted that STELLAR trial participants responded to sotatercept despite receiving greater background therapy at enrolment compared to GRIPHON, where 20% of participants had received no pulmonary hypertension (PH) therapy and 47%

were on monotherapy. In contrast, triple therapy could not have been possible to be used in GRIPHON as selexipag cannot be given alongside IV epoprostenol or inhaled iloprost. In GRIPHON 67% have received monotherapy or no therapy while the remaining 33% received dual therapy. This contrasts to nearly 62% of STELLAR having received triple therapy. If the STELLAR population had been more like GRIPHON's, a potentially greater treatment effect might have been observed. While subgroup analyses in STELLAR suggest no significant differences in outcomes between participants on double versus triple background therapy, the effects on those receiving monotherapy remain inconclusive due to the limited sample size.

These differences may have influenced the observed treatment effects but in order to make a comparison between sotatercept and selexipag, the Company is inferring that the efficacy of sotatercept could extend to a population with less advanced disease progression.

3.5 Critique of the within trial analysis

The company conducted a post-hoc, within-trial analysis of the STELLAR study to assess the outcome of change from baseline in risk status at week 24. This analysis included only participants in each arm who were treated with the following regimens:

- Patients randomised to receive sotatercept: PDE5i + ERA
- Patients randomised to receive placebo: PDE5i + ERA + selexipag

The EAG does not support this approach, as it does not allow for a direct comparison of the treatment effects of sotatercept and selexipag. Participants in the placebo arm were already receiving PDE5 inhibitors, ERAs, and selexipag prior to randomisation, and no active drug was added to these treatments. As a result, any potential treatment effect of selexipag on changes in risk status from baseline at week 24 will not be observable. This method also breaks the randomisation as it is comparing a group who received double combination to a group who received triple combination at baseline. Those who have received triple combination are more heavily treated and therefore were doing less well than those who received double combination.

3.6 Additional work on clinical effectiveness undertaken by the EAG

None

3.7 Conclusions of the clinical effectiveness section

The EAG is satisfied with the level of detail provided in the clinical effectiveness section of the CS. The EAG agrees that the company's analysis suggests sotatercept is more effective than selezipag across all outcomes, except for changes in risk status from baseline at week 24. However, there are some differences in the comparator populations which may influence the effectiveness of sotatercept relative to selezipag.

4 COST EFFECTIVENESS

4.1 EAG comment on company's review of cost-effectiveness evidence

The Company's systematic literature review (SLR) of published economic evaluations is detailed in section 3.1 of their submission document, with further details on identification and selection methods provided in appendix E. They have provided a useful summary of identified cost-effectiveness studies in Table 69 of their main submission document. The inclusion criteria were broad, including all relevant interventions for PAH and all types of economic evaluation. The focus was on modelling approaches and assumptions. The company note that there have been no prior NICE submissions in this area but identified several submissions to other HTA agencies. They also identified a recent economic analysis of sotatercept conducted by the Institute for Clinical and Economic Review in the US.²¹

The EAG is satisfied that the company's SLR is comprehensive and that it has been well conducted. It is clear from the review, that the majority of prior modelling studies in PAH have been structured around WHO functional classification (WHO FC), which the company, and clinical expert feedback to the EAG, suggests is becoming less appropriate as the management of PAH has recently shifted towards a risk stratification approach based on ESC/ERC guidelines.^{12, 13} Therefore, the company's de novo model represents the first economic model to be structured around this newer categorisation.

4.2 Summary and critique of the company's submitted economic evaluation by the EAG

4.2.1 NICE reference case checklist

Table 10 summarises the company's economic analysis against the criteria of the NICE reference case. The following subsections provide a more detailed commentary.

Table 10 NICE reference case checklist

Element of health technology assessment	Reference case	EAG comment on company's submission

Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Aligns broadly with NICE reference case, including health effects for patients and carers.
Perspective on costs	NHS and PSS	The model aligns with the reference case, although has not explicitly assessed and measured impacts on personal social services.
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	Aligns with the reference case
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	The model time horizon is sufficiently long to capture all differences in costs and outcomes
Synthesis of evidence on health effects	Based on systematic review	Aligns with the reference case, but with limited evidence available to inform comparative efficacy on the key outcome around which the model is structured (ESC/ERS risk strata).
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	The main health state utilities align with the reference case, but decrements associated with mode of administration of PAH therapy, and carer utility increments associated with better patient health status, are less in keeping (vignette and SF-6D based, respectively).
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	Yes, for the main health state utilities and carer health state utility, but no for mode of administration vignettes, which were developed with clinical experts
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Yes, utilities and utility decrements have been derived from values based of a representative sample of the UK population
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	The case is met for application of a severity weighing of 1.2 on QALY gains, based on the absolute shortfall versus the expects QALYs of the age/sex matched general population.

Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	The approach to costing is aligned with the NICE reference case, but personal social service resources have not been explicitly included in the analysis.
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Discounting has been appropriately applied and the recommended 3.5% rate.
PSS, personal social services; QALYs, quality-adjusted life years; EQ-5D, standardised instrument for use as a measure of health outcome.		

4.2.2 Model structure

The company have developed a Markov cohort model to address the decision problem as laid out in their submission. Details of the structure are set out in section 3.2.2 of their submission document. An important departure from previously published economic models in PAH, is the structuring around ESC/ERS risk strata states (low, intermediate-low, intermediate-high and high risk), rather than WHO FC (*I II, III and IV*). This is to reflect changes in clinical management, with treatment decisions based on the newer ESC/ERS risk classification,^{12, 13} the goal being to maintain patients in lower risk states for as long as possible. The risk states reflect increasing risks of mortality, which have been incorporated in the model using hazard ratios relative to the low-risk state. A post-transplant state, to account for a small percentage of patients transitioning to lung transplant from the intermediate-high and high risk states, has also been included. The model further stratifies the risk-based health states for sotatercept by whether patients are on or off sotatercept treatment - according to continuation/discontinuation rules applied in the model. A useful overview of the model states is provided in Table 71 of the CS, with schematics provided in Figures 14 and 15 of the CS.

The feedback from clinical experts consulted by company and the EAG support the company's approach to model structure. The EAG note, however, that it does create significant challenges regarding parameterisation of the model with robust estimates of comparative effectiveness (section 4.2.6). For this reason, the EAG queried the possibility of the company providing scenario analyses using model health states based on WHO functional classes, to create better alignment between the model

structure and available comparative clinical effectiveness evidence. This was suggested at the clarification stage. The company did not provide this and reiterated their stance, that their model structure reflects the current treatment/management landscape, and that it has support of clinical experts and NHS England, who were consulted during the scoping phase the appraisal.

The EAG agree that the model, structured around ESC/ERS risk states, is appropriate for capturing the progression of PAH and that it aligns with the contemporary management and treatment guidelines that are now being used in the UK NHS. However, it may be noted that the current commissioning guidelines for comparator treatments do still refer to WHO FC rather than ESC/ERS risk strata. There is a recognised malalignment between these older commissioning guidelines and current practices based on clinical expert feedback.

The EAG have some concerns regarding the different treatment discontinuation rules applied to sotatercept and selexipag (the main comparator treatment) upon exiting the intermediate-low risk health state after 24 weeks. Upon progression to the intermediate-high risk state, sotatercept treated patients are assumed to remain on treatment with no escalation, whereas selexipag patients are discontinued and initiated on infused prostaglandin I₂ (PGI₂) analogues – primarily intravenous prostacyclin analogues (IV PCA), a treatment that attracts a substantial administration utility decrement on top of the effects of disease progression. The EAG asked the company to provide further justification for the asymmetry, and to explore the impact of aligning discontinuation and IV PCA initiation rules between the arms of the model. The company response (clarification letter, B4) focussed more on why selexipag would be discontinued and IV PCA initiated upon progression to the intermediate-high risk state, and less on why those experiencing the same progression on sotatercept would not. The company referred to the World Symposium on Pulmonary Hypertension (WSPH) treatment algorithm, which indicates the addition of sotatercept to PDE5i + ERA as a treatment option in both intermediate-low and intermediate-high risk PAH. It is not clear to the EAG, however, that this would imply no treatment escalation in the intermediate-high risk state for those who have progressed on sotatercept from the intermediate-low risk state. The company did, however, provide the requested scenario whereby sotatercept is discontinued upon

progression to intermediate-high risk and patients are escalated to receive infused PCA in addition to ERA + PDE5i (consistent with stopping criteria for selexipag). On its own, this had a small effect on the ICER, favouring sotatercept.

A further uncertainty, based on clinical expert opinion received by the EAG, is whether sotatercept would be stopped prior to commencing IV PCA in those progressing. The EAGs clinical experts noted the general paradigm of adding treatments rather than substituting them but also noted the limited experience with sotatercept in routine practice. However, maximal therapy with a four drug combination (including IV PCA and activin signalling inhibitor), is indicated as an option for persistent intermediate high or high risk disease according to the World Symposium on Pulmonary Hypertension (WSPH) treatment algorithm.¹³ Some patients in STELLAR will have been receiving this combination, and the more recently published Zenith trial provides further evidence for the safety and efficacy of this combination.¹⁷

A further structural assumption that raised concerns for the EAG, was that patients initiated on treatment with infused PCAs, upon entry into the high-risk state on sotatercept, were given a one-off chance of improving their risk status, whereas those initiating infused PCA in the intermediate-high risk state, following discontinuation of selexipag, are given no chance of improving their risks status. This seemed inconsistent to the EAG based on clinical expert feedback received, and so the company were asked to further justify this and provide a scenario that allows for improvement in risk status following initiation of infused PCA in the selexipag arm. In response, the company acknowledged the inconsistency, but highlighting the added complexity required to account for this transition history in the economic model. Rather than providing a scenario allowing for risk status improvement following initiation of PCA in the selexipag arm, they provided a scenario that removed the possibility of improvement following its initiation in the high risk state in the sotatercept arm. Indeed, they included this assumption in their revised (post-clarification) base case. On its own, this has limited impact on the ICER. The EAG acknowledge the company's reasoning with respect to the added complexity required to allow for risk status improvement following initiation of PCA, but the company's fix has limitations with respect to clinical validity. It likely downplays the clinical

benefit of commencing infused PCA, which may bias cost-effectiveness estimates in favour of sotatercept.

4.2.3 Population

As indicated in their decision problem and section 3.2.1 of their submission document, the company are positioning sotatercept for a subgroup of the licence population. They are focusing on intermediate-low risk patients who would otherwise be treated with selexipag in combination with a PDE5i + ERA. They are omitting low-risk patients, because these patients are at target without the need for additional therapy. The intermediate-high risk group are being omitted, because they are currently initiated on infused PCA, and the company were unable to identify or generate comparative effectiveness evidence for sotatercept versus this comparator. They also note that

The EAG acknowledge and understand the company's reasoning for the proposed positioning but can foresee some issues with it. Clinical expert advice to the EAG suggests that there is also an unmet need for intermediate-high risk patients. Furthermore, some people with intermediate-low risk status will have previously progressed on selexipag and achieved intermediate-low risk status following initiation of treatment with infused PCA. These intermediate-low risk patients, and others who cannot take selexipag for any reason (e.g. patient preference), are also excluded from the economic case. In their response to the clarification letter, the company noted that sotatercept on top of PDE5i + ERA + IV PCA is not indicated in the intermediate-low risk state according to the WSPH treatment algorithm, and that this maximal therapy is only indicated for high-risk patients. They acknowledged that a clinician may want to consider removing IV PCA from intermediate-low risk patients and replacing it with sotatercept but highlighted the infeasibility of an ITC between PDE5i + ERA + sotatercept and PDE5i + ERA + IV PCA. They reiterated that the expected position of sotatercept is in addition to background therapy with PDE5i + ERA in the intermediate-low risk state.

With respect to characteristics of the patient population at the proposed positioning, the company use baseline mean age, proportion female, and the weight distribution of the STELLAR trial cohort. However, they omit US participants from the weight distribution used for weight-based dosing of sotatercept (see section 4.4). The EAG note that the mean age (■) of STELLAR participants is lower than the mean of patients with PAH on treatment in the UK NHS, based on data from the 2019 national PAH audit (age 58 yrs).⁹ Similarly, data from the COMPERA registry suggests a mean age of 61.5 for patients with PAH in the intermediate-low risk state.²²

4.2.4 Interventions and comparators

The intervention is sotatercept at starting dose of 0.3 mg/kg body weight, increasing to a target dose of 0.7 mg/kg administered subcutaneously every 21 days. The increase to target dose occurs from the second administration visit. In alignment with their positioning, this is in combination with background dual therapy of PDE5i + ERA.

In line with their proposed positioning the company have included selexipag added to PDE5i + ERA as the only comparator in the economic model. As noted above, no economic case is being made against treatment combinations that would be initiated for intermediate-high risk patients, or other combinations received by intermediate-low risk patients who are not receiving selexipag; i.e. those on dual therapy who decline selexipag or those already on IV PCA who have achieved intermediate-low risks status following initiation in the intermediate-high risk state.

The EAG acknowledge the choice of comparator is aligned with the company's chosen positioning but do have some concerns that it may exclude several groups with unmet need who could benefit from sotatercept treatment.

4.2.5 Perspective, time horizon and discounting

For costs, the company adopt a UK NHS perspective, including costs of treatment acquisition, monitoring, healthcare events, and costs of mortality. For QALYs, these are mostly derived from patient reported responses to the EQ-5D, valued by the UK general population. Additional QALY benefits to carers associated with remaining in lower risk health states are also included.

The company's perspective is broadly aligned with the NICE reference case, but the EAG note that PSS costs have not been included in the model. This is unlikely to have an important impact on the ICER. The EAG considers the incorporation of carer utilities to be appropriate for this condition.

The model is run for a time horizon of 30 years, with functionality to increase the time horizon to a maximum of 50 years. When applied to the company preferred baseline cohort, mean age 47.9, the model runs up to age 78.

Inspection of the cohort traces show that 0% of either the sotatercept or selexipag modelled cohorts remain alive after 20 years in the model. The EAG are therefore satisfied that a modelled time horizon of 30 years reflects a lifetime horizon in this patient cohort and that it is sufficiently long to capture all relevant costs and QALYs.

A 3.5% annual discount rate is used in the model for costs and QALYs. Discounting is applied continually in each 12-weekly model cycle. For example, the discount factor in cycle 2 (week 12-24) is 0.98, calculated as $1/1.035^{0.46}$.

The EAG are satisfied that continual discounting is reasonable. The decision to discount continually, rather than in whole year increments does not have an important impact on the ICER. The EAG are satisfied that discounting is correctly implemented in the economic model file and is aligned with the NICE reference case.

4.2.6 Treatment effectiveness and extrapolation

Clinical effectiveness parameters for the company's model were informed primarily by analysis of the STELLAR trial data as detailed in section 3.3 of their submission document.

For Sotatercept, the model was populated using transition probabilities between ESC/ERS risk strata states, derived from a mix of different subgroups of the STELLAR trial:

- A subgroup of sotatercept treated patients that were intermediate-low risk at baseline and had available risk strata data at 24 weeks (N=■), were

used to generate short-term (24 week) risk strata transitions. This assumes that transition probabilities for sotatercept added to added to mono or triple background PAH therapy are generalisable to the indicated positioning of those with intermediate-low risk disease currently on background double therapy (PDE5i + ERA). The baseline characteristics of the subgroup, provided by the company at the clarification stage, show that the majority (██████) were in fact on triple combination at baseline.

- Participants form the full ITT analysis set (FAS), with available risk strata data at 12 and 24 weeks (N=██████), have been used to inform 12 weekly transition probabilities for application in the long-term (post-24 weeks). This assumes that that these transition probabilities, for a cohort of mixed risk status and background treatment status at baseline, recently initiated on sotatercept, are generalisable to the target population of those with intermediate-low risk status on dual background therapy at baseline. It also assumes that the 12-24 risk strata transition probabilities are applicable in the long-term.

Selexipag transition probabilities were estimated by applying relative risks of improvement/deterioration to the estimated transition probabilities for sotatercept. In the company base case, these relative risks were estimated from a post-hoc subgroup analysis (N=██████) comparing sotatercept added to dual background therapy (ERA+PDE5i) with placebo added to the triple background therapy combination of ERA+PDE5i+selexipag.

The EAG have several major concerns regarding risk of bias in the company's approach to populating the clinical effectiveness parameters in the economic model:

1. *The use of data from two different subgroups (neither exactly matching the company's proposed target population) to estimate short and long-term transition probabilities introduces potential for bias. The EAG is concerned that onward transition probabilities for sotatercept treated patients, from the intermediate-high to high risk state, reflect the risks of a cohort recently initiated on sotatercept (many in the intermediate-high risk state at baseline). Given the proposed positioning, however, these should ideally reflect the risk of further progression conditional on progression (on sotatercept) from the*

intermediate-low risk state. The EAG asked the company to consider the potential for bias in relying on the different subgroups to generate short and long-term transition probabilities for sotatercept, and to provide an analysis that included transition probabilities derived only from the subgroup aligned more closely with the proposed positioning, i.e. the group with intermediate-low risk status at baseline (Clarification letter, question B7). The company gave a commentary on potential risk of bias, but did not provide the requested scenarios, citing the small numbers available within the restricted subgroups, and the need to optimise the use of available data. The EAG remain concerned that the company approach may overestimate the effect of sotatercept on onward transition probabilities following progression to the intermediate-high risk state. There is also the related issue of whether sotatercept would be continued in the intermediate-high risk state, in patients who have progressed on it from the intermediate-low risk state, and/or whether infused PCA would be initiated in addition to or in place of sotatercept in these patients.

2. *The post-hoc subgroup analysis used to inform the relative risks of risk strata improvement and deterioration with sotatercept versus selexipag is inappropriate for addressing the company's decision problem. The company aim to assess the relative effects of adding sotatercept versus adding selexipag to dual (ERA+PDE5i) background therapy in those with intermediate-low risk PAH. However, their post-hoc subgroup analysis assesses the effect of adding sotatercept to dual background therapy (ERA+PDE5i) versus adding placebo to triple background therapy (ERA+PDE5i+selexipag). Since patients were established and stable on their background therapy at baseline, it is to be expected that adding a new active treatment (sotatercept) will induce a greater initial response than adding placebo. Furthermore, randomisation is broken with this subgroup analysis, and those on triple therapy at baseline may represent a harder to treat group at higher risk of progression compared to those stable on dual therapy. This analysis, therefore, is at high risk of bias in favour of sotatercept, which is propagated through the model by application of the derived relative risks to the short and long-term transition probabilities. The company have, however, provided an alternative scenario analysis, whereby they use relative risks of improvement/deterioration in WHO FC,*

derived from the Bucher ITC with selexipag using data from the STELLAR, GRIPHON and TRACE trials^{15, 19, 20}. Despite the limitations of the ITC, and the focus on WHO FC, these offer a potentially less biased estimate of relative effectiveness than the STELLAR trial post-hoc subgroup analysis.

Unfortunately, they do not align with the model structured around ESC/ERS risk strata, and clinical expert feedback to the EAG suggest that it may be reasonable to expect larger relative effects of sotatercept on 24 week risk strata change than we see for WHO FC, which is a less granular and less sensitive outcome. Nevertheless, they provide potentially less biased estimates of short-term relative effectiveness.

3. *The company's estimated relative risks of improvement/deterioration versus selexipag, were originally applied as hazard (rate) ratios rather than 24-week relative risks (risk ratios). This was incorrect and lead to transition probabilities in the selexipag arm that were inconsistent with the analysis used to inform them. The company addressed this at the clarification stage and provide updated analyses with more appropriate application of the relative risks. However, there is still uncertainty related to the ongoing application of these relative risks to the 'long-term' transition probabilities (beyond 24 weeks) in the model. The relative risks are for improvement/deterioration at 24 weeks following initiation of sotatercept (on ERA+PDE5i) versus initiation of placebo (on ERA+PDE5i+selexipag), or alternatively for change in WHO FC at 24 weeks based on the ITC between STELLAR and GRIPHON/TRACE. Whilst no improvement is allowed in the long-term transition probabilities, and assumed instead to reflect maintenance of risk status, the relative risk of progression may not be appropriate for informing longer term risks of progression on selexipag versus sotatercept. There is also an issue of misalignment between the follow-up time at which the relative risks were derived (24 weeks) and the time interval of the long-term transition probabilities applied in the model (12 weeks). It should also be noted that the relative risks of improvement ([REDACTED]) and deterioration ([REDACTED]), informed by the STELLAR post-hoc subgroup, are imprecise due to the small numbers informing them. There also appears to be an inconsistency in the way that uncertainty around the*

relative risks has been represented in the economic model, with the random draws from the assigned lognormal distributions under-representing the degree of uncertainty based on the reported confidence intervals.

The EAG also asked the company to justify their use of short-term data (12-24 weeks) from STELLAR to inform long-term transition probabilities for sotatercept, when longer-term data are available for patients who have enrolled into SOTERIA (Clarification question B6). In their response, the company flagged the diminishing numbers currently available for analysis at each advancing timepoint in SOTERIA, suggesting there is currently insufficient data to inform transitions probabilities at the most recent interim analysis. They also note the unblinded nature of SOTERIA, which they say could impact the observed transition patterns.

Treatment discontinuation assumptions and data

Treatment continuation/discontinuation in the company's model is modelled as a function of response at 24 weeks (sotatercept only) and subsequent progression of risk status as discussed above (sotatercept and selexipag). The continuation criteria for sotatercept, applied at 24 weeks in the company model, are based on clinical expert advice received by the company. Advice from their clinical experts suggested that sotatercept discontinuation would be considered in those with inadequate response at week 24, defined as clinical worsening on at least one of the three endpoints used in the risk stratification: a higher WHO FC compared to baseline; a $\geq 10\%$ deterioration in 6MWD or NT-proBNP value respectively at Week 24 compared to baseline. Applying these criteria to the intermediate-low risk subgroup from STELLAR (N=█), gives an estimate of █ being discontinued from sotatercept at 24 weeks; split between those in the intermediate-low (█) and intermediate-high risk (█) health states. These patients who discontinue sotatercept at 24 weeks are assigned to the off-sotatercept health states where they are exposed to transition probabilities derived from the placebo arm of STELLAR (low and intermediate low risk health states) or Roman et al 2012²³ (intermediate high and high risk health states) (Company submission, Table 84 and 85). Those on selexipag are assumed to discontinue upon progression to the intermediate-high risk state, aligned with the fact it is not indicated for patients in this risk strata according to the WSPH treatment algorithm.¹³ They are instead initiated on infused PCA.

The EAG are satisfied with the discontinuation assumptions for selexipag and are also broadly happy with the 24-week continuation/discontinuation criteria applied to sotatercept. The EAG have concerns, however, regarding a logical inconsistency between the transition probabilities applied following discontinuation of sotatercept in the intermediate-low risks state (derived from the ITT full analysis set of the placebo arm of STELLAR) and the on-selexipag transition probabilities (derived by applying relative risks to on-sotatercept transition probabilities). These suggest that patients in the intermediate-low risk state who have discontinued sotatercept due to lack of response at 24 weeks, have a better outlook than those remaining on treatment with selexipag. This seems incongruent to the EAG and is likely an artefact of different cohorts being used to inform different transition probabilities, rather than an evidence-based effect. The EAG believe it is more plausible that patients who have discontinued sotatercept due to lack of response at 24 weeks, will face higher risks of progression than those remaining on-treatment with selexipag (intermediate-low state). However, the bias caused by this inconsistency may be small, since it applies only to the small proportion of the cohort modelled to discontinue sotatercept at 24 weeks.

Mortality

The company have modelled all-cause mortality by health state, reflecting the increasing probability of death by increasing ESC/ERS risk status. To do this, they identified published studies that had reported Kaplan-Meier survival data by ESC/ERS risk strata and explored various options for deriving transition probabilities from the selected data source. Details are provided in section 3.3.2 of their submission document (pages 173-184). The company's preferred data source was a published study by Rosenkranz et al., which reported survival data for cohorts of patients identified from the COMPERA PAH registry, defined by baseline risk status and comorbidity status categories.²⁴ The COMPERA registry is the largest PAH registry in the world, originating in Germany but now also including data from other European countries. To derive mortality transition probabilities, they used reported Kaplan-Meier data by risk status for the cohort with no comorbidities, suggesting clinical experts believed this sub-group to be most similar to patients enrolled in the STELLAR trial. They used data on survival after the first follow-up recorded in the

registry, defined as the first follow-up at least 3 months after initiating therapy; This reflecting mortality by risk status for a treated cohort. Following exploration of various options, the company followed the approach of selecting a preferred parametric (Gamma) survival curve fitted independently to the Kaplan-Meier survival data for the low-risk group. They then estimated survival for higher risk groups by applying hazard ratios relative to the low-risk reference curve. The Gamma reference curve was selected for the low-risk group, on grounds it provided the most plausible extrapolation illustrating an increasing hazard over time, which the company state is to be expected in this patient population based on clinical expert opinion. The hazard ratios they apply to this curve, to estimate mortality in higher risk states, were obtained from a Cox proportional hazards model with indicators for risk status included as explanatory variables, applied to reconstructed IPD from Rosenkranz et al.²⁴ (Company submission, Table 88).

At clarification stage the EAG queried the robustness of the targeted review and questioned whether all relevant studies had been identified. Based on the company's clarification, the EAG are satisfied that the two studies retrieved provide the best available data and agree with the company that data from Rosenkranz, based on the COMPERA data are appropriate to inform mortality in the model.

The company also offered an alternative approach using one dependent parametric regression model with non-low-risk states fitted as covariates in the model. Again, using the Rosenkranz et al data, Gompertz and Weibull showed monotonically increasing hazards and were preferred based on clinical expert opinion. Gamma was excluded as it did not show strongly increasing hazards. Gompertz was chosen as it was felt to give more conservative survival benefits for sotatercept.

The company acknowledge that the mortality modelling approach is highly uncertain, and the EAG agree with those concerns. Specific areas of uncertainty include small samples, especially across higher risk groups, uncertainty about the appropriateness of a proportional hazards assumption, and uncertainty about whether hazards of mortality should increase or decrease over time. For example, patients with low-risk who remain low-risk for a substantial amount of time, may be at reduced mortality risk compared to those whose disease progresses more quickly. Furthermore, the

mortality modelling approach does not fully account for the competing risk of disease progression on the estimated HRs leading to significant uncertainty regarding modelled life year gains. The EAG are of the opinion that neither approach is ideal and the impact of both should be explored. The EAG further note that the decision on whether to use HRs or a single dependent model to estimate survival has important implications for modelled life year gains, as demonstrated in the company's survival curves reported in Figure 22 of the company submission. It should also be noted that different mortality modelling approaches, in combination with alternative assumptions about the underlying baseline characteristics of the cohort could also impact on the most appropriate severity weighting for this assessment.

Heart and lung transplant

The company include small probabilities for lung transplantation in the model, under the assumption that these transitions may only occur for people in the intermediate-high or high-risk states (see Table 86 of the company submission). The probabilities were informed by analysis of real-world data from the COMPERA registry. The analysis seems to have focussed on data from an incident cohort of ■■■ PAH patients diagnosed between January 2009 and June 2022 and has assessed time to heart or lung transplant by baseline ESC/ERS risk status.²⁵

The EAG is satisfied with the assumption that the few transplants conducted in PAH patients will be reserved for those with intermediate-high- or high-risk disease. The precise methods and assumptions used to derive the reported transition probabilities are not, however, clearly reported. The Kaplan-Meier data in the unpublished reference provided by the company shows small numbers of events over time, occurring across baseline risk strata. It is not clear to the EAG how these have been converted to transition probabilities from the intermediate high- and high-risk states.

Adverse reactions

Adverse reactions have not been included in the company model, due to lack of comparative data for sotatercept versus selexipag. The company explained in their response to the clarification letter that selexipag was a background therapy in the STELLAR trial, and so adverse events have not been reported for this subpopulation. They also note that severity of adverse events was not reported for selexipag in the

GRIPHON trial,¹⁹ which hampers an indirect comparison with sotatercept. The company consider the exclusion of adverse event data to be conservative, in favour of selexipag, since more serious adverse events have been reported for selexipag than for sotatercept.

The EAG accept this but also note that the implications of some adverse events may have been implicitly captured through the incorporation of hospitalisation rates by health state and treatment arm in the economic model (see below).

PAH hospitalisation

The company include PAH hospitalisation probabilities to all non-death health states during each cycle of the model, indicating an increasing risk of hospitalisation with increasing risk status. These have been derived from analysis of real-world data from the COMPERA registry, for ■■■ patients diagnosed with PAH between January 2009 and June 2022.²⁵ The unpublished accompanying paper, provided in the company references, reports time to first and subsequent hospitalisation events by baseline ESC/ERS risk status, but the methods used to derive hospitalisation transition probabilities from this analysis were not transparently reported. Nevertheless, the results indicate an increasing risk of hospitalisation with increasing disease severity, which is clinically intuitive. The company also apply hazard ratios to the derived risks in the low and intermediate-low risk states, to capture the significant effect of sotatercept versus selexipag on the risk of hospitalisation. This hazard ratio for the effect of treatment was derived from the ITC between the STELLAR and GRIPHON trials.

The EAG has some concern regarding the application of the hazard ratio for the effect of sotatercept versus selexipag on top of the health state specific risks, due to the potential for double counting the effect of sotatercept on this outcome; at least some of the effect will be mediated, and already captured in the model, through the effect of sotatercept on disease progression. The company appear to recognise this and have provided a scenario that removes the additional treatment effect, which has a minor impact on the ICER.

4.2.7 Health related quality of life

The company's economic model incorporated several mechanisms by which treatment can impact on QALYs, specifically:

- 6) The impact of treatment on ESC/ERS health-state, with decreasing HSUVs by increasing risk state derived from the trial data.
- 7) HSUV decrements applied to patients requiring IV PCA treatment, or experiencing PAH related hospitalization, or hospitalization for transplant.
- 8) Utility benefits applied to carers of patients who remain in lower risk health states relative to the utility experienced by carers of patients in the high-risk health state.

The EAG summary and critique of each of these points is addressed in the sections that follow.

Health state utility values by ESC / ERS risk status

The company calculated HSUVs for each health state, defined by ESC / ERS risk strata, using EQ-5D-5L data collected in the STELLAR trial. The NICE DSU recommended crosswalk algorithm was applied to generate EQ-5D-3L HSUVs.²⁶ EQ-5D data were pooled across trial arms from the full analysis set, using unadjusted means. Only participants for whom HSUVs were available both at baseline and 24 weeks were included. One study was identified that reported HSUVs by risk-status, but with a sample of n=6 and n=4 in low and high-risk groups respectively. Other studies were identified through literature review, reporting HSUVs by WHO functional class. A scenario analysis was provided using WHO FC utilities from the STELLAR trial, applied to the most closely corresponding risk strata states.

At the clarification stage, the company provided numbers with available data by risk strata at baseline and 24 weeks, and suggested that the numbers available for the intermediate-low risk and dual background therapy subgroups were too few to inform values by risk strata for these specific subgroups which align more closely with the population of interest (See Table 11 to 13 of the company response to the clarification letter).

Ideally, the EAG believes it may have been preferable for the model HSUVs to reflect 24-week EQ-5D data for a cohort with intermediate-low risk status at baseline, with

adjustment for baseline score. However, the EAG acknowledges that there would be too few observations in some of the risk strata at 24 weeks to adequately inform values. The EAG therefore accepts the company's approach to calculating HSUVs as being reasonable given the available data. The EAG note that these values will to some extent capture the impact of IV PCA use on health-related quality of life. However, since the proportion of IV PCA at baseline (████) was reasonably flat across baseline risk strata, it seems unlikely that differences between the mean HSUVs are driven by differences in proportions on IV PCA across the risk states.

Patient utility decrements for PAH related events (IV PCA treatment, transplant and PAH related hospitalisation)

The company apply utility decrements for these additional events in the model.

For IV PCA administration, the company used a study in which health state descriptions reflecting ERA/PDE5i modes of administration (twice daily oral drug, nebulization, continuous subcutaneous infusion, or continuous intravenous infusion with a permanent catheter) were developed based on literature review and interviews with three clinical experts. The states were otherwise identical in terms of describing PAH, its symptoms, and impact on functioning. A sample of the UK general population (n=150) finally completed a TTO valuation interview for the resultant health state vignettes, following the standard protocol for the valuation of EQ-5D-5L health states. The difference in mean TTO score between the oral administration and the continuous IV infusion state (-0.307) is applied in the company model as a utility decrement for those on IV PCA. In response to clarification queries, the company provides an additional scenario analysis where a decrement equal to 50% of that estimated from Davies et al was applied in the model.

The EAG raises several points of critique regarding the IV administration utility decrement applied in the model:

- *The approach taken by Davies et al., 2018,²⁷ does not align with the NICE reference case for utility estimation. The health states were not reported by patients using a generic descriptive system but rather developed with clinical experts. Whilst the resultant vignettes were valued by a UK general population sample, it is unclear how closely these values would align with the values of EQ-*

5D health states reported directly by patients with PAH on different modes of treatment administration. The EAG note that it might have been possible for the company to inform the utility impact of IV PCA by regression analysis of the EQ-5D scores from STELLAR on health state and IV PCA use, which would more closely align with the NICE reference case. For these reasons, the EAG is not satisfied that the magnitude of utility decrement applied in the economic model appropriately reflects the impact of IV PCA on quality of life and it is also concerned about transferability from the source study to the modelled setting.

- *The EAG's clinical experts note that, in UK clinical practice, IV PCA treatment would be targeted at patients who are most likely to be able to tolerate the treatment and who have a reasonable chance of achieving an improvement in quality of life. A systematic review of clinical effectiveness evidence has shown that treatments like epoprostenol and iloprost can significantly improve patient relevant outcomes such as 6-minute walk distance, with a tendency for improvement in WHO FC.²⁸ For example, one RCT of inhaled iloprost compared to placebo, reported improvements in 6m walk test, NYHA FC and EQ-5D based utilities between baseline and 12 weeks. EQ-5D utilities increased from 0.49 to 0.58 between baseline and 12 weeks for iloprost but remained unchanged in the placebo group.²⁹ Related economic modelling raises questions about the cost-effectiveness of these treatments in addition to supportive care, but modelling consistently demonstrated QALY gains compared to supportive care for all explored treatments across a range of scenario analyses undertaken.²⁸ Whilst the EAG acknowledges that there may be some potential utility loss associated with the additional burden of IV PCA administration and restrictions on usual activities, these would likely be offset by the utility gains from treatment benefit. The EAG acknowledges that ESC / ERS health state utility values might not fully capture the quality-of-life reduction from IV PCA administration in higher risks states, given that a substantial proportion of patients in STELLAR were on parenteral PCA at baseline. However, the magnitude of any uncaptured decrement is unclear, and on balance the EAG feels that the company has not provided sufficient evidence to justify the application of an additional utility decrement for IV PCA administration in the model. This is particularly so, in the*

context of the model's simplifying assumption, of not allowing any improvement (regression) in risk status following initiation of IV PCA.

For clinical events, i.e., (PAH hospitalisation and lung / heart transplant), a one-off utility decrement was applied in each model cycle to the proportion of the cohort in the selexipag and sotatercept arms of the model experiencing a PAH related hospitalisation or transplant. The utility decrement of -0.105 was identified from a targeted review performed by the company to identify utility decrements associated with hospitalisation in PAH. The company considered hospitalisation for HF to be a reasonable proxy for PAH and chose a reported utility decrement for hospitalisation within the previous 30 days, from a study by McMurray et al., for application in the model.³⁰

The source study is based on EQ-5D estimates of the utility impact of hospitalisation in heart failure populations. The EAG are satisfied that in the absence of other data, HF may be a reasonable proxy for PAH. However, the EAG have two concerns with the approach taken. First, there is a risk that applying additional utility decrements for PAH hospitalisation partially double counts the utility impact of these events in the model; i.e. The model HSUVs may already capture the utility impact of increasing hospitalisation risk by ESC/ERS risk status, and so the beneficial effect of sotatercept on this outcome, mediated through its effect on progression risk. Secondly, the magnitude of the utility decrement that should be applied is unclear. For example, the source study identified by the company also shows a much smaller utility decrement for hospitalisations occurring in the previous 30 to 90 days (-0.054).

Furthermore, the EAG note that the company submission suggests that the utility decrement for hospitalisation should be applied to the proportion of the cohort experiencing a hospitalisation event in each model cycle for the duration of the model cycle (12 weeks, ~90 days). Inspection of the economic model shows that PAH hospitalisation decrements have not been adjusted for cycle-length, potentially further over-estimating the QALY losses of PAH hospitalisation (equivalent to assuming patients are at the reduced level of utility for a full year following hospitalisation). On balance, the EAG believes it may be reasonable to apply utility decrements in the model, but at a reduced level and for the duration of the model cycle: calculated as $(0.333 \times -0.105 + 0.667 \times -0.054)$, reflecting the 30 and 90 day estimates of utility loss

reported in McMurray et al³⁰. Despite uncertainties in the approach taken, the EAG acknowledges that the magnitude of utility decrement applied to hospitalisation events is not a major driver of cost-effectiveness results.

Adverse event disutilities

The company has chosen not to include any adverse event disutilities in the economic model on the grounds that the proportion of participants experiencing severe adverse reactions was similar in the sotatercept and placebo groups of the trial.

EAG agree with the company's statement in the submission that excluding adverse events is a conservative assumption. The EAG's clinical expert advice is that selexipag treated patients would be expected to have more frequent adverse events, and of greater severity when compared to sotatercept in UK clinical practice. Any bias associated with the company's omission of adverse event costs and disutilities is likely to bias results in favour of selexipag in the economic model.

Impact on carer quality of life

The company included carer utility in their model, using evidence from a published paper exploring the relationship between patient and carer health state utility based on SF-6D index scores.³¹ This suggests an improvement of 0.1 in patient health state utility is associated with an increase of 0.0123 in carer utility. Rather than using this to generate decrements associated with more severe health states, as is more common in HTA, the company has generated carer utility increments associated with better patient health states with reference to the worst (high-risk) health state. This, they note, avoids the paradox of patient mortality resulting in reduced carer disutility and QALY losses. The company base case assumes one carer per patient, but noting uncertainty associated with the values used, which are based on SF-6D and not specific for PAH, they assess the impact of removing carer utility in a scenario analysis.

The EAG acknowledges that there is likely to be an impact of PAH on carer's quality of life, particularly in higher risk health states where patients require additional support with everyday tasks. The approach taken by the company is novel in HTA, applying a carer benefit rather than utility decrement. The EAG considers the

approach taken to be reasonable, but makes two specific observations that lead to uncertainty in the magnitude of carer impact applied in the economic model:

- *The ratio of the utility impact for carers: patients is obtained from SF-6D utilities, which are out with the NICE reference case, but applied to EQ-5D data from the STELLAR trial in each health state. This generates some uncertainty regarding the magnitude of carer utility impact applied in the model. However, the impact of this issue alone on the ICER is likely to be minimal, and reasonable given that the EAG are unaware of any EQ-5D carer QoL impacts calculated using the same methodology.*
- *The approach taken is based on a study linking patient utility impacts to carer utility impacts, using longitudinal data across 13 waves of “understanding society”. The study is based on respondents self-identifying as a carer, and it is unclear how the impact on their quality of life translates to people caring for patients with PAH. This is a substantial uncertainty, and it is unclear whether any bias would favour sotatercept or selexipag when applied in the economic model.*

Age and sex-adjustment of health state utility values over time in the model.

The company’s base case analysis applies a multiplicative approach to age adjust health state utility values over each subsequent year in the economic model.

The EAG agrees that the company’s approach is consistent with NICE DSU recommendations³² and prefers this to the company’s alternative scenario approach of capping utilities at general population norms.

Summary of company and EAG preferred HSUV assumptions

A summary of the company and EAG preferred HSUVs is provided in Table 11. The net impact of applying all the EAG preferred assumptions is a substantial increase in the ICER for sotatercept compared to selexipag.

Table 11 Summary of company and EAG preferred HSUVs

	Company preferred approach	EAG preferred approach	EAG brief commentary
ESC / ERS health state utility values	HSUVs based on pooled baseline and 24-week data from the STELLAR trial	As per company approach	Whilst the EAG would ideally prefer the use of 24-week utilities, amongst patients with low-intermediate risk at baseline, adjusted for baseline EQ-5D utility, the preferred approach would be limited by small numbers of participants which would increase uncertainty.
Utility decrement for IV PCA	-0.307	0	EAG approach removes the risk of double counting and acknowledges that net impact may not be negative (treatment benefits would be anticipated to offset additional burden of treatment management).
Utility decrement for PAH related clinical events (hospitalisation / transplant)	-0.105	-0.071	EAG approach calculates a reduced decrement based on the utility decrements estimated for hospitalization within 30 days and between 30 and 90 days in the source document, and applies this for the duration of model cycles (12 weeks).
Carer utilities	Utility benefits of 0.036, 0.023 and 0.013 for low, low-intermediate and high-intermediate compared to high risk.	As per company approach	The company's approach is reasonable.

Key: EAG, external assessment group; HSUV, health state utility value; IV PCA, intravenous prostacyclin A; PAH, pulmonary arterial hypertension; ESC, European Society of Cardiology; ERS, European Respiratory Society.

4.2.8 Resources and costs

The company's economic model incorporates the following categories of costs:

- 1) Treatment acquisition and administration costs for the sotatercept and selexipag interventions.
- 2) Treatment acquisition and administration costs for other treatments, including infused PCA.
- 3) Health care resource use costs applied by ESC / ERS health state.
- 4) PAH related hospitalisation costs.
- 5) Lung transplant costs.

Treatment acquisition costs (Sotatercept and Selexipag)

Sotatercept treatment acquisition costs are calculated using weight-based dosing, as described in the STELLAR trial protocol. An initial one-off loading dose of 0.3 mg/kg is applied to cover the first 21 days of treatment. The loading dose is then followed by a dose of 0.7 mg/kg every 21 days thereafter, unless dose adjustments are required. Dose adjustments were not, however, considered in the company's approach to calculating sotatercept treatment acquisition costs in the model.

Sotatercept is available as 45mg and 60mg subcutaneous injections, with a list price of £5,422.50 and £7,230.00 respectively, equating to a cost per mg of £120.50. The company's economic model applies a simple discount, which is confidential. For the weight-based costing, the company calculate the minimum number and combination of 45mg and 60mg vials required for 7kg weight bands as described in Table 100 of the company submission. The proportion of the treated cohort in each weight band is calculated using a log normal distribution defined using the average weight and standard deviation (SD) of the non-US population in the STELLAR trial, (■■■■■ kg; SD=■■■■■ kg). It is not possible to vial share as treatment is provided at home and so the company include wastage in treatment acquisition cost calculations.

The EAG are satisfied that the dosing approach taken within the STELLAR protocol is aligned with the SmPC and reflects the intended use of sotatercept in UK clinical practice. The EAG were not convinced that the company's decision to exclude trial participants from the US when calculating the average weight for treatment acquisition cost calculations was well justified. The company's justification for removal of US participants to estimate average weight is that US participants have on

average higher weight than UK participants. CDC data for the USA, published in 2015/16, show a mean weight amongst US adults aged 40-59 was 91kg (male) and 80kg (female).³³ By comparison, Health Survey for England 2021 data show that adults aged 45-54 were 89.1kg (male), and 75.2kg (female).³⁴ Whilst there is a difference between average weights in the US and UK, these are likely to be offset by the inclusion of several other countries with lower average weights than in the UK. Therefore, by excluding countries with average weights higher, but not lower than the UK, the company's approach may under-estimate patient weight and consequently treatment acquisition costs in the UK. In response to a clarification question, the company explained that it was not meaningful to calculate average weight based on UK recruited participants due to the small sample of [REDACTED]. The company provided a comparison of several different calculations for average weight in response to a clarification question, summarised in Table 12 below. The EAG agree that a sample of [REDACTED] is too small to inform treatment costs but welcome the alternative scenarios to assess uncertainty in the ICER. The alternative scenarios produce a small increase in the ICER.

Table 12 Comparison of alternative estimates of patient weight used in treatment acquisition cost calculations [reproduced from Tables 17 and 18 of the company response to the clarification letter]

	Patient weight (kg)				Mean treatment acquisition costs per 12-week cycle
	N	Mean (SD)	Median	Min, Max	
All countries (ITT)	323	72.3 (18.32)	68.2	38,0, 141.3	
Non-USA participants					
Intermediate-low risk (all countries)					
Double PAH background therapy (all countries)					
Triple PAH background therapy (all countries)					

ITT, intention to treat; PAH, peripheral arterial hypertension

Treatment acquisition costs for the comparator (selexipag) are applied in the ESC/ERS low-risk and intermediate low-risk states only. The cost of treatment is calculated based on self-administration of oral tablets, taken at home, at a dose of 1.6mg twice daily. Selexipag costs are not incurred in the proportion of the sotatercept cohort who are off treatment. The model does not include any within state treatment discontinuation or dose adjustments. There is no adjustment for wastage given the oral administration approach. The unit cost per pack of 60 x 1.6mg tables is £3000, sourced from the BNF, leading to a treatment acquisition cost of £8,400 per 12-week model cycle. Confidential prices for selexipag are available to the NHS and are reported in a confidential appendix to this report.

The EAG's clinical experts consider the dosing and treatment acquisition costs for selexipag to be appropriate and reflective of UK clinical practice. The EAG note, however, that the company has not included any dose adjustments for sotatercept, selexipag or any other treatments included in the model. This could be a cause for concern as it could over cost and bias against treatments with lower RDIs or where dose reductions might be required more regularly. The magnitude and direction of any bias is unclear as it depends on the respective RDIs of sotatercept and selexipag treatment.

Treatment administration costs (Sotatercept and Selexipag)

A one-off cost is applied in the first model cycle, in the intermediate-low risk health state to capture sotatercept treatment initiation costs. Based on expert opinion from one UK centre, 3.5 face-to-face appointments were assumed (2 to seek agreement and training, 1 for first administration, with 50% requiring an appointment for the second administration). Appointments were costed at £95.80 per appointment for a non-consultant led, multi-professional, non-admitted, first face-to-face attendance - General Medicine clinic (code: WF01B). It was assumed that a further 5 remotely conducted blood tests would be required to monitor for adverse events, at a cost of £20 per blood test.

The EAG's clinical expert advisors noted that treatment administration costs for sotatercept would likely vary across the country, would be dependent on local practices at individual centres, and may depend on the level of experience using the treatment in clinical practice. However, on balance, it was felt that the treatment initiation costs included in the company's model would be an under-estimate of real-world clinical practice in England because outpatient consultations would most likely take place in consultant led respiratory / physiology clinics which incur higher average costs (£168.28) than the £95.80 used in the company base case analysis.

There was further uncertainty amongst the EAG's clinical experts with respect to whether sotatercept administration would require a day-case admission, or whether this would take place on an outpatient ward. There was also some uncertainty with respect to the number of supervised administrations that might be required, with one of the EAG's experts feeling that a total of 5 appointments may be required for

sotatercept initiation. The EAG has therefore conducted a further exploratory scenario that includes 2 outpatient attendances + 3 day-case admissions for monitored treatment administration over the first three doses, costing using HRG code: DZ19N (Other Respiratory Disorders without Intervention, with CC Score 0-4). The alternative assumptions are explored in an EAG scenario analyses, but it is noted that these assumptions are not a major driver of the ICER.

The EAG and company preferred unit costs of treatment initiation with sotatercept are compared in Table 13.

Table 13 Comparison of company and EAG preferred sotatercept treatment administration costs

Resource use	Company preferred assumption	EAG alternative assumptions
Outpatient appointments to agree to treatment, provide training and monitor initial administration.	3.5 attendances at £95.80, sourced from NHS reference costs 2022-23 (Non-Consultant Led General Medicine WF01B Multiprofessional Non-Admitted Face-to-face Attendance, First) ^A	3.5 attendances at £168.28, sourced from NHS reference costs 2022-23 (Consultant led, respiratory / physiology service [code 341], non-admitted face-to-face attendance, follow-up [code WF01A]).

^A Note that the EAG were unable to reproduce the company's unit cost of £95.80 given the information provided in the company submission against the quoted NHS reference costs for 2022/23.

The assumed treatment initiation and monitoring costs are substantially higher for selexipag than for sotatercept in the economic model. A one-off outpatient visit to a non-consultant led respiratory physiology service (WF01B, non-admitted first face-to-face attendance), is applied at a cost of £147.30. A further 8 video / telephone calls are assumed in the first 4 months, at a cost of £82.46 per call, leading to a total initiation cost of £806.98. Further routine monitoring call by PAH nurses are then

assumed to be incurred once every 3 months for 90% of patients, and once per month for 10%, leading to a recurring cycle specific cost of £32.98 per 12-week cycle.

The EAG consider the treatment administration costs for selexipag to be reasonable, but as per the clinical expert advice provided to the EAG for sotatercept, it was felt that outpatient appointments would take place in a consultant led respiratory / physiology clinic as opposed to a non-consultant led clinic assumed in the company submission. The EAG therefore also prefers to use a unit cost of £168.28 rather than £147.30 for selexipag treatment administration related appointments. The impact of this change on the ICER is minimal.

Health state specific treatment acquisition and administration costs

The company economic model assumes that all PAH patients, regardless of intervention (sotatercept / selexipag) or ESC / ERS health state require ERA (Bosentan, Ambrisentan, Macitentan) and PDE5i (Sildenafil citrate, Tadalafil) treatments. It is assumed that the proportion receiving the treatments is 100% across all health states and treatment arms.

The EAG note that there may be some variability in the distribution of drugs and dosages for ERA and PDE5i treatments in UK clinical practice. However, the EAG's clinical experts are satisfied with the company's assumption that these treatments are not dependent on initial intervention (sotatercept or selexipag) or on the ESC / ERS risk categorisation. Any biases associated with variation in the distribution of treatment type amongst ERA or PDE5i treatments are therefore minimal as they apply to both arms of the economic model and across all risk health states.

In addition to ERA and PDE5i treatments, it is assumed that all (100%) patients in the high-risk state will receive infused PCA (PGI2 analogue) treatment. Further, it is assumed that the proportion of the sotatercept cohort who have discontinued sotatercept, and the selexipag cohort will also receive PGI2 analogues in the intermediate-high risk state. Three treatment options are considered within the economic modelling in the company base case:

- Epoprostenol IV (61.94%),

- Treprostinil IV (4.95%),
- Iloprost inhaled (33.11%)

The distribution of infused PCA treatments is the same regardless of risk state or initial treatment with sotatercept / selexipag.

The inclusion of infused PCA costs in the intermediate-high risk state for the selexipag arm, but not for those receiving sotatercept in the sotatercept arm means that PGI2 (prostacyclin) analogues are an important driver of cost-effectiveness in the economic model. Whilst the EAG are satisfied that the company's administration costs of PGI2 treatments is reasonable, there are several uncertainties regarding proportions on treatment, treatment by state, and dosing regimens that all have a large impact on the ICER:

- *The decision to exclude PGI2s from the intermediate-high risk state in the sotatercept arm could create a bias in favour of sotatercept. The EAG's clinical advisors felt that it would be plausible that patients would receive PGI2 treatment alongside sotatercept in clinical practice, if they progress on sotatercept from the intermediate-low risks state. The company's response to the clarification letter (Table 3) also confirms that some patients in the STELLAR trial received IV or inhaled PCA at baseline (██████ in the sotatercept arm in the intermediate-low risk group; ██████ in the placebo group). The published manuscript in fact states that 39.9% of STELLAR participant were on infused PCA at baseline. The EAG acknowledges that the definitions of double and triple therapy subgroups for the company's effectiveness assessment of sotatercept vs. selexipag preclude the use of PGI2 at baseline. However, the EAG does not consider this adequate justification for the exclusion of PGI2 costs from the intermediate-high risk state whilst on sotatercept. The EAG therefore explores scenario analyses that include these costs, maintaining a more consistent approach to that taken for selexipag.*
- *The ICER is also sensitive to decisions about the proportion of patients who receive infused PGI2 treatments in the model. The company assume this to be 100% in the base case. The EAG's clinical experts explained that infused*

PGI2 treatment may not be considered suitable for all patients (e.g. due to frailty of comorbidities), and of those offered treatment not all accept it due to the frequent administration of inhaled options and the burden of IV PCA. The EAG's experts felt that, whilst uncertain, it would be reasonable to assume that 85% receive treatment with infused PGI2 analogues upon progression based on these judgments, but that in real world practice it may be lower than this. Company and EAG preferred proportions requiring PGI2 treatments, by treatment arm of the model and health state are summarised in Table 14.

Table 14 Comparison of company and EAG preferred proportions of patients receiving PGI2 treatments

	Company base case	EAG clinical expert opinion
Low risk	0%	0%
Intermediate low risk	0%	0%
Intermediate high risk - sotatercept	0%	39.9%
Intermediate high risk – sotatercept discontinued	100%	85%
Intermediate high risk – selexipag	100%	85%
High risk – sotatercept	100%	85%
High risk – sotatercept discontinued	100%	85%
High risk selexipag	100%	85%

- *The EAG note that the distribution of infused PCA treatment types is sourced from a PAH registry and is therefore the most representative source of UK clinical practice treatment distribution. The EAG's clinical experts also explain that the high proportion receiving Epoprostenol IV compared to Treprostinil IV is to be expected and likely reflective of UK clinical practice. However, there is greater uncertainty with regards to the most appropriate*

dosing regimens applied in the cost calculations for IV epoprostenol and treprostinil. For example, for epoprostenol IV, the company's target dose was obtained from a previous study, quoted in a CADTH appraisal of selexipag, with a mid-point dose 50ng/kg/minute across a range of 40-60ng/kg/min. However, the EAG note that this is substantially higher than the target dose specified in the ESC / ERS guidelines (midpoint of target dose range = 23ng/kg/min). Conversely, the EAG noted that the company's costing of treprostinil was an under-estimate compared to the guideline target dose. The company applies a dose of 30ng/kg/min, whereas the EAG prefers the mid-point of the target dose stated in the guidelines of 42.5ng/Kg/min (midpoint of 25-60 ng/kg/min). The EAG's clinical experts advised that dosage varies across UK clinical practice, but that the dosage used in the company submission seemed overly high, particularly for epoprostenol. The EAG's experts also advised that, on average, treprostinil dosage would typically be twice that of epoprostenol. For these reasons the EAG preferred base case applies the dosages from the ESC / ERS clinical guidelines, as stated in Table 19 of the guideline document. The EAG further notes that the starting dose for IV PCA is much lower than the target dose. However, up-titration of dosage is likely to happen quite quickly (according to EAG clinical expert opinion, by 2ng/kg/min every 48hours for the first few weeks until on 10-12 ng/kg/min, then more slowly by 2ng/kg/min every month until target dose is reached). Not modelling up-titration of dose is, therefore, likely to slightly over cost IV PCA, but conversely covers potential for doses to increase beyond the target dose in the longer term.

The company and EAG preferred 12-week costs on treatment for PGI2 treatments are summarised in Table 15. The impact of applying the EAG preferred assumptions is a substantial increase in the ICER for sotatercept.

Table 15 Comparison of company and EAG preferred PGI2 costs

Treatment	Company preferred 12-weekly cost	EAG preferred 12-weekly cost
Epoprostenol IV	£12,633	£5,811

Treprostinil IV	£3,571	£5,058
Iloprost (inhaled)	£9,244	£9,244

IV, intravenous

Additional costs included in the model

The economic model includes costs of clinical events including PAH hospitalisation, lung / heart transplant and death (see Table 107 of the company submission) as well as a recurring cost per cycle for selexipag, and PCA treatments (Table 106).

Whilst there are some uncertainties regarding the magnitude of resource incurred, the EAG do not have any major concerns with the inclusion of these costs in the economic model. However, the EAG does re-iterate the concern raised in Section 4.2.6 that applying an additional relative risk for hospitalisation may partly double count the impact of sotatercept treatment on hospitalisation costs. However, any bias is likely to be small in magnitude and not a major driver of the ICER.

Adverse event costs

Adverse event costs were not included in the model.

As noted in the EAG critique of adverse event disutilities above, the exclusion of adverse event costs from the economic model is likely to create a bias of uncertain magnitude in favour of selexipag in the economic model.

Summary of company and EAG preferred costing assumptions

A summary of the company and EAG preferred costing assumptions is provided in Table 16. The net impact of applying all the EAG preferred assumptions is a substantial increase in the ICER for sotatercept compared to selexipag.

Table 16 Summary of company and EAG preferred costing assumptions

Costing assumption / approach	Company preferred approach	EAG preferred approach	Justification of EAG preferred approach
Sotatercept treatment administration costs	Apply average weight from STELLAR ITT analysis set, with exclusion of US participants	Apply average weight from STELLAR ITT analysis set without exclusion	Inclusion of participants from countries with higher average weight, but inclusion of participants from countries with lower average weight risks underestimating the average weight for treatment cost acquisition in economic modelling.
Sotatercept treatment administration costs	Based on 3.5 non-consultant led outpatient appointments with multiprofessional team	Based on 3.5 consultant led outpatient appointments in a respiratory / physiology clinic	EAG clinical expert advice is that all outpatient clinics where PAH treatments are initiated on treatment would be consultant led and take place in respiratory / physiology clinics. The EAG preferred approach maintains consistency between sotatercept and selexipag treatment initiation setting.
Use of PGI2 treatments in the intermediate-high risk state on sotatercept	Excluded	Included for 39.9%	EAG clinical expert advice was that patients could be co-treated in clinical practice with sotatercept and PGI2.
Proportion of patients receiving PGI2 treatment	100%	85%	Based on EAG clinical expert opinion, only patients anticipated to benefit would be offered treatment in

			UK clinical practice. EAG's experts felt that, whilst uncertain, on average 70-80% of patients may be offered PGI2, and that 80-90% of those would accept treatment. Taking an average of these ranges gives: $0.75 \times 0.85 = 0.6375$.
PGI2 analogue dosing assumptions	Sourced from study quoted in CADTH appraisal of sotatercept	Sourced from ESC / ERS clinical guidelines.	ESC / ERS clinical guidelines are more representative of UK clinical practice.

Key: EAG, external assessment group; PGI2, prostaglandin I₂; IV PCA, intravenous prostacyclin A; PAH, pulmonary arterial hypertension; ESC, European Society of Cardiology; ERS, European Respiratory Society

5 COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

The Company presented their base case cost-effectiveness findings in section 3.10 of their submission document. However, they revised these in response to the clarification letter. The updated results are provided at the end of the company's response to clarification letter, under the heading of "Updated results". The changes made to the model at this stage included:

- correcting a utility value for the intermediate-high risk health state
- applying the relative risks of improvement/deterioration with selexipag versus sotatercept as risk ratios rather than rate ratios
- removing the one-cycle probability of improvement for patients discontinuing sotatercept and initiating PGI2 analogues (to be consistent with the selexipag arm).

These combined changes resulted in a modest reduction in the base case ICER versus selexipag, reproduced in Table 17 below. These results are confidential and include a simple PAS discount for sotatercept. Selexipag is also available at a discounted price to the NHS, but the company do not have access to this. The EAG will reproduce all necessary results, inclusive of the PAS discount for selexipag, in a confidential appendix. The results are presented with and without a severity modifier of 1.2, for which the EAG agrees the case is met based on absolute and proportional shortfall, under the company base case assumptions and population characteristics (based on the STELLAR baseline characteristics). It should be noted, however, that the case for a 1.2 multiplier may not hold when population age and sex are aligned with the those reported in the UK National Audit of Pulmonary Hypertension, and other assumptions are modified.⁹

Table 17 Deterministic base case results (Source: Table 21 of the company response to the clarification letter)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)	ICER with severity modifier (£/QALY)
Sotatercept + dual therapy	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	2.53	[REDACTED]	[REDACTED]
Selexipag + dual therapy	[REDACTED]	6.37	[REDACTED]	=	=	=	=	=

ICER: incremental cost-effectiveness ratio; LYG: life-years gained; QALY: quality-adjusted life-years

5.2 Company's sensitivity analyses

The company have also provided updated probabilistic sensitivity analysis, one-way sensitivity analysis and scenario analyses at end of their response to the clarification letter (see Tables 23-28 and Figures 3-5 of the company response to clarification). Their base case probabilistic ICER is slightly higher than their deterministic ICER (Company response to clarification, Table 23). Based on one-way sensitivity analysis, the ICER appeared most sensitive to variation in the proportion on PGI2 analogues who receive epoprostenol (in the high-risk or intermediate-high risks states), the drug acquisition cost of iloprost (an inhaled PGI2 analogue), and the relative risk of worsening for sotatercept versus selexipag.

In addition to their original scenario analyses (Company response to clarification, Table 27), the company included a table of new scenarios in response to requests made by the EAG (Company response to clarification, Table 28). Across all scenarios considered, the results display most upward uncertainty to changes in the approach used to model the increasing risk of mortality by risk strata, and the inclusion of sotatercept in combination with PGI2 in the high-risk state of the model (base case assumes sotatercept is discontinued before PGI2 initiation).

5.3 Model validation and face validity check

The company describe steps taken to validate their economic model in section 3.14 of their submission document. They note that clinical expert input was sought to ensure that inputs and assumption were relevant to UK clinical practice, and to ensure that model output was

clinically plausible. They describe two advisory boards and input from seven clinical experts. Technical validation is stated to have followed the guidance of the ISPOR and Society for Medical Decision Making Joint Task Force on Modelling Good Research Practices.³⁵

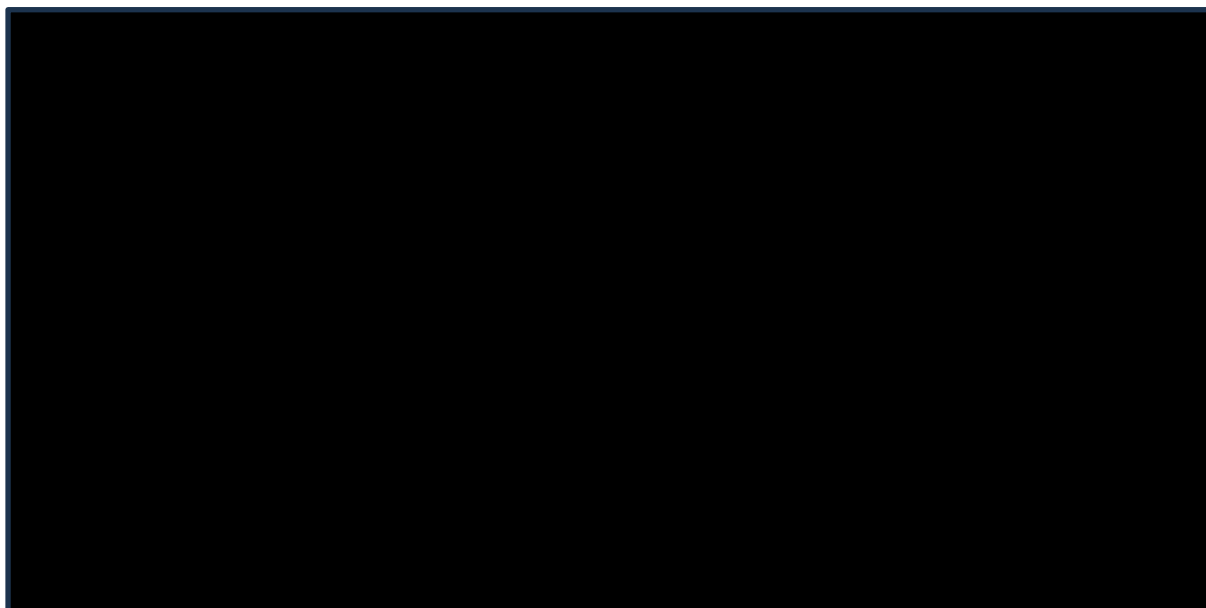
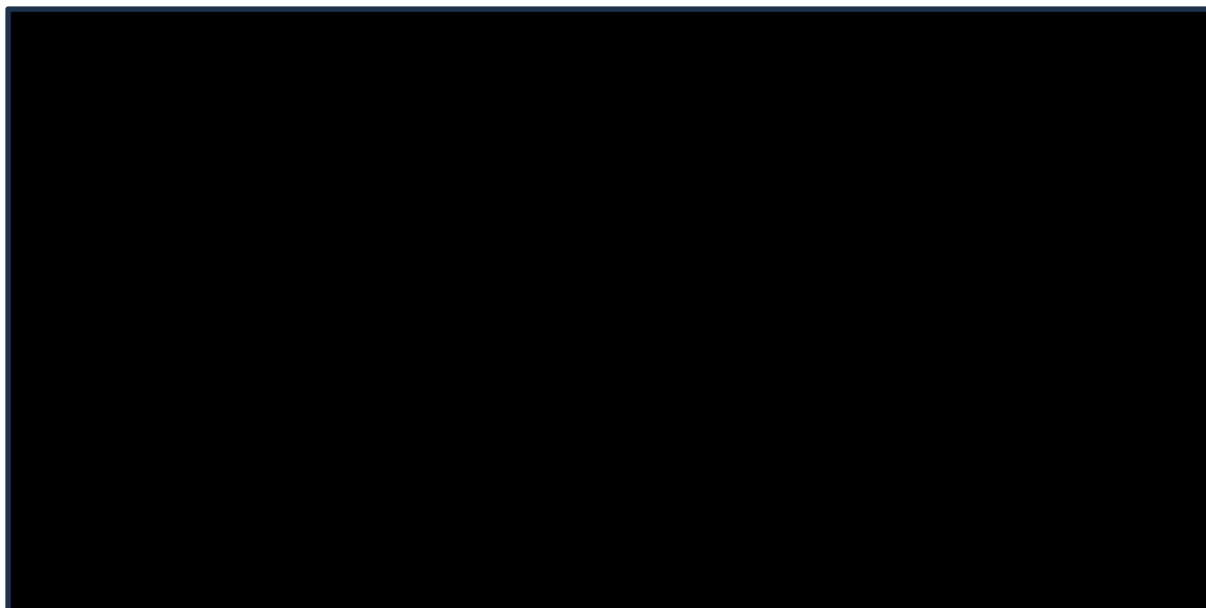
The EAG has conducted its own model checks, through a detailed walk through of the model's logic, calculations and cohort trace. The EAGs review of the model identified a few apparent errors or inconsistencies in the calculations:

1. Relative risks (24-week risk ratios for improvement/deterioration) were originally treated as rate ratios. The company amended this at the clarification stage, but uncertainty remains regarding their ongoing application to longer-term transition probabilities beyond the 24 week period in the model.
2. Utility decrements for hospitalisation did not appear to be adjusted for cycle length in the model, leading to potential overestimation of QALY losses attributable to these events.
3. Lognormal distributions for the relative risks appeared incorrectly specified in the model, with the random draws consequently providing poor coverage of the relevant confidence intervals for point estimates.

The EAG have explored amendments to these issues in chapter 6, all of which have only a minor impact on the ICER.

Beyond the issues noted above, the EAG has concerns regarding the clinical validity of the model output. The modelling assumptions, and ongoing application of 24-week relative risks to long-term transition probabilities, result in patients in the selexipag arm progressing rapidly to the intermediate-high- and high-risk states; [REDACTED] have progressed to these more advanced states by approximately three years. Once in these states, all patients are assumed to have commenced PGI2 analogous, but without the possibility of any improvement in health status. In fact, those on IV PCA (the majority) are assigned a substantial utility decrement. Thus, patients in the selexipag arm of the model spend most of the time horizon in the intermediate-high and high risks states receiving a high-cost treatment for limited clinical benefit (Figure 3). These implications were discussed with the EAGs clinical experts and deemed to be overly pessimistic.

Figure 3 Markov cohort traces for the (a) sotatercept and (b) selexipag arm of the company model under company base case assumptions (Source: Company model)

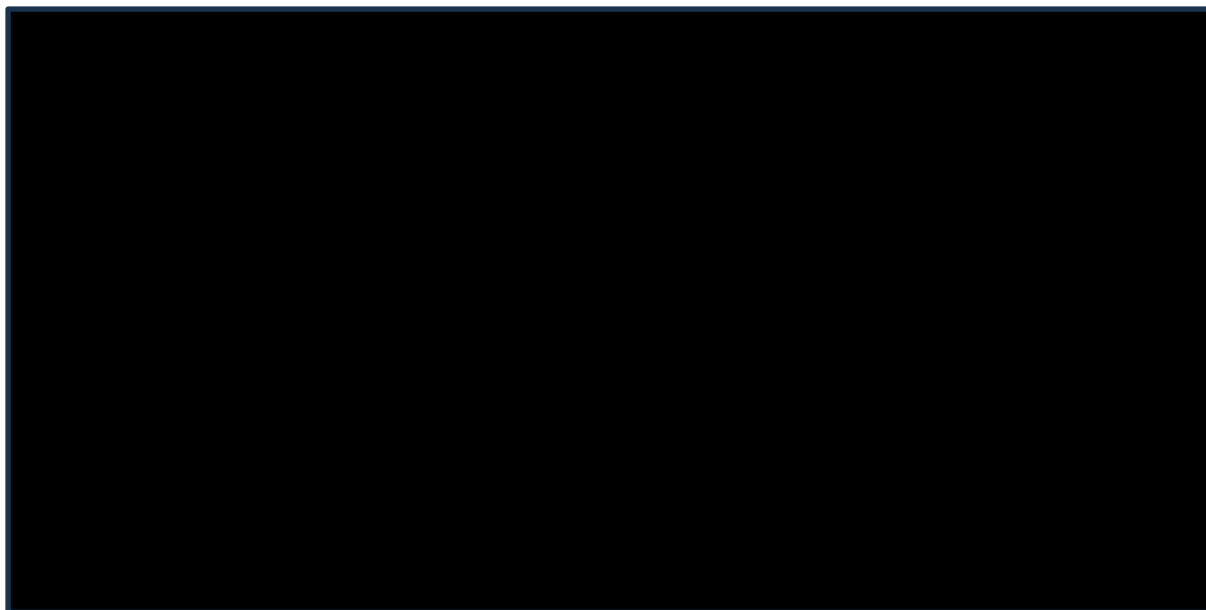


BGT, background therapy; sota, sotatercept

To get a handle on clinical validity, the EAG compared the model output against the primary outcome reported in the GRIPHON trial of selexipag.¹⁹ This was a composite time to event end point of death from any cause, disease progression, or worsening of PAH that resulted in hospitalization, initiation of parenteral prostanoid therapy or long-term oxygen therapy, or the need for lung transplantation or balloon atrial septostomy. Disease progression was defined as a decrease from baseline of at least 15% in the 6-minute walk distance (accompanied by a

worsening in WHO functional class (for the patients with WHO functional class II or III at baseline) or the need for additional treatment of pulmonary arterial hypertension (for the patients with WHO functional class III or IV at baseline). The EAG acknowledge that this outcome is not based on ESC/ERS risk status, and reflects the whole trial population of GRIPHON, and so is not directly comparable with the model output. However, it includes key elements of the model outcome and shows the proportion that were considered alive a progression-free (and not experience hospitalisation for PAH or requiring initiation of parenteral PGI2 analogues) over a substantial duration of follow-up on selexipag. The difference between the model's projection of alive and free from progression (of ESC/ERS risk status), and the observed outcome of GRIPHON, is marked (Figure 4). Clinical expert advice to the EAG suggests that whilst the primary outcome of GRIPHON may underestimate progression on ESC/ERS risk strata, it would not be unreasonable to expect that up to 50% could be maintained in the intermediate-low risk state at three years from initiating selexipag, suggesting the company base case may be substantially underestimating the clinical benefit of selexipag. For this reason, the EAG assess a series of scenarios that factor down the relative risks of progression applied to selexipag in the model (see Chapter 6).

Figure 4 Alive in progression free states (Model output versus GRIPHON Primary Outcome for selezipag)



PO, primary outcome; OS, overall survival

6 EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the EAG

The results of further scenario analyses conducted by the EAG around the company base case are described in Table 18 below, including the rational and cross reference to the preceding sections of the report where the issue being addressed is discussed. Results of these analyses are presented in Table 19 and discussed in section 6.2. The results of these scenarios are deterministic. Probabilistic results are found to be close to the deterministic ICERs in this model.

Table 18 EAG scenarios around the company corrected base case (explanation)

Scenario number	Parameters/assumption	Company base case	Alternative Scenario	Rationale	Relevant section of EAG report
1	Hospitalisation QALY loss calculations	Calculation does not appear to account for model cycle length	Calculation adjusted to the 12 week cycle length.	QALY losses will be overestimated as utility decrement is implicitly assumed to last 1 year when it is stated to apply for the cycle.	4.2.7
2	Long-term transition probabilities from intermediate-high to high risks state on sotatercept	Based on observed 12–24-week transitions in the sotatercept arm of the STELLAR full analysis set	Explore impact of increasing this risk of progression, by a) 50% and b) 100%, conditional on already having progressed on sotatercept from the intermediate-low state.	The company apply long-term transition probabilities that are not consistent with a population starting in the intermediate-low risk state.	4.2.6
3a	Short term Transition probabilities on selexipag (baseline to 24 weeks)	Estimated by applying STELLAR subgroup derived relative risks of improvement/deterioration for selexipag versus sotatercept, to corresponding 24-week transition probabilities for sotatercept	Use more conservative relative risks of WHO FC improvement/deterioration from the company ITC, as proxies for effects of selexipag versus sotatercept on ESC /ERS risk status	Relative risks from the post hoc subgroup of STELLAR, comparing ERA+PDE5i+sotatercept versus ERA+PDE5i+selexipag+placebo, are at risk of bias.	4.2.6
3b			Apply mid-point of relative risks from the STELLAR subgroup and the ITC.		
4a	Long-term transition probabilities on selexipag (beyond 24 weeks)	Estimated by applying STELLAR subgroup derived relative risks of improvement/deterioration, to corresponding 12 weekly transition probabilities on sotatercept	Use more conservative relative risks of WHO FC improvement /deterioration from the company ITC, as proxies for effects of selexipag versus sotatercept on ESC /ERS risk status transitions	The relative risks derived from both sources, reflect relative treatment effects between baseline and 24 weeks, and their suitability for ongoing application beyond 24 weeks, to 12 weekly transition probabilities, is highly uncertain.	4.2.6
4b			Apply 50% of the relative risk reduction/increases from ITC of WHO FC		

4c			Apply 25% of the relative risk reduction/increases from ITC of WHO FC		
5	Transition probabilities from the intermediate-low risk state for patients who discontinue sotatercept at 24 weeks.	Informed by counts in the placebo arm of the full analysis set of STELLAR (intermediate-low)	Assumed equal to the selexipag arm for intermediate-low to intermediate high	Data appear inconsistent, suggesting lower risks of progression from the intermediate-low state for those discontinued from sotatercept at 24 weeks, compared to those on selexipag.	4.2.6
6	Transition probabilities following initiation of PGI2 analogues	Uses TPs between WHO FC states from a published model, as proxies for TPs between ESC/ERS risk strata, but assumes no chance of improvement in risk state.	Unable to assess this due to the need for substantial model restructuring...	Likely underestimates the clinical benefit of commencing infused PCA, which may bias in favour of sotatercept	4.2.6
7a	Proportion of patients commencing PGI2 analogues in both arms of the model	Selexipag arm: 100% upon progression to the intermediate high-risk state; sotatercept arm: 100% upon progression to the high-risk state.	Selexipag arm: 85% upon progression to the intermediate high-risk state; Sotatercept arm: 85% upon progression to the high-risk state. Remainder stay on prior therapy.	Based on EAG clinical expert opinion, that not all patients are considered suitable for infused PGI2 analogues (e.g. due to frailty) and not all accept the treatment if they are.	4.2.8
7b			Selexipag arm: 72% upon progression to the intermediate high-risk state; Sotatercept arm: 72% upon progression to the high-risk state. Remainder stay on prior therapy.		
7c			Selexipag arm: 63.75% upon progression to the intermediate high-risk state; Sotatercept arm: 63.75% upon progression to the high-risk state. Remainder stay on prior therapy.		
8a	Initiation of PGI2 analogues for patients in the sotatercept arm	Upon progression to the high-risk state, instead of sotatercept	Upon progression to the intermediate-high risk state, instead of sotatercept	Clinical expert advice to the EAG suggests PGI2 analogues would be offered earlier to suitable patients who progresses on sotatercept from the intermediate-low to the intermediate-high risk state.	4.2.6 and 4.2.8

8b			100% upon progression to the intermediate-high risk state, on top of sotatercept	Maximal therapy (ERA+PDE5i+IV PCA + activin signalling inhibitor) is indicated in the WSPH treatment algorithm for persistent intermediate-high risk patients, and this maximal therapy has been used in STELLAR and the recently published ZENITH trial.	
8c			40% upon progression to the intermediate-high risk state, on top of sotatercept	40% aligns with percentage of the full analysis population receiving Prostacyclin infusion therapy at baseline in STELLAR, from which transition probabilities have been derived	
9a	Utility decrement of hospitalisation	Decrement of 0.105 per 12-week model cycle	Reduced utility decrements of 0.071 ($=0.33 \times 0.105 + 0.67 \times 0.054$) over the 12-week model cycle	McMurray et al. report a decrement of 0.105 for hospitalisation within 30 days and a lower decrement of 0.054 for hospitalisation within 30-90 days	4.2.7
9b			Removed		
10a	Utility decrement of IV PCA administration	Decrement of 0.307 applied for duration of time on treatment with IV PCA	Apply a 50% reduction to the decrement ($=0.154$) for duration of time on treatment with IV PCA	Applying the company's utility decrement for IV PCA, in combination with the simplifying assumption that PGI2 initiation can provide no improvement in risk status, overemphasises the negative aspects and downplays the benefits of PGI2 analogues	4.2.7
10b			Remove utility decrement of IV PCA		
11	Utility increments for carers of patients in less severe states	Increments of 0.036, 0.023 and 0.013 applied for time in the low, intermediate-low and intermediate-high risks states, respectively	Removed	Uncertain relationship informed by SF-6D index scores in patients and carers across a range of conditions – not specific to PAH	4.2.7
12	Treatment acquisition costs for sotatercept	Company applies patient average weight, excluding US participants	Patient average weight from all STELLAR trial participants	The company approach may underestimate average UK weight by removing patients from countries with higher average weight but	4.2.8

				retaining patients from countries with potentially lower average weight.	
13	Resource use for initiation of sotatercept	3.5 face-to-face appointments were assumed (2 to seek agreement and training, 1 for first administration, with 50% requiring an appointment for the second administration). Appointments were costed at £95.80 per appointment for a non-consultant led, multi-professional, non-admitted, first face-to-face attendance General Medicine clinic (code: WF01B). A further 5 blood tests are assumed to monitor for adverse events at a cost of £20 per blood test.	Two face-to-face appointments to seek agreement and training, at a consultant led respiratory / physiology clinic (2 x follow up attendance) + 3 day-case admissions (respiratory / physiology clinic) to administer the first three doses.	EAG's clinical expert opinion was that, with limited inexperience with sotatercept in current clinical practice, centres in England may admit patients as day cases for monitoring of the first three administrations of sotatercept. And that preceding outpatient contacts would be at consultant led clinics. Whilst contacts could take place at either cardiology or respiratory clinics, for consistency with costing of selexipag, costs of respiratory / physiology clinics are applied.	4.2.8
14	Resource use for initiation of selexipag	Initial outpatient appointment non consultant led	Initial outpatient appointment consultant led	Reasonable to assume all treatment initiation in this setting would be under leadership of a consultant.	4.2.8
15	Dosage of PGI2 analogue preparations	Epoprostenol: 50ng/kg/min Treprostinil: 30ng/kg/min Inhaled Iloprost: 5.0 micrograms x (6-9 times per day).	Epoprostenol: 23ng/kg/min (midpoint of 16-30ng/kg/min). Treprostinil: 42.5ng/kg/min (midpoint of 25-60 ng/kg/min). Inhaled Iloprost: 5.0 micrograms x (6-9 times per day), as per company base case. EAG preferred doses obtained from ESC / ERS guidelines, table 19, targeted maintenance dose.	The EAG prefers to apply dosing for PGI2 analogues as per the ESC / ERS guidelines (See Table 19 of guidelines). Clinical expert advice to the EAG is that dosage varies widely across clinical practice, but that the dosage used in the company submission is an overestimate of what would be a targeted maintenance dose in UK clinical practice. The EAG's experts also advised that the dose of Treprostinil would be, on average, twice that of epoprostenol.	4.2.8

CONFIDENTIAL UNTIL PUBLISHED

Key: EAG, external assessment group; HSUV, health state utility value; IV PCA, intravenous prostacyclin A; PAH, pulmonary arterial hypertension; ESC, European Society of Cardiology; ERS, European Respiratory Society; ITC, indirect treatment comparison; PGI₂, prostaglandin I₂; QALY, quality adjusted life year; WHO FC, World Health Organisation Functional Class; ERA, endothelin receptor antagonists; PDE5i, phosphodiesterase 5 inhibitors

6.2 Impact on the ICER of additional clinical and economic analyses undertaken by the EAG

The impact of the EAGs scenarios on the ICER are provided in Table 19. The greatest impacts are seen: when applying more conservative treatment effects for sotatercept versus selexipag on the long-term transition probabilities for risk status deterioration/improvement (scenarios 4a-4c); when reducing the proportion of patients who initiate treatment with infused PGI2 analogues in according to model's assumptions (scenario 7a-7c); when assuming infused PGI2 analogues are initiated in addition sotatercept for a proportion who progress to the intermediate-high risk state (scenarios 8b and 8c); when the Utility decrement of IV PCA administration is reduced or removed (scenarios 10a and 10b); and when the IV preparation of PGI2 analogues are aligned with European guidelines (Scenario 15). It may be noted that the case for QALY severity multiplier of 1.2 continues to apply across all these individual scenarios, based on the absolute shortfall compared to the age/sex matched general population.

Table 19 EAG scenarios around the company base case (results)

Scenario number	Parameters/assumption	Company base case assumption	Alternative Scenario	Inc. cost	Inc. LYs	Inc. QALYs	ICER (£/QALY)	ICER with severity modifier (£/QALY)
Company base case (post-clarification)				██████	████	2.53	██████	██████
1	Hospitalisation QALY loss calculations	Calculation does not appear to account for model cycle length	Calculation adjusted to the 12 week cycle length.	██████	████	2.48	██████	██████
2a	Long-term transition probabilities from intermediate-high to high risks state on sotatercept	Based on observed 12–24-week transitions in the sotatercept arm of the STELLAR full analysis set	Increase by 50%	██████	████	2.42	██████	██████
2b			Increase by 100%	██████	████	2.32	██████	██████
3a	Short term Transition probabilities on selezipag (baseline to 24 weeks)	Estimated by applying STELLAR subgroup derived relative risks of improvement/deterioration for selezipag versus sotatercept, to corresponding 24-week transition probabilities for sotatercept	Use more conservative relative risks of WHO FC improvement/deterioration from the company ITC, as proxies for effects of selezipag versus sotatercept on ESC /ERS risk status	██████	████	2.43	██████	██████
3b			Apply mid-point of relative risks from the STELLAR subgroup and the ITC.	██████	████	2.47	██████	██████
4a	Long-term transition probabilities on selezipag (beyond 24 weeks)	Estimated by applying STELLAR subgroup derived relative risks of improvement/deterioration, to corresponding 12 weekly transition probabilities on sotatercept	Use more conservative relative risks of WHO FC improvement/deterioration from the company ITC, as proxies for effects of selezipag versus sotatercept on ESC /ERS risk status	██████	████	2.28	██████	██████

4b			Apply 50% of the relative risk reduction/increases from ITC of WHO FC	██████	████	1.89	██████	██████
4c			Apply 25% of the relative risk reduction/increases from ITC of WHO FC	██████	████	1.71	██████	██████
5	Transition probabilities from the intermediate-low risk state for patients who discontinue sotatercept at 24 weeks.	Informed by counts in the placebo arm of the full analysis set of STELLAR (intermediate-low)	Assumed equal to the selezipag arm for intermediate-low to intermediate high	██████	████	2.51	██████	██████
6	Transition probabilities following initiation of PGI2 analogues	Uses TPs between WHO FC states from a published model, as proxies for TPs between ESC/ERS risk strata, but assumes no chance of improvement in risk state.	It has not been possible to conduct a scenario to address this limitation due to the need for substantial structural changes to the model. It is expected that it would increase QALYs and costs (through impact on survival) in both arms, but more so in the selezipag arm.					
7a	Proportion of patients receiving PGI2 analogues in both arms of the model	Selexipag arm: 100% upon progression to the intermediate high-risk state; Sotatercept arm: 100% upon progression to the high-risk state.	Selexipag arm: 85% upon progression to the intermediate high-risk state; Sotatercept arm: 85% upon progression to the high-risk state. Remainder stay on prior therapy	██████	████	2.39	██████	██████
7b			Selexipag arm: 72% upon progression to the intermediate high-risk state; Sotatercept arm: 72% upon progression to the high-risk state. Remainder stay on prior therapy	██████	████	2.27	██████	██████
7c			Selexipag arm: 63.75% upon progression to the intermediate high-risk state; Sotatercept arm: 63.75% upon progression to the high-risk state. Remainder stay on prior therapy.	██████	████	2.19	██████	██████

8a	Initiation of PGI2 analogues for patients in the sotatercept arm	Upon progression to the high-risk state, instead of sotatercept	Upon progression to the intermediate-high risk state, instead of sotatercept	██████	████	1.79	██████	██████*
8b			Upon progression to the intermediate-high risk state, in addition to sotatercept	██████	████	2.02	██████	██████
8c			For 40% upon progression to the intermediate-high risk state, in addition to sotatercept	██████	████	2.33	██████	██████
9a	Utility decrement of hospitalisation	Decrement of 0.105 per 12-week model cycle	Reduced utility decrements of 0.071 ($=0.33 \times 0.105 + 0.67 \times 0.054$) over the 12-week model cycle	██████	████	2.51	██████	██████
9b			Removed	██████	████	2.47	██████	██████
10a	Utility decrement of IV PCA administration	Decrement of 0.307 applied for duration of time on treatment with IV PCA	Apply a 50% reduction to the decrement ($=0.154$) for duration of time on treatment with IV PCA	██████	████	2.06	██████	██████
10b			Remove utility decrement of IV PCA	██████	████	1.59	██████	██████
11	Utility increments for carers of patients in less severe states	Increments of 0.036, 0.023 and 0.013 applied for time in the low, intermediate-low and intermediate-high risks states, respectively	Removed	██████	████	2.42	██████	██████
12	Treatment acquisition costs for sotatercept	Company applies patient average weight, excluding US participants	Patient average weight from all STELLAR trial participants	██████	████	2.53	██████	██████
13	Resource use for initiation of sotatercept	See table X, scenario 15 above	See Table 18, scenario 15 above	██████	████	2.53	██████	██████

CONFIDENTIAL UNTIL PUBLISHED

14	Resource use for initiation of selexipag	Initial outpatient appointment non consultant led	Initial outpatient appointment consultant led	██████	████	2.53	██████	██████
15	Dosage of PGI2 analogue preparations	Epoprostenol: 50ng/kg/min Treprostinil: 30ng/kg/min	Epoprostenol: 23ng/Kg/min Treprostinil: 42.5ng/Kg/min	██████	████	2.53	██████	██████

Key: *Company produced scenario that the EAG has been unable to replicate

6.3 EAG's preferred assumptions

Reflecting on critique provided in chapter 4, and the model validation section in chapter 5, the EAG prefer to make the following changes to the company base case, based on the justifications provided in Table 18 above.

The first relates to a minor calculation inconsistency that the EAG identified in the company model, relating to scenario analyses 1-3 described in Table 18.

1. Hospitalisation QALY loss calculations adjusted to the 12 week cycle length of the model.

The further changes relate to modelling assumptions and inputs as follows:

2. Short term transition probabilities for selexipag informed by application of WHO FC relative risks derived from the ITC with GRIPHON – less risk of bias than the post-hoc subgroup analysis and to these produce model projections that are more in keeping with EAG clinical expert expectations for selexipag.
3. Long-term transition probabilities for selexipag derived by applying half of the [REDACTED] relative risk reduction of disease progression observed for sotatercept versus selexipag at 24 weeks based on the ITC with GRIPHON ([REDACTED] relative risk reduction); i.e. the relative risks of deterioration on selexipag versus sotatercept = $1/[REDACTED] = [REDACTED]$
4. Removal of the independent effect of sotatercept versus selexipag on hospitalisation rate in the low and intermediate-low risk states, to avoid double counting of this benefit which is captured through the effect of sotatercept on progression to higher risks states.
5. Removal of the IV PCA administration disutility (to compensate for the lack of clinical benefit included in the model for these treatments)
6. Reduce the utility decrement associated with hospitalisation from 0.105 to 0.071, applied for the duration of the cycle in which events occur – to align more closely with the values for time periods since hospitalisation reported in the original evidence source for this input.
7. Assume PGI2 analogues are added to sotatercept for 39.9% of patients upon progression to the intermediate-high risk state on sotatercept – this aligns with the proportion on infused prostacyclin therapy in the STELLAR cohort used to inform transition probabilities from the intermediate-high risk state and aligns better with clinical expert opinion provided to the EAG.

8. Assume 85% rather than 100% of patients receive PGI2 analogues following progression to the intermediate-high risk or high risk state on selexipag (or to the high risk state on sotatercept), calculated based on EAG clinical expert opinion that not all patient considered suitable and not all accept it. The remainders stay on prior therapy.
9. Dosing assumptions for epoprostenol and Treprostinil based on target maintenance doses in ESC / ERS guidelines: 23ng/kg/min and 42.5/kg/min, respectively.
10. Assume consultant led OP appointments to initiate treatment on sotatercept and selexipag but retain company assumptions about number of contacts. There appears to be variation in the expected resource utilisation between clinical experts, but it seems reasonable to expect all treatment initiation contacts in this setting would be consultant led.

The cumulative impact of these changes on the ICER, in the order described above, is presented in Table 20. In addition to the central EAG base case, recognising the substantial uncertainty surrounding the relative effect of sotatercept versus selexipag on ESC/ERS risk strata deterioration, the EAG also present two alternative base cases:

1. A lower ICER alternative, whereby the full relative risk reduction for deterioration derived from the ITC between STELLAR and GRIPHON for WHO FC, is applied as the proxy for the long-term effect of sotatercept versus selexipag on ESC /ERS risk status deterioration
2. A higher ICER alternative, whereby 25% of the relative risk reduction for deterioration derived from the ITC between STELLAR and GRIPHON for WHO FC, is applied as the proxy for the long-term effect of sotatercept versus selexipag on ESC /ERS risk status deterioration

The three alternative base cases are presented more fully in Table 21, and Figure 5 illustrates the proportion of the cohort who remain alive in the low or intermediate-low risks states in the selexipag arm versus the sotatercept arm for each of these alternatives and the company base case.

Probabilistic results for the central EAG base case are provided in Table 22 and figures 6 and 7.

The EAG ICERs are substantially higher than under the company case assumptions, this is driven mostly by 5 key changes:

1. Reduced relative effects of sotatercept versus selexipag on the long-term transition probabilities of ESC/ERS risk status deterioration
2. Removal of IV PCA administration disutility
3. The assumption that PGI2 analogues would be added to sotatercept for 39.9% upon progression to the intermediate-high risk state
4. The assumption that a reduced percentage of patients (85% rather than 100%) would receive PGI2 analogues following progression on sotatercept or selexipag to high risk and intermediate high risk states respectively
5. Applying dosing assumptions for epoprostenol and Treprostinil based on target maintenance doses in ESC/ERS treatment guidelines: 23ng/kg/min and 42.5ng/kg/min, respectively

To further assess uncertainty around the EAG base case, the results of further scenario analysis have been provided in Table 23. Notable reductions in the ICER are observed with: The application of the dependent Gompertz model to inform mortality risks for the four risk strata; the removal of the PGI2 analogues for those remaining on sotatercept in the intermediate-high risk state; reinstatement of the utility decrement of IV PCA administration (= -0.307). When applying baseline patient characteristics (age and sex) based on UK National Audit of Pulmonary Hypertension, the ICER drops from the central EAG base case, but the severity modifier no longer applies due there being a smaller absolute QALY shortfall compared to age matched general population for this older cohort.

Table 20 Results using EAG's preferred model assumptions with and without severity weighting (x1.2 where appropriate)

Preferred assumption	Exploratory analysis number (Table 18)	Incremental costs (£)	Incremental QALYs	Cumulative ICER (£/QALY)	Cumulative ICER with severity modifier (£/QALY)
Company base case		██████	2.53	██████	██████
Hospitalisation QALY loss calculations adjusted to the 12 week cycle length	1	██████	2.48	██████	██████
Selexipag short term (24 week) transition probabilities informed by application of WHO FC relative risks from the ITC with GRIPHON	3a	██████	2.39	██████	██████
Selexipag long-term transition probabilities derived by applying half of the █████ relative risk reduction of disease progression for sotatercept versus selexipag at 24 weeks, based on the ITC with GRIPHON (=█████ relative risk reduction)	4b	██████	1.51	██████	██████
Remove independent effect of sotatercept versus selexipag on hospitalisation in the low and intermediate-low risk state	NA	██████	1.50	██████	██████
Remove IV PCA administration disutility (to compensate for the lack of clinical benefit included in the model for these treatments)	10b	██████	0.87	██████	██████

Reduce utility decrement associated with hospitalisation from 0.105 to 0.071, applied for the duration of the cycle in which events occur	9a	██████	0.87	██████	██████
Assume PGI2 analogues added to sotatercept for 40% upon progression to the intermediate-high risk state on sotatercept	8b	██████	0.87	██████	██████
Assume 85% rather than 100% of patients receive PGI2 analogues following progression on sotatercept or selezipag to high risk and intermediate high risks state respectively	7b	██████	0.87	██████	██████
Dosing assumptions for epoprostenol and Treprostinil based on target maintenance doses in ESC / ERS guidelines: 23ng/Kg/min and 42.5ng/Kg/min, respectively	15	██████	0.87	██████	██████
Assume consultant led OP appointments to initiate treatment on sotatercept and selezipag but retain company assumptions about number of contacts		██████	0.87	██████	██████
EAG base case (mid)		██████	0.87	██████	██████
EAG base case (low)	6a	██████	1.24	██████	██████
EAG base case (high)	6c	██████	0.68	██████	██████

Table 21 Alternative EAG Deterministic base case results

Treatment	Total Costs (£)	Total QALYs	Inc. Cost	Inc. QALYs	ICER	Shortfall multiplier	ICER with multiplier
EAG base case (mid)							
Sotatercept + dual BGT	██████	████	██████	0.87	██████	████	██████
Selexipag + dual BGT	██████	████					
EAG base case (low)							
Sotatercept + dual BGT	██████	████	██████	1.24	██████	████	██████
Selexipag + dual BGT	██████	████					
EAG base case (high)							
Sotatercept + dual BGT	██████	████	██████	0.68	██████	████	██████
Selexipag + dual BGT	██████	████					

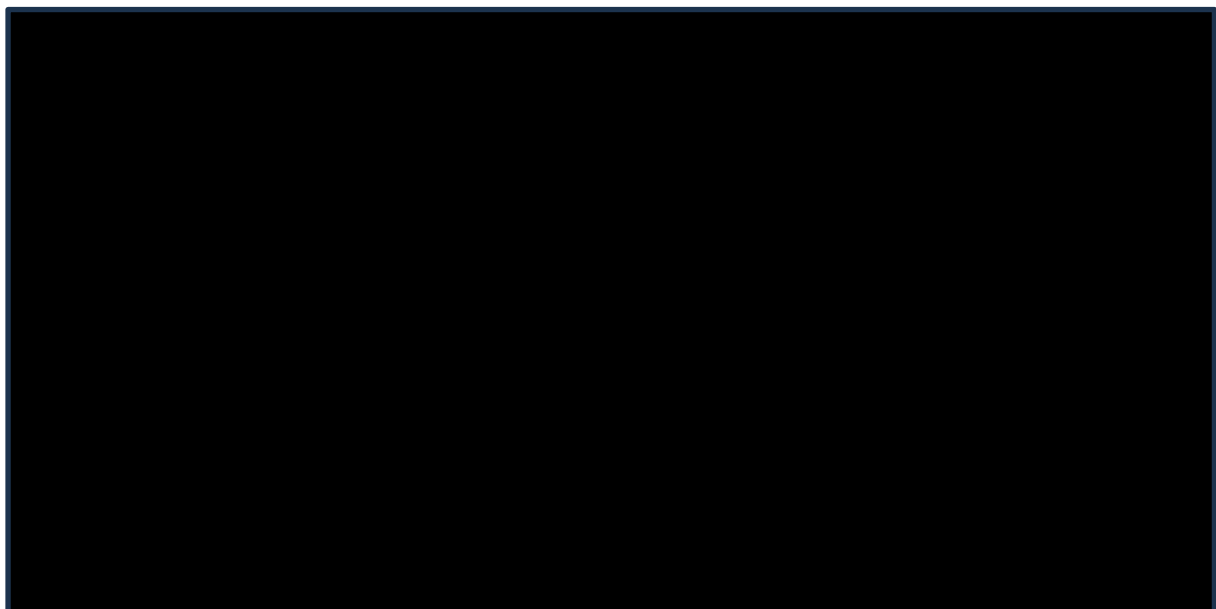
Figure 5 Percentage of patients remaining alive in low-risk or intermediate-low risk states for the company base case and alternative EAG base case options

Table 22 Probabilistic EAG base case (Mid) results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)	ICER with severity modifier (£/QALY)	ICER weighted according to modifier (£/QALY)
Sotatercept	██████	█	████	██████	█	████	██████	██████	██████
Selezipag	██████	█	████	█	█	█	█	█	█

Figure 6 Cost-effectiveness scatter plot (1000 iterations) – EAG base case (Mid) with severity modifier applied

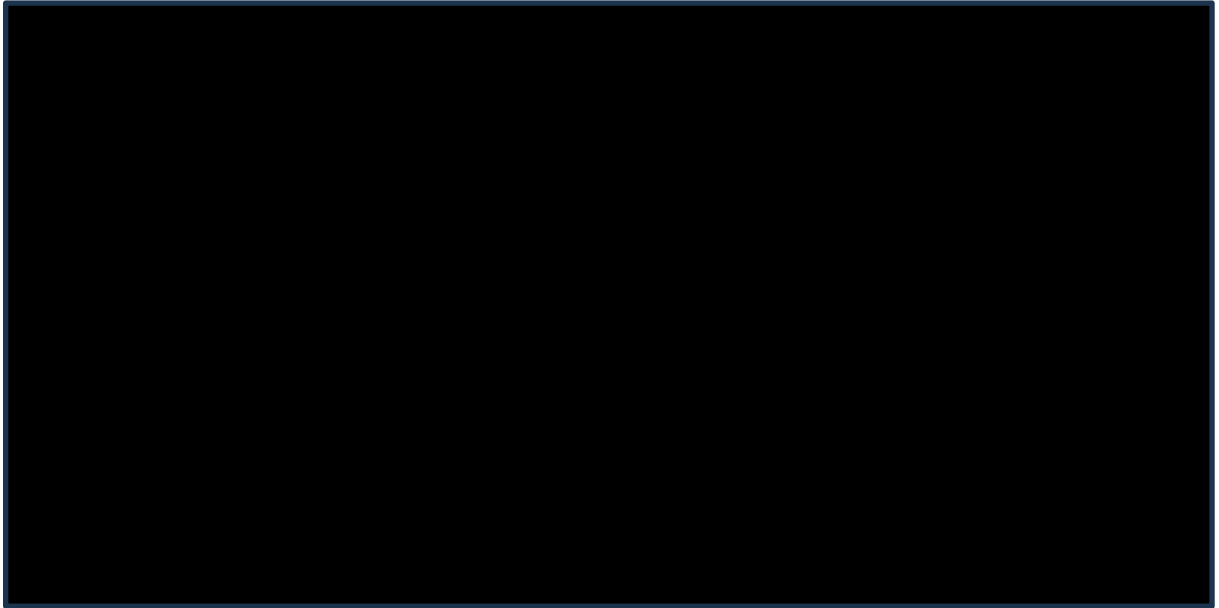


Figure 7 Cost-effectiveness acceptability curves (1000 iterations) – EAG base case (Mid) with severity modifier applied



Table 23 Scenario analysis around the EAG base case (Mid)

Alternative Scenario	EAG base assumptions	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)	ICER with severity modifier (£/QALY)
EAG base case (mid)		██████	0.87	██████	██████
Overall survival – Dependent Gompertz model for the four risk strata	Single gamma model for the low risk group, with hazard ratios from cox regression applied	██████	1.14	██████	██████
No PGI2 analogues initiated upon progression to intermediate-high state on sotatercept	PGI2 analogues for 39.9% in addition to sotatercept in the intermediate-high state	██████	0.87	██████	██████
Percentage of patients receiving PGI2 analogues upon progression to intermediate-high (selexipag arm only) or high (=72%)	= 85%	██████	0.87	██████	██████
Percentage of patients receiving PGI2 analogues upon progression (=63.75%)		██████	0.87	██████	██████
Percentage of patients receiving PGI2 analogues upon progression (=100%)		██████	0.87	██████	██████
Apply utility decrement of IV PCA administration (= -0.307)	No IV administration decrement applied	██████	1.20	██████	██████
Weight based distribution for calculating the dose of sotatercept – STELLAR ITT population	Non-US population of STELLAR	██████	0.87	██████	██████
Baseline characteristics based on UK National Audit of Pulmonary Hypertension*	Baseline characteristics of STELLAR	██████	0.56	██████	██████

Overall survival – Dependent Gompertz model for the four risk strata + Baseline characteristics based on UK National Audit of Pulmonary Hypertension*	Single gamma model for the low risk group, with hazard ratios from cox regression applied + Baseline characteristics of STELLAR	██████	0.88	██████	██████

* Severity modifier no longer applies

6.4 Conclusions of the cost effectiveness section

The company's cost effectiveness case is complicated by the recently established clinical management pathways, shifting towards treatment decisions based on ESC/ERS risk stratification. The company has structured their model around health states defined by the four ESC/ERS risks strata, which does reflect the emerging routine clinical practice in the UK NHS. However, this has created substantial challenges regarding parameterisation of the model with robust estimates of comparative effectiveness for the key outcome around which the model is structured. This results in substantial uncertainty around the company's cost-effectiveness case. Adding to this, the company have made several assumptions which the EAG believe lack strong clinical justification, leading to further uncertainty: The key areas of uncertainty for consideration include:

- Limitations of the evidence for the comparative efficacy of sotatercept versus selexipag on ESC/ERS risk strata deterioration/improvement, relying on a post-hoc subgroup analysis of STELLAR, comparing ERA + PDE5i + sotatercept with ERA + PDE5i + selexipag + placebo.
- Ongoing application of a short-term (24 week) relative risk of deterioration, to generate long-term transition probabilities for selexipag, which the EAG believe results in an overly pessimistic extrapolation of risk status deterioration for selexipag.
- The simplifying assumption of allowing no clinical improvement following initiation of PGI2 analogues, which lacks plausibility and likely biases against selexipag.
- Timing and approach to PGI2 initiation in the sotatercept arm of the model, which is inconsistent with clinical expert feedback received by the EAG.
- The proportion of patients who receive PGI2 analogues, when indicated according to the modelling assumptions, which based on EAG clinical expert feedback should be less than 100%.
- The weight-based doses applied to the IV PCA preparations included in the model are high compared to European guidelines and clinical expert feedback received by the EAG.

8 References

1. Prins KW, Thenappan T. World Health Organization Group - pulmonary hypertension: epidemiology and pathophysiology. *Cardiol Clin*. 2016;**34**(3):363-74.
2. Farber HW, Loscalzo J. Pulmonary arterial hypertension. *N Engl J Med*. 2004;**351**(16):1655-65.
3. Macaulay TE, Covell MB, Pogue KT. An update on the management of pulmonary arterial hypertension and the pharmacist's role. *J Pharm Pract*. 2016;**29**(1):67-76.
4. Levine DJ. Pulmonary arterial hypertension: updates in epidemiology and evaluation of patients. *Am J Manag Care*. 2021;**27**(3 Suppl):S35-s41.
5. Guignabert C, Savale L, Boucly A, et al. Serum and pulmonary expression profiles of the activin signaling system in pulmonary arterial hypertension. *Circulation*. 2023;**147**(24):1809-22.
6. Pulmonary Hypertension Association UK. What it means to live with PH today. Sheffield, UK: PHA UK; 2017. Available from: <https://www.phauk.org/research-survey-work/living-ph-report/>. (Accessed 4 May 2025)
7. Pulmonary Hypertension Association UK. The true emotional impact of pulmonary hypertension. Sheffield, UK: PHA UK; 2019. Available from: <https://www.phauk.org/living-with-pulmonary-hypertension/the-emotional-impact-of-ph-survey-results/>. (Accessed 4 May 2025)
8. Pulmonary Hypertension Association UK. Breathless, not voiceless: what it means to live with PH today. Sheffield, UK: PHA UK; 2023. Available from: <https://www.phauk.org/research/pha-uk-led-research/breathless-not-voiceless-what-it-means-to-live-with-ph-today/>. (Accessed 4 May 2025)
9. NHS England. National Audit of Pulmonary Hypertension Great Britain, 2018-19 (10th annual report). London: NHS Digital; 2019. Available from: <https://digital.nhs.uk/data-and-information/publications/statistical/national-pulmonary-hypertension-audit/2019>. (Accessed 4 May 2025)
10. Thenappan T, Ormiston ML, Ryan JJ, Archer SL. Pulmonary arterial hypertension: pathogenesis and clinical management. *BMJ*. 2018;**360**:j5492.
11. Canadian Agency for Drugs and Technology in Health. CADTH Therapeutic Review Report: Drugs for pulmonary arterial hypertension: comparative efficacy, safety, and cost-effectiveness - recommendations report. Ottawa: Canadian Agency for Drugs and Technology in Health; 2015. Available from: https://www.cadth.ca/sites/default/files/pdf/TR0006_PAH_Recs_Report.pdf. (Accessed 8 May 2025)

12. Humbert M, Kovacs G, Hoeper MM, et al. 2022 ESC/ERS guidelines for the diagnosis and treatment of pulmonary hypertension. *Eur Heart J*. 2022;**43**(38):3618-731.
13. Chin KM, Gaine SP, Gerges C, et al. Treatment algorithm for pulmonary arterial hypertension. *Eur Respir J*. 2024;**64**(4).
14. NHS England. National Pulmonary Hypertension Audit, 15th Annual Report. London: NHS Digital; 2024. Available from: <https://digital.nhs.uk/data-and-information/publications/statistical/national-pulmonary-hypertension-audit/15th-annual-report>. (Accessed 8 May 2025)
15. Hoeper MM, Badesch DB, Ghofrani HA, et al. Phase 3 trial of sotatercept for treatment of pulmonary arterial hypertension. *N Engl J Med*. 2023;**388**(16):1478-90.
16. Preston IR, Badesch D, Ghofrani H-A, et al. A long-term follow-up study of sotatercept for treatment of pulmonary arterial hypertension: interim results of SOTERIA. *Eur Respir J*. 2025:2401435.
17. Humbert M, McLaughlin VV, Badesch DB, et al. Sotatercept in patients with pulmonary arterial hypertension at high risk for death. *N Engl J Med*. 2025.
18. MSD. STELLAR (MK-7962-003) subgroup and post-hoc analyses (06-DEC-2022 LPLV) [data on file]. 2024.
19. Sitbon O, Channick R, Chin KM, et al. Selexipag for the treatment of pulmonary arterial hypertension. *N Engl J Med*. 2015;**373**(26):2522-33.
20. Howard LS, Rosenkranz S, Frantz RP, et al. Assessing daily life physical activity by actigraphy in pulmonary arterial hypertension: insights from the randomized controlled study with selexipag (TRACE). *Chest*. 2023;**163**(2):407-18.
21. Institute for Clinical and Economic Review. Sotatercept for pulmonary arterial hypertension - final evidence report. Boston, MA: ICER; 2024 January 8. Available from: https://icer.org/wp-content/uploads/2023/05/PAH_Final-Evidence-Report_For-Publication_01082024.pdf. (Accessed 16 May 2025)
22. Hoeper MM, Pausch C, Olsson KM, et al. COMPERA 2.0: a refined four-stratum risk assessment model for pulmonary arterial hypertension. *Eur Respir J*. 2022;**60**(1).
23. Roman A, Barberà JA, Escribano P, et al. Cost effectiveness of prostacyclins in pulmonary arterial hypertension. *Appl Health Econ Health Policy*. 2012;**10**(3):175-88.
24. Rosenkranz S, Pausch C, Coghlan JG, et al. Risk stratification and response to therapy in patients with pulmonary arterial hypertension and comorbidities: A COMPERA analysis. *J Heart Lung Transplant*. 2023;**42**(1):102-14.
25. MSD. COMPERA registry: Association between long-term clinical outcomes and risk status / WHO-FC in PAH patients (survival FC risk change) [Data on File]. 2023.

26. Hernández Alava M, Pudney S, Wailoo A. Estimating EQ-5D by age and sex for the UK. 2022.
27. Davies EW, Llewellyn S, Beaudet A, Kosmas CE, Gin-Sing W, Doll HA. Elicitation of health state utilities associated with the mode of administration of drugs acting on the prostacyclin pathway in pulmonary arterial hypertension. *Patient Preference Adherence*. 2018;1079-88.
28. Chen YF, Jowett S, Barton P, et al. Clinical and cost-effectiveness of epoprostenol, iloprost, bosentan, sitaxentan and sildenafil for pulmonary arterial hypertension within their licensed indications: a systematic review and economic evaluation. *Health Technol Assess*. 2009;13:49.
29. Olschewski H, Simonneau G, Galiè N, et al. Inhaled iloprost for severe pulmonary hypertension. *N Engl J Med*. 2002;347(5):322-9.
30. McMurray JJV, Trueman D, Hancock E, et al. Cost-effectiveness of sacubitril/valsartan in the treatment of heart failure with reduced ejection fraction. *Heart*. 2018;104(12):1006-13.
31. Pennington BM, Alava MH, Strong M. Unpaid caring and health-related quality of life: longitudinal analysis of Understanding Society (the UK Household Longitudinal Survey). *Value Health*. 2025;28(1):138-47.
32. Ara R, Wailoo A. NICE DSU Technical Support Document 12: the use of health state utility values in decision models. Sheffield, UK: Decision Support Unit, University of Sheffield; 2011. Available from: <https://www.sheffield.ac.uk/media/34240/download?attachment>. (Accessed 8 May 2025)
33. Fryar C, Kruszon-Moran D, Gu Q, Ogden C. Mean body weight, height, waist circumference, and body mass index among adults: United States, 1999–2000 through 2015–2016: National Health Statistics Reports; no 122. Hyattsville, MD: National Center for Health Statistics; 2018. Available from: <https://www.cdc.gov/nchs/data/nhsr/nhsr122-508.pdf>. (Accessed 4 May 2025)
34. NHS Digital. Health Survey for England: overweight and obesity in adults. 2022. Available from: <https://files.digital.nhs.uk/03/D7F19D/HSE-2021-Overweight-and-obesity-tables.xlsx> (Accessed 4 May 2025).
35. Eddy DM, Hollingworth W, Caro JJ, et al. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. *Med Decis Making*. 2012;32(5):733-43.

Single Technology Appraisal

Sotatercept for treating pulmonary arterial hypertension [ID6163]

EAG report – factual accuracy check and confidential information check

“Data owners may be asked to check that confidential information is correctly marked in documents created by others in the evaluation before release.” (Section 5.4.9, [NICE health technology evaluations: the manual](#)).

You are asked to check the EAG report to ensure there are no factual inaccuracies or errors in the marking of confidential information contained within it. The document should act as a method of detailing any inaccuracies found and how they should be corrected.

If you do identify any factual inaccuracies or errors in the marking of confidential information, you must inform NICE by the end of **2 June** 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.

Please underline all confidential information, and information that is submitted as [REDACTED] should be highlighted in turquoise and all information submitted as '[REDACTED]' in pink.

Issue 1 Trials included in the indirect treatment comparison

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
In the EAG report section 1.1 <i>Overview of the EAG's key issues</i>, in the second paragraph (lines 20-21) it is stated that "An indirect treatment comparison with selexipag is made using data from the GRIPHON trial." However, for several outcomes analysed as part of the indirect treatment comparison data on selexipag from the GRIPHON and/or TRACE trials are used, as noted in section 3.4 of the EAG report and in section 2.10.1 of the Company submission.	This text should be reworded to "An indirect treatment comparison with selexipag is made using data from the GRIPHON and TRACE trials."	Factual accuracy correction of the indirect treatment comparison methods/input data.	Amendment accepted

Issue 2 Intervention name

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
In the EAG report section 3.2 <i>Critique of trials of the technology of interest, the company's analysis and interpretation (and any standard meta-analyses of these)</i>, in the final paragraph it is stated that <i>"However, they noted that the study participants have a longer disease duration, better cardiac index, and more stable disease than patients who would be eligible for etanercept treatment in clinical practice."</i> The intervention name here should be sotatercept and not etanercept.	This text should be reworded to <i>"However, they noted that the study participants have a longer disease duration, better cardiac index, and more stable disease than patients who would be eligible for sotatercept treatment in clinical practice."</i>	Factual accuracy correction of the intervention name.	Amendment accepted

Issue 3 Dizziness adverse event rate numbers

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
In the EAG report section 3.2.4 <i>Adverse events</i> , Table 8 <i>Summary of key adverse events reported in STELLAR</i> , the row for “Dizziness” shows incorrect data (it appears that the data from the “Tooth infection” row from the Company submission Appendix D, section D1.1.1, Table 39 has been used here as opposed to the data from the dizziness row).	The numbers in this row should be replaced with those shown in the “Dizziness” row of the Company submission Appendix D, section D1.1.1, Table 39.	Factual accuracy correction of adverse event rate numbers.	Amendment accepted

Issue 4 Indirect treatment comparison included trials summary demographics and baseline characteristics data

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
In the EAG report section 3.3 <i>Critique of trials identified and included in the indirect comparison and/or multiple treatment</i>	The numbers in this table should be aligned with those in Table 1 of the Sitbon et al. 2015 publication (DOI: 10.1056/NEJMoa1503184).	Factual accuracy correction of demographics and baseline characteristics data of trials included in the	Amendments accepted

comparison, Table 9 <i>Summary demographics and baseline characteristics of the trials included in the indirect treatment comparison</i>, several of the numbers shown to do not match those reported in the source publication (i.e. Sitbon et al. 2015 for the GRIPHON trial data) which these should correspond to. In particular: The GRIPHON trial numbers for “<i>Length of time since PAH diagnosis (years)</i>” (selexipag arm and placebo arm numbers have been switched around). The GRIPHON trial numbers for “<i>Use of medications for PAH, n (%)</i>” (All Patients numbers are shown in the Placebo arm columns).		indirect treatment comparison.	

Issue 5 Spelling of pivotal trial

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
In the EAG report section 1.5, <i>The cost-effectiveness evidence: summary of the EAG's key issues</i> , Issue 5 Timing and approach to PG12 initiation in the sotatercept arm of the model, is inconsistent with clinical expert feedback received by the EAG, the pivotal trial name is spelt incorrectly as "STEALLAR".	The spelling of the pivotal trial name on EAG report page 11 should be replaced with "STELLAR".	Incorrect spelling of pivotal trial.	Amendment accepted

Issue 6 Potentially misleading statement regarding efficacy data used in the economic model

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
In the EAG report section 1.5, <i>The cost-effectiveness evidence: summary of the EAG's key issues</i> , Issue 5 Time and approach to PG12 initiation in the sotatercept arm of the model, is inconsistent with clinical expert feedback received by the EAG, the EAG propose an alternative approach whereby PG12 analogues are given in addition to sotatercept in a proportion of patient who	The company propose that the EAG reconsider this statement on page 11 and corresponding report sections 4.2.2 and 4.2.6, or otherwise remove it, to avoid an incorrect interpretation of the trial data.	Incorrect interpretation of the STELLAR trial data used in the economic model, and EAG alternative approach scenario, that is potentially misleading.	We acknowledge company's point and have edited this statement to highlight the assumptions the EAG are making. By the same token, the intermediate high risk

progress to intermediate high-risk. The report states “This is fact may me more consistent with the efficacy data used to inform the onward transition probabilities from the intermediate-high risk state in the sotatercept arm, which have been based on the full analysis population STEALLAR, a sizable proportion of whom would have been on triple background therapy (including PG12 analogues) in addition to sotatercept.” However, the company believe that this statement is misleading and an incorrect interpretation of the clinical trial data. In the intervention arm, participants in the STELLAR trial received sotatercept in addition to background therapy, of whom some participants were on triple background therapy including PG12 analogues (ie ERA+PDE5i+PCA). In contrast, the scenario the EAG describe in their approach to Issue 5 is the addition of PG12 analogues to patients on ERA+PDE5i+sotatercept, suggesting that the efficacy of adding PG12 analogues to patients on dual background therapy plus sotatercept is equivalent to the addition of sotatercept to patients on triple background therapy including PG12 analogues.			group (at 12 weeks), being used to inform long-term transition probabilities are not primarily a cohort that has progressed from the intermediate-low risk state on sotatercept. The EAG approach assumes that adding PGI2 analogues for patients who have progressed to the intermediate-high risk state on ERA+PDE5i+sotatercept, has similar efficacy to adding sotatercept for those with intermediate-high risk status on triple background therapy (including PG12 analogues).

Issue 7 Spelling of risk-based mortality source

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
In the EAG report section 4.2.6, Treatment effectiveness and extrapolation, on page 64 the source of the risk-based mortality data is spelt incorrectly on two occasions as “Rosenkrantz”.	This statement on page 64, lines 17-20, should be replaced with “<i>Based on the company’s clarification, the EAG are satisfied that the two studies retrieved provide the best available data and agree with the company that data from Rosenkranz, based on the COMPERA data are appropriate to inform mortality in the model.</i>” The statement on page 64, lines 23-25, should be replaced with “<i>Again, using Rosenkranz et al data, Gompertz and Weibull showed monotonically increasing</i>	Incorrect spelling of source for risk-based mortality.	Amendment accepted

	<i>hazards and were preferred based on clinical expert opinion.”</i>		

Issue 8 ‘Off-sotatercept’ health state transition probabilities

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
In the EAG report section 4.2.6 Treatment effectiveness and extrapolation, sub-heading Treatment discontinuation assumptions and data (Page 63), it is stated that “patients who discontinue sotatercept at 24 weeks are assigned to the off-sotatercept health states where they are exposed to transition probabilities derived from the placebo arm of STELLAR” This is true for patients in the intermediate-low ‘off-sotatercept’ health state. However, 40% of patients who discontinue sotatercept at week 24 are in the intermediate-high risk health state where they are assigned transition probabilities from Roman et al. 2012 to reflect that these patients would initiate parenteral prostacyclins in line with the WSPH treatment algorithm.	This text should be reworded to “<i>These patients who discontinue sotatercept at 24 weeks are assigned to the off-sotatercept health states where they are exposed to transition probabilities derived from either the placebo arm of STELLAR (low and intermediate low risk health states) or Roman et al 2012 (intermediate high and high risk health states) (Company submission, Table 84 and 85).</i>”	Factual accuracy correction of the transition probabilities assigned to ‘off-sotatercept’ health states in the company economic model.	Change accepted. We have also slightly edited the subsequent critique section, which overlooked the fact that those discontinuing sotatercept in the intermediate-high risk state do in fact initiate infused PCA.

Issue 9 'Off-sotatercept' health state transition probabilities

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
The EAG report section 4.2.8, subheading Health state specific treatment acquisition and acquisition and administration costs, discusses the EAG preferred approach to calculating drug acquisition costs on treatment for PG12 analogues using the year 1 target dose from the ESC/ERS guidelines. However, in the accompanying model provided by the EAG containing the EAG preferred scenarios, there is a minor error in the calculation. In the tab 'Direct cost inputs' cells V27 and V28 calculate the dose as 'mg/day' using a patient weight of ■■■kg instead of the correct value of ■■■kg as per the patient weight in the model using the non-US population from the STELLAR trial. Albeit minor, this calculation error underestimates the drug acquisition costs for Epoprostenol and Treprostinil.	The company propose that this calculation amendment is made in the EAG base case model (ID6163 Sotatercept in PAH_UK_CEM_CiC_CQs_17Apr2025_EAG_modified_base_case – 220525 [CON]) and EAG base case PSA model (ID6163 Sotatercept in PAH_UK_CEM_CiC_CQs_17Apr2025_EAG_modified_base_case_PSA – 220525 [CON]), and subsequently Table 15 of the EAG report, and all corresponding scenario results, be updated to align with correct calculation output.	Minor inaccuracy in the calculation of drug acquisition costs for PG12 analogues in the updated EAG economic model.	We accept this change and have updated the results for consistency. Please note, however, that the exact ■■■kg value was back calculated from the company's original excel based calculation of prostacyclin costs. The company's PGI2 costs in the model are consistent a body weight of ■■■kg exactly, not ■■■kg.

Issue 10 Approach to application of relative risks for calculating selexipag transition probabilities

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
In the EAG report section 5.3, Model validation and face validity check, (page 88) the text summarises apparent errors or inconsistencies found in the company model. The report states “The calculation of transition probabilities for selexipag, by application of relative risks to the transition probabilities for sotatercept, included adjustment for the probability of staying in state on sotatercept. The EAG were unclear on the reasoning for this, which inferred the relative risks of improvement and deterioration that are less than those estimated from the data.” However, the company methodology for applying to relative risks to derive the transition probabilities for selexipag is neither an error nor an inconsistency. The rationale for this approach being due to the fact that the relative risks for improvement or deterioration are not specific to each risk strata, but rather the general relative risks across all strata. This is accounted for in the proportional redistribution approach taken by the company by first adjusting the	The company propose that this statement be removed for the EAG report section 5.3, along with the corresponding scenario (scenario number 1) in the report section 6.1 (Table 18) and section 6.2 (Table 19).	The company believe that the modelling approach taken for deriving selexipag transition probabilities using relative risks is methodologically robust, and thus the identification of this as an apparent error in the company model is believed to be incorrect.	With the further explanation, the EAG accept the adjustment from the company and remove the relevant sections and analyses from the EAG report. It also required replacing Figure 5, to show updated traces retaining the company approach to relative risk application. The impact on the ICER is small.

sotatercept transitions by the relative risk and then normalizing the results to derive the transitions for each risk strata health state. Conversely, the alternative approach taken by the EAG, decreasing the ICER in favour of sotatercept, is in fact in danger of overinterpreting the relative risks at a state level that could lead to scenarios of implausible, and sometimes negative, selezipag transition probabilities.			

Issue 11 Application of hospitalisation disutility in the company model

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
In the EAG report section 5.3, Model validation and face validity check, (page 88) the text summarises apparent errors or inconsistencies found in the company model. The report states “Utility decrements for hospitalization, and utility increments reflecting changes in carer HRQoL, did not appear to be adjusted for cycle length in the model, leading to overestimation of QALY losses/gains attributable to these elements.”	The company propose that this statement is removed from the EAG report section 5.3, along with the corresponding scenario (scenario number 2) in section 6.1 (Table 18) and section 6.2 (Table 19).	The company believe that the approach to applying hospitalisation utility decrements is methodologically robust, and thus the identification of this as an error in the company model is incorrect.	Not accepted. The hospitalisation disutility is an EQ-5D decrement on the scale full health (1) to death (0), for people with heart failure who have experienced hospitalisation within 30 days. It needs to be multiplied by duration in years for which it

However, the hospitalization utility decrements are considered to be a disutility per event, rather than per year, and are applied in the company model to the number of patients experiencing a new hospitalization event in each 12-week cycle. Therefore, the calculation is already adjusted for cycle length and does not require further adjustment as suggested by the EAG and applied in their additional scenario analysis. The EAG approach in scenario number 2 applies an additional cycle length adjustment which results in an underestimation of the utility decrement associated with hospitalisation. This scenario incorrectly increases the ICER in favour of selexipag.			lasts, to transform it into a QALY decrement for adding to the QALY total. It is being multiplied by the event risk per cycle, but the company implicitly assumes that the utility decrement of each event lasts one year – not the 12 weeks (one cycle) that the submission suggests. No evidence is presented to justify the utility decrement of hospitalisation lasting for the duration of one year. The paper from which it is taken shows a trajectory of improvement with increasing time from hospitalisation. And applying it for longer than a cycle risks double counting the utility impact of progression.

			An appropriate duration adjustment has been made for the IV administration disutility (to generate QALY losses), and for the carer increments (see below) but not for the hospitalisation disutility.

Issue 12 Application of carer HRQoL increments in the company model

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
In the EAG report section 5.3, Model validation and face validity check, (page 88) the text summarises apparent errors or inconsistencies found in the company model. The report states “Utility decrements for hospitalization, and utility increments reflecting changes in carer HRQoL, did not appear to be adjusted for cycle length in the model, leading to overestimation of QALY losses/gains attributable to these elements.”	The company proposes that this statement is removed from the EAG report section 5.3, along with the corresponding scenario (scenario number 3) in section 6.1 (Table 18) and section 6.2 (Table 19).	The company believe that the current approach to applying carer HRQoL increments relative to high-risk health state is methodologically robust and does account for cycle length, thus the identification of this as an error in the company model is incorrect.	We accept this correction, having overlooked the sum product calculations in rows 9 and 10.

However, the company believe the utility increments for carer HRQoL relative to the high risk state is adjusted for in the company model ('QALYs calc' sheet, rows 9 and 10 whereby the SUMPRODUCT accounts for cycle length). Therefore, EAG scenario number 3 applies an additional cycle length adjustment on top of the current cycle length adjustment which results in an underestimation of the carer utility increments in lower risk health states. This scenario incorrectly increases the ICER in favour of selezipag.			

Issue 13 Assumption of model impact if allowing for clinical improvement following initiation of PG12 analogues

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
EAG key issue 4, the simplifying assumption of not allowing for clinical improvement following initiation of PG12 analogues, page 10 and subsequently discussed in report section 4.2.2 and 4.2.6, states "The EAG believe that allowing for risk status improvement in the model would increase the ICER for sotatercept versus selezipag".	The text should be reworded to "The impact on the ICER of sotatercept versus selezipag is unknown where the model to allow for risk status improvement upon initiating PG12 in both arms."	The current statement is potentially misleading and cannot be substantiated.	Accepted but with rewording: "The impact on the ICER of sotatercept versus selezipag is, however, unpredictable, as allowing for improvement may

The EAG hypothesise that allowing for a one-cycle opportunity for improvement of patients initiating parenteral PCA in intermediate-high risk, following discontinuation of selexipag, would increase the ICER. However, it is not possible to implement this scenario in the current economic model and therefore the impact, directionality and severity, on the ICER is unknown and this conclusion cannot be drawn from the current evidence. If patients discontinuing selexipag and initiating PCA in intermediate-high risk have a one-cycle opportunity to improve to intermediate-low, they would remain on an expensive and highly burdensome PCA treatment whilst also remaining in a lower health state for an extended period of time, thus increasing the time spent on PCA analogue treatment. It is unclear whether the additional QALY benefits would outweigh the additional costs. In addition, as discussed in the company response to clarification questions, aside from the structural complexities required to implement this scenario, there would be an outstanding data gap to inform transition probabilities from this new intermediate-			disproportionately increase costs and QALYs in the selexipag arm relative to the sotatercept arm. To counter the potential biases the EAG prefer at least not to apply the utility decrement applied to IV PCA in the intermediate-high and high risk states of the model.”

low risk “on-PCA” health state; further capping the ability to hypothesise on the potential impact. Therefore, it is possible that the company model structure simplification may favour either the selezipag or sotatercept arm, and were this afore mentioned scenario to be feasible the ICER for sotatercept versus selezipag could be impacted in either direction.			

Issue 14 The proportion of patients who receive PG12 analogues

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
EAG Key Issue 6, The proportion of patients who receive PG12 analogues, when indicated according to the modelling assumptions, is too high at 100%, page 12 and subsequently report section 4.2.2 and 4.2.8 states “Clinical expert feedback to the EAG suggests that these treatments are reserved for those who are most likely to tolerate them and therefore benefit from them clinical practice, suggesting 70%-80% may be offered and 80-90% may accept it.”. Furthermore, a scenario,	The company proposes that the EAG reconsider and revise, or remove, this analysis to appropriately reflect the treatment pathway per the clinical guidelines.	The EAG preferred approach to applying PG12 analogues to a proportion of patients in the corresponding intermediate-high risk and high-risk health states, and justification for doing so, lacks validity with respect to clinical guidelines and risks underestimating the true costs of PG12 analogues.	Having consulted with our clinical experts again, they are of the opinion that the 63.75% may be an underestimate of the proportion who will go on to receive infused PG12. However, the EAG clinical experts maintain that these treatments are not

included in the EAG base case, use proportion of 63.75% patients that receive PCA upon escalation to the relevant health state in each arm. The concern with this approach is three-fold. Firstly, the ESC/ERS guidelines state that all patients determined to be of intermediate-high risk should be recommended the addition of PG12 analogues, therefore the only additional factor for consideration should be whether or not a patient can tolerate the PG12 analogues. Applying two separate filters for a patient being offered and a patient accepting the PG12 analogue may underestimate the use of parenteral PCAs. Secondly, the company acknowledges that not all patients may be able to tolerate intravenous PG12 analogues, due the high patient burden; therefore may instead be offered an inhaled alternative. This has been accounted for in the company model whereby the treatment distribution of PG12 analogues is split by IV Epoprostenol (61.94%), IV Treprostinil (4.95%) and inhaled Iloprost (33.11%) in line with the treatment distribution in UK			suitable or acceptable for all/ The EAG therefore amend this to 85% as a plausible (but possibly generous) estimate of the proportion receiving PG12 upon progression. The remaining 15% are assumed to remain on their current therapy (selexipag or sotatercept on top of dual background therapy). The discontinuation of the prior therapy was an oversight in the original EAG scenarios. The relevant sections of the report and results have been updated to reflect the

clinical practice, taken from the National Audit of Pulmonary Hypertension. Finally, in the EAG alternate approach a proportion of 63.75% receiving PG12 analogues is applied to health state where parenteral PCA is initiated in the respective treatment arms. However, in the selexipag arm, in this assumption it is implicit that the additional 36.25% of patients who would either not be offered nor accept PG12 analogues upon progression to intermediate-high risk, and discontinuing selexipag, return to dual background therapy (ERA + PDE5i) with no additional therapy being added on. This lacks clinical validity and is in conflict with the ESC/ERS guidelines and WSPH treatment algorithm whereby patients in the intermediate-high or high-risk strata would be on a triple combination therapy. This assumption is likely to underestimate the costs in these health states.			revised assumptions. The scenario has less impact on the ICER when the prior therapy is appropriately maintained for those who do not receive infused PGI2 analogues. We have also added a scenario that increases the proportion to 100% from the EAG base case. This is an area of uncertainty that would benefit from real-world data and/or further clinical expert opinion, in terms of how the guidelines are used in routine NHS practice.

(please cut and paste further tables as necessary)

Location of incorrect marking	Description of incorrect marking	Amended marking									EAG response
[ID6163] sotatercept – EAG report to NICE_210525 [CON], Page 100 (Table 19)	Results for scenario number 4b have been highlighted but not underlined.	4a	Long-term transition probabilities from intermediate-high to high risks state on sotatercept	Based on observed 12–24-week transitions in the sotatercept arm of the STELLAR full analysis set	Increase by 50%	██████	██████	2.42	██████	██████	Amended as suggested (now 2b)
		4b			Increase by 100%	██████	██████	2.32	██████	██████	

(Please add further lines to the table as necessary)