NICE National Institute for Health and Care Excellence

# Single Technology Appraisal

# Zanubrutinib for treating Waldenstrom's macroglobulinaemia [ID1427]

# Committee Papers

NICE National Institute for Health and Care Excellence

**NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE**

**SINGLE TECHNOLOGY APPRAISAL**

# Zanubrutinib for treating Waldenstrom's macroglobulinaemia [ID1427]

**Contents:**

The following documents are made available to consultees and commentators:

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

1. **Company submission** from BeiGene
2. **Clarification questions and company responses**
3. **Patient group, professional group and NHS organisation submissions** from:
   a. British Society Haematology and Royal College of Pathologists
   b. WMUK and Lymphoma Action
4. **Evidence Review Group report** prepared by Kleijnen Systematic Reviews
5. **Evidence Review Group report – factual accuracy check**
6. **Technical engagement response from company**
7. **Technical engagement responses and statements from experts:**
   a. Dr Dima El-Sharkawi, Haematology Consultant – clinical expert, nominated by the Royal College of Pathologists
   b. Ronald Presswood – patient expert, nominated by WMUK
   c. Dr Shirley D'Sa, Consultant Haematologist – clinical expert, nominated by WMUK
   d. Jane Nicholson – patient expert, nominated by WMUK (**see item 3b*)
8. **Technical engagement responses from consultees and commentators:**
   a. Janssen-Cilag
9. **Evidence Review Group critique of company response to technical engagement** prepared by Kleijnen Systematic Reviews
   a. Revised ERG analyses with new PAS

*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

# Zanubrutinib for treating Waldenström's macroglobulinaemia [ID1427]

## Document B

## Company evidence submission

**March 2021**

| File name | Version | Contains confidential information | Date |
|---|---|---|---|
| ID1427 Zanubrutinib Document B 20Jan2022 | 1.1 | Yes | 20 January 2022 |

# Contents

# List of tables

# List of figures

# Abbreviations

| | |
|---|---|
| Admin | administration |
| AE | adverse event |
| AESI | adverse event of special interest |
| AIC | Akaike information criteria |
| ALT | alanine transaminase |
| ASCT | autologous stem cell transplantation |
| AST | aspartate aminotransferase |
| BCR | B-cell antigen receptor |
| BCSH | British Committee for Standards in Haematology |
| BDR | bortezomib, rituximab and dexamethasone |
| BIC | Bayesian information criteria |
| BID | twice daily |
| BLNK | B cell linker |
| BNF | British National Formulary |
| BR | rituximab and bendamustine |
| BSA | body surface area |
| BSC | best supportive care |
| BTK | Bruton's tyrosine kinase |
| $Ca^{2+}$ | calcium |
| CD19 | cluster of differentiation 19 |
| CE | Conformité Européenne (European Conformity) |
| CHMP | Committee for Medicinal Products for Human Use |
| CHOP | cyclophosphamide, doxorubicin, vincristine and prednisone |
| CI | confidence interval |
| Clad-R | cladribine and rituximab |
| CLL | chronic lymphocytic leukaemia |
| CMH | Cochran-Mantel-Haenszel |
| CR | complete response |
| CT | computerised tomography |
| CVP | cyclophosphamide, vincristine and prednisolone |
| *CXCR4* | C-X-C motif chemokine receptor 4 |
| CYP | cytochrome P450 |
| CYP3A | cytochrome P4503A |
| DAG | 1,2 di-acyl glycerol |
| DAPS | directly accessed pathology services |
| DOR | duration of response |
| DRC | dexamethasone, rituximab and cyclophosphamide |
| DSA | deterministic sensitivity analysis |
| DSC | Decision Support Unit |
| ECOG PS | Eastern Cooperative Oncology Group Performance Status |
| EMA | European Medicines Agency |
| EORTC QLQ-C30 | European Organisation for Research and Treatment of Cancer |
| EQ-5D-3L | EuroQol 5-Dimensions 3-Level |
| EQ-5D-5L | EuroQol 5-Dimensions 5-Level |
| ERG | Evidence Review Group |
| ESHAP | etoposide, solu-medrone, cytarabine, cisplatin |
| ESMO | European Society for Medical Oncology |
| FAS | full analysis set |
| FCR | fludarabine, cyclophosphamide and rituximab |
| FR | fludarabine and rituximab |
| HR | hazard ratio |
| HRQoL | health-related quality of life |
| HRU | healthcare resource use |

| | |
|---|---|
| HTA | health technology assessment |
| IBR | ibrutinib |
| ICER | incremental cost-effectiveness ratios |
| IDARAM | idarubicin, methotrexate, cytarabine and dexamethasone |
| Ig | immunoglobulin |
| IKK | I kappa B kinase |
| Inv | investigator |
| IPD | individual patient-level data |
| IPSSWM | International Prognostic Scoring System for Waldenström's Macroglobulinaemia |
| IQR | interquartile range |
| IRC | independent review committee |
| ITC | indirect treatment comparison |
| ITT | intention-to-treat |
| IV | intravenous |
| IWWM | International Workshop on Waldenström's Macroglobulinemia |
| IWWM-6 | Sixth International Workshop on Waldenström's Macroglobulinemia |
| IWWM-7 | Seventh International Workshop on Waldenström's Macroglobulinemia |
| KM | Kaplan-Meier |
| LMM | linear mixed effects model |
| LPL | lymphoplasmacytic lymphoma |
| LS | least square |
| LY | life year |
| LYN | LYN proto-oncogene |
| MAA | marketing authorisation application |
| Max | maximum |
| MAIC | matching adjusted indirect comparisons |
| MCL | mantle cell lymphoma |
| MHRA | Medicines and Healthcare Products Regulatory Agency |
| Min | minimum |
| MRR | major response rate |
| *MYD88* | myeloid differentiation primary response gene 88 |
| *MYD88*$^{MUT}$ | myeloid differentiation primary response gene 88 mutant |
| *MYD88*$^{WT}$ | wild-type myeloid differentiation primary response gene 88 |
| n | number of patients in the category |
| N | number of patients evaluable |
| N/A | not applicable |
| NE | not evaluable |
| $n_{eff}$ | effective sample size |
| NFκB | nuclear factor kappa B |
| NFAT | nuclear factor of activated T cells |
| NHL | non-Hodgkin's lymphoma |
| NHS | National Health Service |
| NICE | National Institute of Health and Care Excellence |
| No. | number |
| NR | not reported |
| NYHA | New York Heart Association |
| OD | once daily |
| OR | overall response |
| ORR | objective response rate |
| OS | overall survival |
| PD | progressive disease |
| PFS | progression-free survival |
| PH | proportional hazard |

| | |
|---|---|
| Pl3K | phosphatidylinositol-4,5-bisphosphate 3-kinase |
| PIP3 | phosphatidylinositol (3,4,5)-trisphosphate |
| PKC | protein kinase C |
| PLC | phospholipase C |
| PN | peripheral neuropathy |
| PP | per-protocol |
| PPS | post-progression survival |
| PR | partial response |
| PRO | patient-reported outcome |
| PSM | partitioned survival model |
| PSSUR | Personal Social Services Research Unit |
| Pt | patient |
| PT | Preferred Term |
| QALY | quality-adjusted life year |
| QLQ-C30 | quality of Life Questionnaire core-30 |
| QoL | quality of life |
| QTcF | T interval corrected for heart rate using Fridericia's formula |
| R | rituximab or Randomised |
| R-CHOP | rituximab, cyclophosphamide, doxorubicin, vincristine and prednisone |
| R-ESHAP | rituximab, etoposide, solu-medrone, cytarabine and cisplatin |
| R-IDARAM | rituximab, idarubicin, methotrexate, cytarabine and dexamethasone |
| R/R | relapsed/refractory |
| RAP | rapGTP-binding protein |
| RCT | randomised controlled trial |
| SAE | serious adverse event |
| SD | standard deviation |
| SE | standard error |
| SLR | systematic literature review |
| SmPC | summary of product characteristics |
| SMQ | Standardized Medical Dictionary for Regulatory Activities Query |
| SOC | System Organ Class |
| SPEP | serum protein electrophoresis |
| SYK | spleen tyrosine kinase |
| TA | technology appraisal |
| TN | treatment naïve |
| TRAE | treatment-related adverse event |
| TTD | time to discontinuation |
| UK | United Kingdom |
| ULN | upper limit of normal |
| Unk | unknown |
| US | United States |
| VGPR | very good partial response rate |
| VR | bortezomib and rituximab |
| vs | versus |
| WHIM | warts, hypogammaglobulinemia, infections, myelokathexis |
| WM | Waldenström's macroglobulinaemia |
| WMUK | Waldenström's Macroglobulinaemia United Kingdom |
| WT | wild type |
| WTP | willingness-to-pay |
| ZANU | zanubrutinib |

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

## B.1.1 Decision problem

The submission covers the technology's full marketing authorisation for this indication (treatment of adult patients with Waldenström's macroglobulinaemia [WM] who have received at least one prior therapy, or first-line treatment for patients unsuitable for chemo-immunotherapy). A summary of the decision problem is provided in Table B.1.1.

**Table B.1.1. 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 WM:<br>• who have had at least 1 prior therapy, or<br>• whose disease is untreated, for whom chemo-immunotherapy is unsuitable | As per scope | N/A |
| Intervention | Zanubrutinib | As per scope | N/A |
| Comparator(s) | Treatment without zanubrutinib:<br>• For people who have had at least one prior therapy:<br>○ BR<br>○ DRC<br>○ FR<br>○ FCR<br>○ Clad-R<br>○ ASCT in people for whom ASCT is suitable<br>• For people for whom chemo-immunotherapy is unsuitable:<br>○ chlorambucil<br>○ rituximab monotherapy<br>○ BSC including blood product transfusions, plasma exchange, granulocyte stimulating factors and intravenous Ig infusions | Treatment without zanubrutinib:<br>• BR<br>• DRC<br>• Ibrutinib | Other than BR and DRC, it was not possible to conduct comparisons with chemotherapy regimens or BSC, due to a lack of data in the literature to enable comparison of zanubrutinib with the comparators of interest (see Appendix D).<br><br>However, BR and DRC currently represent the two most common regimens for the first-line treatment of WM in patients considered fit enough to tolerate them (13.1% and 16.2%, respectively [see Section B.1.3.5.2]). In addition, BR and DRC are the third- and second- most common second-line regimens, respectively, behind ibrutinib (18.2%).[1]<br><br>Ibrutinib is also included as a comparator, given that:<br>• Registry data indicates that BTK inhibitors (currently only ibrutinib is available) are an emerging standard of care in patients who have had ≥1 prior therapy, with ibrutinib being the most frequently used treatment in |

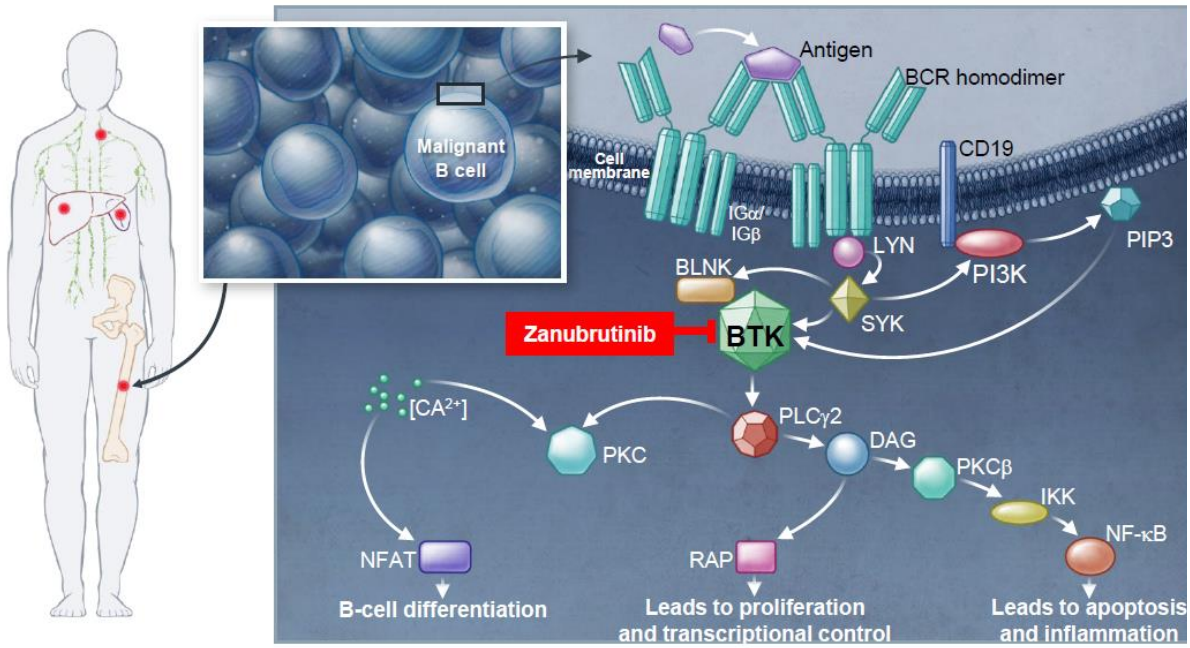
| | | | |
|---|---|---|---|
| | | | clinical practice (approximately 18.2% of cases).[1]<br>• Ibrutinib is the only comparator for which direct head-to-head evidence is available – the safety and efficacy of zanubrutinib versus ibrutinib were evaluated in the largest Phase 3 trial of BTK inhibitors in WM (BGB-3111-302 [ASPEN]),[2] which forms the primary source of clinical evidence for this submission<br>• Although ibrutinib is currently recommended for use in the CDF, the data collection arrangement for ibrutinib was anticipated to conclude in September 2020,[3] and NICE is subsequently due to update the guidance for ibrutinib in WM |
| Outcomes | • OS<br>• PFS<br>• Response rates (ORR, MRR, VGPR/CR)<br>• Time to next treatment<br>• Duration of response/remission<br>• Adverse effects of treatment<br>• HRQoL | • Response rates (ORR, MRR, VGPR/CR)<br>• Duration of response<br>• PFS<br>• OS<br>• Time to next treatment<br>• HRQoL<br>• Adverse effects of treatment | N/A |

Abbreviations: ASCT = autologous stem cell transplantation; BR = bendamustine and rituximab; BSC = best supportive care; CDF = Cancer Drugs Fund; CHOP = cyclophosphamide, doxorubicin, vincristine and prednisone; Clad-R = cladribine and rituximab; CR = complete response; DRC = dexamethasone, rituximab and cyclophosphamide; FCR = fludarabine, cyclophosphamide and rituximab; FR = fludarabine and rituximab; HRQoL = health-related quality of life; Ig = immunoglobulin; N/A = not applicable; NICE = National Institute for Health and Care Excellence; MRR = major response rate; ORR = objective response rate; OS = overall survival; PFS = progression-free survival; VGPR = very good partial response rate; WM = Waldenström's macroglobulinaemia

## B.1.2 Description of the technology being appraised

Zanubrutinib is an orally administered, highly selective, small-molecule inhibitor of Bruton's tyrosine kinase (BTK).[2, 4] Zanubrutinib forms a covalent bond with a cysteine residue in the BTK active site, leading to inhibition of BTK activity. BTK is a signalling molecule of the B-cell antigen receptor (BCR) and cytokine receptor pathways. In B cells, BTK signalling results in activation of pathways necessary for B-cell proliferation, trafficking, chemotaxis, and adhesion. In nonclinical studies, zanubrutinib inhibited malignant B-cell proliferation and reduced tumour growth.[4] A summary of the mechanism of action of zanubrutinib is presented in Figure B.1.1.

**Figure B.1.1. Mechanism of action of zanubrutinib**




Abbreviations: BCR = B-cell antigen receptor; BLNK = B cell linker; BTK = Bruton's tyrosine kinase; CA2+ = calcium; CD19 = cluster of differentiation 19; DAG = 1,2 di-acyl glycerol; IKK = I kappa B kinase; LYN = LYNproto-oncogene; Src family tyrosine kinase; NFκB = nuclear factor kappa B; NFAT = nuclear factor of activated T cells; PI3K = phosphatidylinositol-4,5-bisphosphate 3-kinase; PIP3 = phosphatidylinositol (3,4,5)-trisphosphate; PKC = protein kinase C; PLC = phospholipase C; RAP = RapGTP-binding protein also known as Ras-relatedprotein; SYK=spleen tyrosine kinase
Source: Hendricks et al., 2011[5]

Zanubrutinib is specific and selective for BTK and was designed to minimise off-target inhibition of other kinases. As such, it has the potential to improve outcomes and reduce side effects compared with existing therapies for WM.[6, 7] A summary of zanubrutinib is provided in Table B.1.2 and the summary of product characteristics (SmPC) is included in Appendix C.

**Table B.1.2. Technology being appraised**

| | |
|---|---|
| **UK approved name and brand name** | Zanubrutinib (Brukinsa®) |
| **Mechanism of action** | Zanubrutinib is a small-molecule inhibitor of BTK. Zanubrutinib forms a covalent bond with a cysteine residue in the BTK active site, leading to inhibition of BTK activity. BTK is a signalling molecule of the BCR and cytokine receptor pathways. In B cells, BTK signalling results in activation of pathways necessary for B-cell proliferation, trafficking, chemotaxis, and adhesion. In nonclinical studies, zanubrutinib inhibited malignant B-cell proliferation and reduced tumour growth.[4] |
| **Marketing authorisation/CE mark status** | Zanubrutinib was reviewed by the EMA for patients with WM who have received at least one prior therapy or are unsuitable for chemoimmunotherapy. The application was submitted in May 2020, with CHMP positive opinion granted in September 2021.<br><br>Following CHMP opinion, an MAA was submitted to MHRA via the Reliance Route and UK approval was granted in December 2021. |
| **Indications and any restriction(s) as described in the SmPC** | Proposed indication:<br>Zanubrutinib as a single agent is indicated for the treatment of adult patients with WM who have received at least one prior therapy, or in first line treatment for patients unsuitable for chemo-immunotherapy.[4] |
| **Method of administration and dosage** | The recommended daily dose of zanubrutinib is 320 mg, taken orally either OD (four 80 mg capsules) or BID (two 80 mg capsules).[4] |
| **Additional tests or investigations** | N/A |
| **List price and average cost of a course of treatment** | £4,928.65 per pack of 120 80 mg capsules. |
| **Patient access scheme (if applicable)** | N/A |

Abbreviations: BID = twice daily; BCR = B-cell antigen receptor; BTK = Bruton's tyrosine kinase; CE = Conformité Européenne (European Conformity); CHMP = Committee for Medicinal Products for Human Use; EMA = European Medicines Agency; MAA = marketing authorisation application; MHRA = Medicines and Healthcare Products Regulatory Agency; N/A = not applicable; OD = once daily; SmPC = summary of product characteristics; WM = Waldenström's macroglobulinaemia; UK = United Kingdom

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

### B.1.3.1 Disease overview

WM is a rare, heterogeneous, incurable lymphoplasmacytic lymphoma (LPL),[8, 9] characterised by overproduction and bone-marrow infiltration of monoclonal immunoglobulin M (IgM)-secreting lymphoplasmacytic cells.[10]

The L265P point mutation in myeloid differentiation primary response gene 88, *MYD88*$^{MUT}$, has been identified in more than 90% of patients with WM, making it a useful diagnostic indicator.[11, 12] *MYD88*$^{MUT}$ leads to activation of the transcription factor nuclear factor kappaB (NF-κB) via BTK, and is thought to be the main driver of BTK activation resulting in increased cell survival and proliferation.[13, 14] Other *MYD88* mutations or a wild-type *MYD88* gene (*MYD88*$^{WT}$) may be found in 5–10% of patients with WM[12] and non-L265P mutation variants

(V217F, S219C M232T and S243N) have also been identified in 1–2% of patients with WM.[15, 16] Additionally, a truncated C-X-C motif chemokine receptor 4 (*CXCR4*) mutation is found in approximately 30% of WM patients[11]—mutations in the *CXCR4* gene correlate with a more active disease course, resistance to ibrutinib therapy, and hyperviscosity syndrome.[14] Patients with both *MYD88* and *CXCR4* mutations have worse outcomes following ibrutinib treatment than patients with the *MYD88* mutation alone, exhibiting a delay in attaining a major response and fewer major responses or very good partial responses (VGPRs).[17] Furthermore, the presence of a *CXCR4* mutation has been associated with earlier disease progression in patient with *MYD88*$^{MUT}$ treated with ibrutinib alone or with ibrutinib and rituximab when compared with patients with mutated *MYD88* alone.[17]

### B.1.3.2 Epidemiology

WM is a rare disease, representing 1–2% of all cases of non-Hodgkin's lymphoma.[12] The age-standardised annual incidence of WM in England is 0.55 per 100,000 people, which with a national population of 56.3 million[18] translates to approximately 310 new cases of WM each year.[19] This is consistent with the 353 newly diagnosed cases of WM registered in England in 2017.[20]

WM is most prevalent in men,[1] with age-adjusted incidence rates of 7.3 and 4.2 per million among males and females, respectively.[12] WM is also most prevalent in the elderly,[12] with incidence rates increasing dramatically with increasing age;[13, 14, 21] there is a 95-fold higher incidence of WM in those aged 80–90 years than those aged <50 years[13] and median age at WM diagnosis is 71 years.[22]

### B.1.3.3 Symptomology and presentation

WM is an indolent disease;[9] approximately 25% of patients are asymptomatic at diagnosis.[13] However, patients that present with symptoms report non-specific constitutional symptoms such as fatigue and B symptoms (weight loss, fever and night sweats).[13] As the disease progresses, symptoms reflect infiltration of haematopoietic tissue by LPL and IgM paraprotein deposition and autoimmune activity.[13] Clinical features may include cytopenia (i.e. thrombocytopenia, anaemia, leukopenia) and peripheral neuropathy.[13, 14] Anaemia can occur as IgM concentration increases, leading to increased oncotic pressure and ultimately increased plasma volume; blood loss and iron deficiency secondary to mucosal bleeding or decreased iron absorption can also lead to anaemia.[23] Although involvement of extramedullary sites is rare, the most commonly involved of these sites are the lungs, soft tissues, central nervous system, kidneys and bones.[14]

Upon laboratory evaluation, IgM paraprotein levels typically range from 0.1 to >8.0 g/dL with hyperviscosity seen in patients with IgM levels >6.0 g/dL.[14] Approximately 30% of patients with WM present with hyperviscosity syndrome, which can manifest as neurological symptoms and mucosal bleeding due to the raised IgM interfering with haemostasis.[23] Other reported signs and symptoms related to increased IgM concentrations include coagulation disorders, cryoglobulinemia, neuropathy and cold agglutinin haemolytic anaemia.[14]

### B.1.3.4 Burden to patients, carers and society

WM is an incurable disease with a median OS of 18.5 years in symptomatic patients.[1] In an analysis of UK registry data from 671 patients with WM, 118 patients (18%) died between 1978 and 2019, equating to a 5-year OS of 90.5% and 10-year OS of 79.4%. Patients in a

higher International Prognostic Scoring System for WM (IPSSWM) risk category had a significantly reduced 5-year OS rate than those in a lower IPSSWM risk category (p<0.005); 92% for the low risk group, 79% for the intermediate group and 38% for the high risk group.[24]

Survival in WM is impacted by sex and age at diagnosis; median OS is lower in women than in men (19.4 years versus 29.5 years) and is also lower in patients diagnosed at >65 years than those diagnosed at <65 years (14.6 years versus 29.5 years, respectively).[1] In addition to age, anaemia is also considered an important prognostic factor. Haemoglobin levels of <9 to 12 g/dL are proposed to decrease survival rates. Additionally, high serum $\beta_2$-microglobulin (>3 to 3.5 mg/L), IgM (>7 g/dL), and low platelet count (<100,000/µL) are associated with poor prognosis.[25] Prognostic factors for WM are outlined in Table B.1.3.

**Table B.1.3. Prognostic factors for WM**

| Factors associated with prognosis | Value |
|---|---|
| Age, years | >65 |
| Haemoglobin, g/dL | ≤11.5 |
| Platelet count, N/µL | ≤100,000 |
| $\beta_2$-microglobulin, mg/L | >3 |
| Monoclonal IgM, g/dL | >7 |

Abbreviations: IgM = immunoglobulin M; N = number of platelets; WM = Waldenström's macroglobulinaemia
Source: Morel et al., 2009[25]

The IPSSWM uses these criteria to assign a score and a corresponding risk group (low, intermediate and high). High-risk patients have three or more prognostic factors and a 5-year overall survival of 36%; intermediate-risk patients have two prognostic factors or age >65 and a 5-year overall survival of 68%; low-risk patients have one or no prognostic factors (excluding age >65) and a 5-year overall survival of 87% (Table B.1.4).[25]

**Table B.1.4. IPSSWM staging criteria**

| Risk group | Risk factors present | 5-year OS, % |
|---|---|---|
| Low risk | 0 or 1 (except age) | 87 |
| Intermediate risk | Age or 2 | 68 |
| High risk | ≥3 | 36 |

Abbreviations: IPSSWM = International Prognostic Scoring System for Waldenström's macroglobulinaemia; OS = overall survival
Source: Morel et al., 2009[25]

Although WM is an indolent disease, it is characterised by symptomatic disease recurrences that have a detrimental impact on quality of life (QoL) and daily living.[26] Symptoms of WM, including recurring infections due to leukopenia, or fatigue and weakness due to anaemia, have been shown to impact QoL.[26, 27] Complications of WM can include renal disease, peripheral neuropathy (PN), histological transformation and secondary malignancies; WM may require prompt treatment to avoid irreparable organ damage or fatal complications, such as in the case of hyperviscosity syndrome.[9] PN has been shown to significantly impact QoL; in a study of patients with WM, PN was associated with reduced cognitive function (p=0.0031) and a greater perception of anxiety (p=0.0015).[28] Additionally, an analysis of patient-reported outcomes collected within the WMUK Rory Morrison Registry up to 2018 (a registry with a total of 579 WM patients registered from 19 hospitals across the UK) showed that approximately 10–20% of patients, regardless of when diagnosed, were experiencing anxiety.[1] The indolent disease course and symptoms associated with WM can lead to fear of relapse

following treatment,[29] a common factor that impacts the sense of well-being of patient with WM.[30]

Studies on the economic burden of WM are limited.[31] However, current treatments for WM have a considerable economic impact. Long-term use of treatments that do not target disease specific abnormalities (e.g. prolonged chemotherapy treatment) can lead to serious adverse events (SAEs) in patients with WM,[32] potentially resulting in increased healthcare expenditure.[33] In England, the National Health Service (NHS) resource use associated with WM (primary diagnosis) was 5,384 hospital episodes and 2,609 bed days in 2018-19.[34]

### B.1.3.5 Clinical pathway of care

#### *B.1.3.5.1 Diagnostic pathway*

The diagnosis of WM is based on the histopathological confirmation of bone marrow infiltration by LPL and the detection of any amount of monoclonal IgM protein, confirmed by immunofixation. Identification of *MYD88*$^{MUT}$ status can be helpful for differential diagnosis from other morphologically similar diseases. However, *MYD88*$^{MUT}$ alone is not diagnostic of WM.[12]

European Society for Medical Oncology (ESMO) clinical guidelines recommend:

- Review of familial history for WM and other B cell lymphoproliferative disorders
- Review of symptoms (B symptoms, organomegaly, hyperviscosity symptoms, neuropathy, Raynaud's disease, peripheral oedema, skin abnormalities, dyspnoea)
- Fundoscopic examination if IgM is high and hyperviscosity is suspected
- Laboratory testing to ascertain complete blood count, complete metabolic panel, serum Ig levels (IgA, IgG and IgM), serum and urine electrophoresis with immunofixation, serum $\beta_2$-microglobulin level and viral serology (hepatitis B, hepatitis C, human immunodeficiency virus)
- Bone marrow aspiration and biopsy to ascertain immunohistochemistry and *MYD88*$^{MUT}$ status
- Computerised tomography (CT) of the chest, abdomen and pelvis[12]

#### *B.1.3.5.2 Treatment pathway*

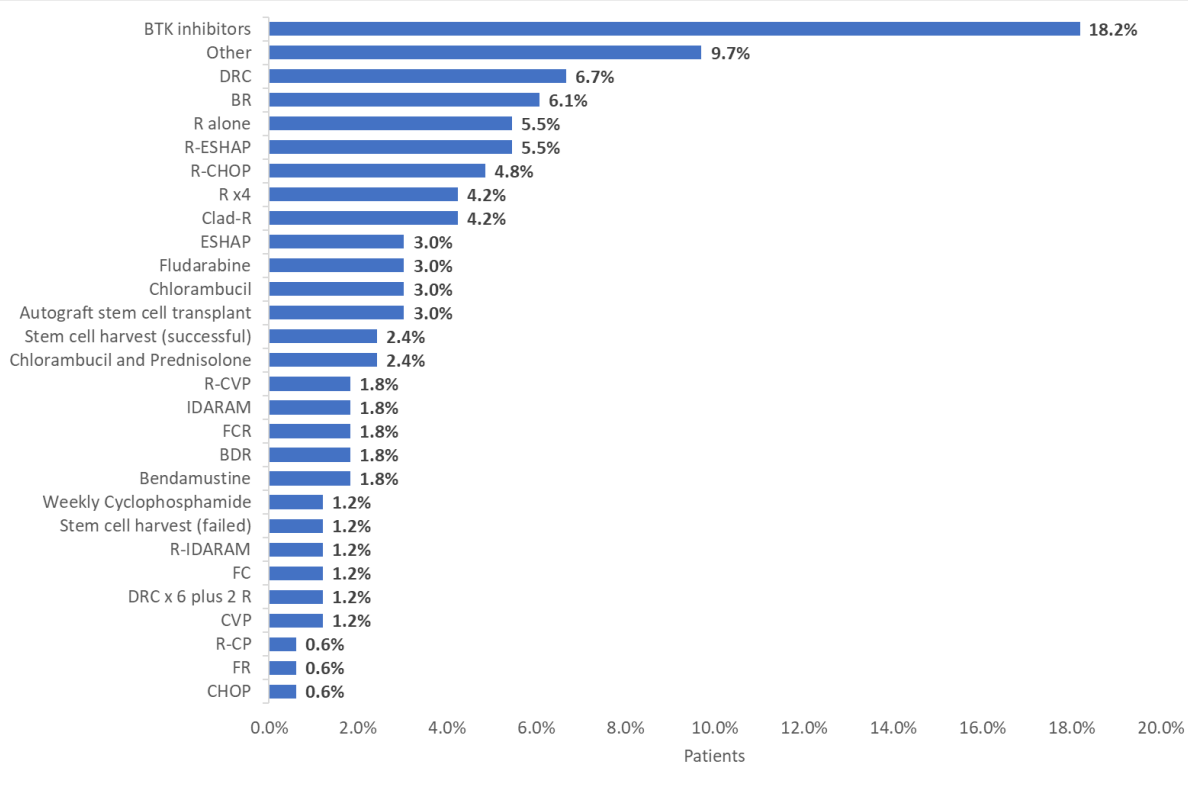
At present there is no established standard of care for WM in England.[1] The British Committee for Standards in Haematology (BCSH) issued guidelines on the diagnosis and management of WM in 2014 that recommend patients with symptomatic WM should receive a rituximab-containing regimen, including dexamethasone + rituximab + cyclophosphamide (DRC), bendamustine + rituximab (BR), fludarabine + rituximab (FR), fludarabine + cyclophosphamide + rituximab (FCR) and cladribine + rituximab (Clad-R).[19] Chlorambucil is also recommended in those not suitable for chemotherapy.[19]

When diagnosed with WM, most patients do not require immediate treatment and are monitored in clinic on a *watch and wait* or *active surveillance* approach, the length of which can vary. Patients are monitored for symptoms of WM and once these develop, treatment is initiated.[1]

There is significant variability in the treatments prescribed for WM in the UK. In the Rory Morrison Registry 2018 report, the two most common first-line regimens for patients considered fit enough to tolerate them were DRC (n=51/314, 16.2%) and BR (n=41/314, 13.1%).[1] The registry also demonstrates the heterogeneity in second-line regimens used in

the UK (Figure B.1.2). The most frequently used second-line treatment was a BTK inhibitor (i.e. ibrutinib) in 18.2% of cases, followed by DRC (6.7%) and BR (6.1%).[1] The development of ibrutinib, a first in class BTK inhibitor available through the Cancer Drugs Fund (CDF), has marked a paradigm shift in treatment of WM and has rapidly become standard treatment for patients with relapsed/refractory WM.[1] However, as ibrutinib became available in 2017, it is not included in the current BSCH guidelines for WM which were developed in 2014.[19] Therefore, the ESMO 2018 guidelines provide the most up-to-date recommendations for the treatment of WM.[12]

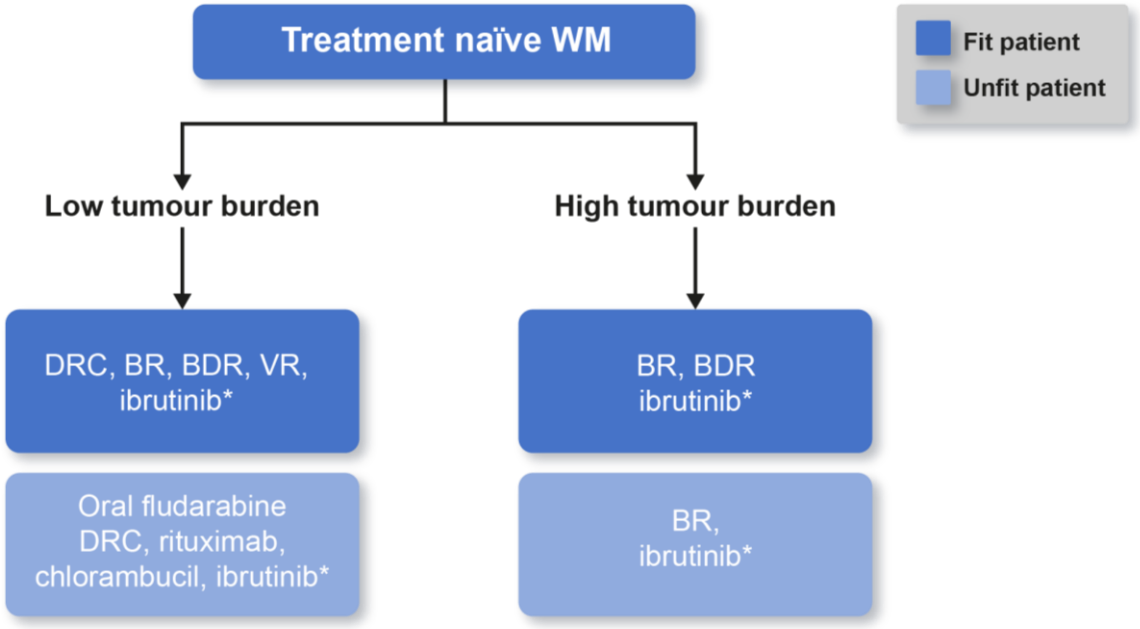
**Figure B.1.2. Patients with WM undergoing second-line treatment**




Abbreviations: BDR = bortezomib, dexamethasone and rituximab; BR = bendamustine and rituximab; BTK = Bruton's tyrosine kinase; Clad-R = cladribine and rituximab; CVP = cyclophosphamide, vincristine and prednisolone; DRC = dexamethasone, rituximab and cyclophosphamide; F(C)(R) = fludarabine, (cyclophosphamide), (rituximab); R = rituximab; (R)-CHOP = (rituximab), cyclophosphamide, doxorubicin, vincristine and prednisolone; (R)-C(V)P = (rituximab), cyclophosphamide, (vincristine) and prednisolone; (R)-ESHAP = (rituximab), etoposide, solu-medrone, cytarabine and cisplatin; (R)-IDARAM = (rituximab), idarubicin, methotrexate, cytarabine and dexamethasone; VR = bortezomib and rituximab; WM = Waldenström's macroglobulinaemia
Source: WMUK, 2018[1]

ESMO guidelines advise that choice of therapy is highly personalised and should be determined by the patient's age, fitness, *MYD88*$^{MUT}$ status, prior therapies and existing comorbidities.[12] Recommended treatment options for treatment-naïve patients include chemo-immunotherapy (e.g. rituximab in combination with alkylating agents) or rituximab monotherapy (Figure B.1.3).

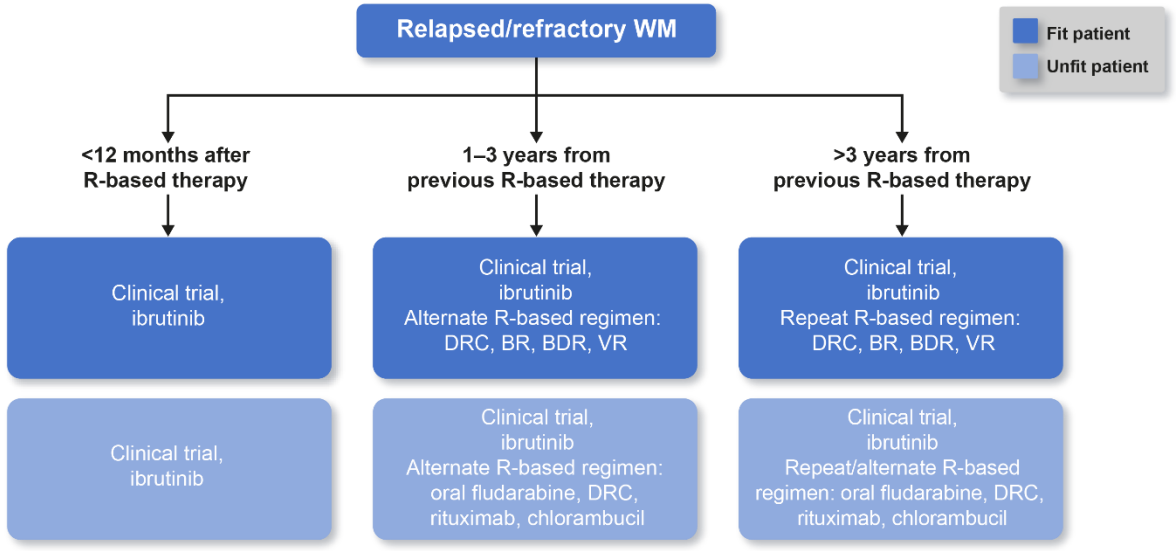
**Figure B.1.3. Treatment pathway for treatment-naïve patients with WM**




*In treatment-naïve patients who are not suitable for chemotherapy
Abbreviations: BDR = bortezomib, dexamethasone and rituximab; BR = bendamustine and rituximab; DRC = dexamethasone, rituximab and cyclophosphamide; VR = bortezomib and rituximab; WM = Waldenström's macroglobulinaemia
Source: Kastritis et al., 2018[12]; NICE, 2017[35]

As WM is incurable, almost all patients will eventually relapse;[9] in relapsed/refractory WM, options include ibrutinib and alternative/repeat rituximab-based regimens, such as BR, DRC, bortezomib, dexamethasone and rituximab, and bortezomib and rituximab (Figure B.1.4).[12]

**Figure B.1.4. Treatment pathway for relapsed/refractory patients with WM**




Abbreviations: BDR = bortezomib, dexamethasone and rituximab; BR = bendamustine and rituximab; DRC = dexamethasone, rituximab and cyclophosphamide; R = rituximab; VR = bortezomib and rituximab; WM = Waldenström's macroglobulinaemia
Source: Kastritis et al., 2018[12]; NICE, 2017[35]

### B.1.3.6 Unmet need

Effective treatment options for WM are limited across all lines of therapy.[16] No established treatment approach for WM has curative potential[36] with some patients requiring >8 lines of therapy.[1] Once chemo-immunotherapy options are exhausted (e.g. rituximab combinations such as BR and DRC), there is an unmet need for additional treatments for relapsed/refractory patients. For treatment-naïve patients with WM unsuitable for chemo-immunotherapy, options are limited to best supportive care and ibrutinib.[12]

Chemo-immunotherapy options for WM, such as BR and DRC, have poor tolerability and have shown reduced OS and PFS in real-world studies compared with clinical trials.[37-39] These rituximab-based regimens were originally developed for other diseases and are used off-label to manage disease symptoms rather than targeting WM-specific pathways. Off-label treatment with traditional agents is associated with high levels of toxicity and poor tolerability. Rituximab is known to induce IgM flares. and rituximab combinations, such as BR and DRC, are associated with myelosuppression and cytopenias.[40]

Ibrutinib, the only current treatment specifically licensed for WM, has shown discontinuation rates of up to 22% of patients within one year of initiation.[41] In addition, 73% of patients experienced an IgM rebound (≥25% increase in serum IgM) following discontinuation of ibrutinib, with almost half of those occurring within 4 weeks of discontinuation.[41] IgM rebound led to reduced response to salvage therapy compared with patients without IgM rebound, and patients who did not respond to salvage therapy had an increased risk of death following ibrutinib discontinuation.[41] Additionally, use of ibrutinib is associated with treatment limiting adverse events (AEs) such as bleeding and atrial fibrillation, which requires additional treatment and strict monitoring by a cardiologist.[42] There is a particular unmet need in patients with *MYD88*$^{WT}$ WM, where ibrutinib has been found to demonstrate a shorter median survival and a lower probability of response than in those with *MYD88*$^{MUT}$.[16]

In conclusion, as WM is largely a disease of the elderly[25] there is a need for new treatment options that are well tolerated and suitable for those who are immunosuppressed or who have considerable comorbidities. New BTK inhibitors with improved pharmacological properties resulting in sustained disease control and with greater selectivity are required, leading to a superior safety profile to ibrutinib in the treatment of WM.

### B.1.3.7 Place of zanubrutinib in the treatment pathway

Zanubrutinib is a new, effective treatment choice for those with WM who are unsuitable for chemo-immunotherapy or with relapsed/refractory disease, irrespective of *MYD88* status.[2, 43] Zanubrutinib is expected to provide an additional treatment option alongside ibrutinib in the WM arm of the non-Hodgkin's lymphoma National Institute of Health and Care Excellence (NICE) pathway. [redacted]

Zanubrutinib represents a much-needed alternative to ibrutinib and offers sustained efficacy and improvement in QoL from baseline with a more favourable safety and tolerability profile, as well as a lower rate of discontinuation due to AEs compared with ibrutinib.[2]

## B.1.4 Equality considerations

There are no known equality issues relating to the use of zanubrutinib in patients with WM.

# B.2. Clinical effectiveness

## B.2.1 Identification and selection of relevant studies

A systematic literature review (SLR) was conducted to identify and summarise the available randomised controlled trial (RCT) evidence for the current and future treatment options for adults with Waldenström's macroglobulinaemia (WM) who have had at least one prior therapy, or whose disease is untreated, for whom chemo-immunotherapy is unsuitable. Full details of the methodology and the results of the SLR are detailed in Appendix D.

## B.2.2 List of relevant clinical effectiveness evidence

The efficacy of zanubrutinib has been evaluated in the pivotal Phase 3 BGB-3111-302 study (ASPEN; NCT03053440).[2] An overview of ASPEN is provided in Table B.2.1.

**Table B.2.1. Clinical effectiveness evidence**

| | | | | | |
|---|---|---|---|---|---|
| **Phase 3** | | | | | |
| **Study** | Study BGB-3111-302 (ASPEN; NCT03053440)[2] | | | | |
| **Study design** | Multicentre, randomised, open-label, Phase 3 trial | | | | |
| **Population** | Patients with WM who are relapsed/refractory or treatment naïve and considered to be unsuitable for chemotherapy | | | | |
| **Intervention(s)** | Zanubrutinib 160 mg BID to progression | | | | |
| **Comparator(s)** | Ibrutinib 420 mg OD to progression | | | | |
| **Indicate if trial supports application for marketing authorisation** | Yes | ✓ | Indicate if trial used in the economic model | Yes | ✓ |
| | No | | | No | |
| **Rationale for use/non-use in the model** | Pivotal Phase 3 trial supporting this indication | | | | |
| **Reported outcomes specified in the decision problem** | • **OS**<br>• **PFS**<br>• Response rate<br>• Time to next treatment<br>• DOR/remission<br>• **Adverse effects of treatment**<br>• **HRQoL** | | | | |
| **All other reported outcomes** | • Symptom resolution<br>• Serum IgM<br>• **TTD** | | | | |
| **Phase 1/2** | | | | | |
| **Study** | Study BGB-3111-AU-003 (NCT02343120)[44] | | | | |
| **Study design** | Multicentre, Phase 1/2 trial | | | | |
| **Population** | Patients with WM who are relapsed/refractory or treatment naïve | | | | |
| **Intervention(s)** | Zanubrutinib 40 mg OD, 80 mg OD, 160 mg OD, 320 mg OD or 160 mg BID | | | | |
| **Comparator(s)** | N/A | | | | |
| **Indicate if trial supports application for marketing authorisation** | Yes | ✓ | Indicate if trial used in the economic model | Yes | |
| | No | | | No | ✓ |
| **Rationale for use/non-use in the model** | Phase 1/2 trial supporting the evidence for the intervention within the indication | | | | |
| **Reported outcomes specified in the decision problem** | • Response rate<br>• OS<br>• DOR | | | | |

| | |
|---|---|
| | • PFS |
| **All other reported outcomes** | N/A |

Abbreviations: BID = twice daily; DOR = duration of response; HRQoL = health-related quality of life; OD = once daily; OS = overall survival; TTD = time to discontinuation; WM = Waldenström's macroglobulinaemia
Source: Tam et al., 2020[2]; Trotman et al., 2020[44]

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

The safety and efficacy of zanubrutinib have been evaluated in a comprehensive clinical trial programme, including the largest Phase 3 trial of BTK inhibitors in WM and the first head-to-head comparison of BTK inhibitors in any disease (the Phase 3 BGB-3111-302 [ASPEN] study; see Figure B.2.1).[2] ASPEN is the primary source of clinical evidence for this submission (see Sections B.2.6.1 and B.2.10.1), with supplemental long-term efficacy data provided by the Phase 1/2 BGB-3111-AU-003 study (see Section B.2.6.2), and pooled safety data presented for all WM patients (see Section B.2.10.2). A summary of methodology of ASPEN is provided here, with methodology of BGB-3111-AU-003 summarised in Appendix L.

### B.2.3.1 Study design and objectives

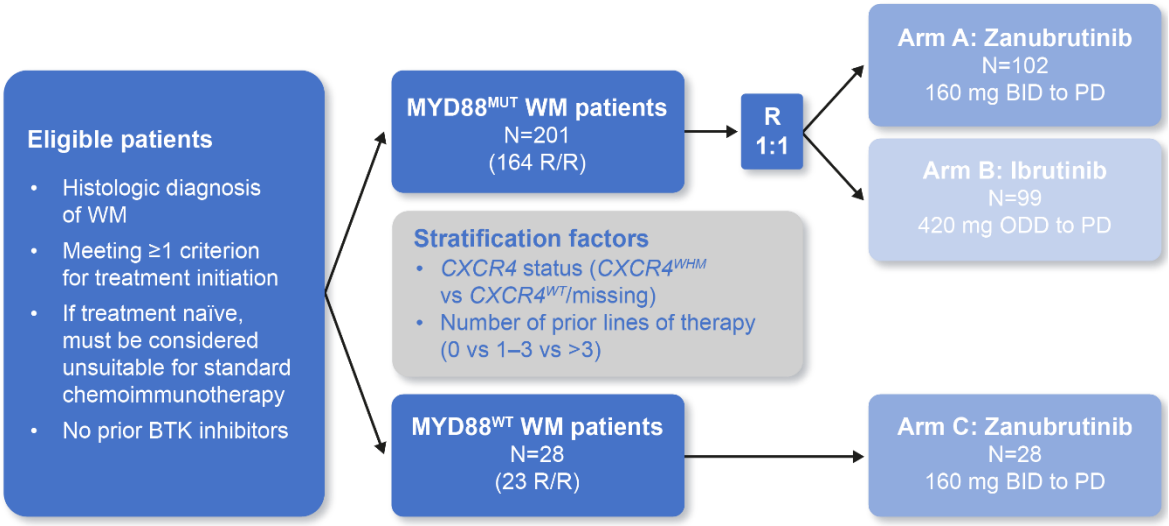
ASPEN is an ongoing Phase 3, open-label, two-arm, multicentre, randomised study of zanubrutinib versus ibrutinib for the treatment of WM in patients with relapsed/refractory disease, or who are treatment naïve and ineligible for chemoimmunotherapy.[2] The study was designed with two cohorts, according to *MYD88* status:

- Cohort 1: patients with *MYD88*$^{MUT}$; randomised to either zanubrutinib or ibrutinib{Tam, 2020 #48
- Cohort 2: patients with *MYD88*$^{WT}$; assigned to zanubrutinib.{Tam, 2020 #48}

The primary objective of ASPEN was to compare the efficacy and safety of zanubrutinib with ibrutinib in patients with *MYD88*$^{MUT}$ WM (relapsed/refractory arm of Cohort 1). The anti-cancer activity and safety of zanubrutinib in patients with *MYD88*$^{WT}$ WM (Cohort 2) was assessed as an exploratory objective.[2]

An overview of ASPEN study design is presented in Figure B.2.1, with a summary of methodology provided in Table B.2.2.

**Figure B.2.1. ASPEN study design (Study BGB-3111-302; NCT03053440)**




Abbreviations: BID = twice daily; BTK = Bruton tyrosine kinase; *CXCR4* = C-X-C Motif Chemokine Receptor 4; *MYD88*$^{MUT}$ = myeloid differentiation primary response gene 88 mutant; N = number of patients evaluable; OD = once daily; PD = progressive disease; R = randomised; R/R = relapsed/refractory; WM = Waldenström's macroglobulinaemia; WT = wild type
Source: Tam et al., 2020[45]

**Table B.2.2. Summary of methodology of ASPEN (Study BGB-3111-302; NCT03053440)**

| | |
|---|---|
| **Study design** | Multicentre, randomised, open-label, Phase 3 trial |
| **Locations (number of patients recruited)** | Australia (68), UK (33), Italy (27), Spain (24), US (21), Poland (19), Greece (13), Czech Republic (9), Sweden (7), Netherlands (5), Germany (2) and France (1) |
| **Study status** | Ongoing<br>• First patient treated: 25 January 2017<br>• Data cut-off date: 31 August 2019 |
| **Key eligibility criteria** | • Men and women aged ≥18 years<br>• Clinical and definitive histologic diagnosis of WM that is either treatment naïve and unsuitable for standard chemoimmunotherapy, or relapsed/refractory<br>• Meet at least one criterion from the Seventh IWWM<br>• ECOG Performance Status 0–2<br>• No prior exposure to a BTK inhibitor<br>• No WM central nervous system involvement |
| **Study treatments** | • Arm A: zanubrutinib 160 mg BID (N=102)<br>• Arm B: ibrutinib 420 mg OD (N=99)<br>• Arm C: zanubrutinib 160 mg BID (N=28) |
| **Concomitant medication** | **Permitted:**<br>• Corticosteroids (short-term administration)<br>**Disallowed:**<br>• Anti-cancer therapy (other than zanubrutinib or ibrutinib)<br>• Warfarin or other vitamin K antagonists<br>• Strong CYP3A inhibitors or inducers<br>• Fish oil and vitamin E supplements |
| **Primary outcomes** | • Rate of CR or VGPR, as assessed by IRC with adaptation of the response criteria updated at the Sixth IWWM |
| **Secondary outcomes** | • MRR (CR, VGPR or PR) as assessed by IRC<br>• Rate of CR or VGPR as assessed by the investigator<br>• DOR as assessed by IRC |

| | |
|---|---|
| | • DOR as assessed by the investigator<br>• PFS as assessed by the IRC<br>• PFS as assessed by the investigator |
| **Exploratory outcomes** | • CR/VGPR rate, MRR, ORR, PFS, DOR, and OS, as assessed by the IRC and by the investigator in patients with *MYD88*$^{WT}$ WM (Cohort 2)<br>• OS (Cohort 1)<br>• Time to response for MRR, according to *CXCR4* sequence (*CXCR4*$^{WHIM}$ versus *CXCR4*$^{WT}$) (Cohort 1)<br>• Time to next treatment (Cohort 1)<br>• MRR according to *CXCR4* sequence (*CXCR4*$^{WHIM}$ versus *CXCR4*$^{WT}$) in subjects with *MYD88*$^{MUT}$ WM (Cohort 1)<br>• IgM reduction over time |
| **PROs** | • EORTC QLQ-C30<br>• EQ-5D-5L |
| **Safety outcomes** | • AEs |
| **Pre-planned subgroups** | Subgroup analyses of the primary and selected secondary endpoints were conducted by:<br>• Gender<br>• Age (≤65 years versus >65 years; >75 years versus ≤75 years)<br>• Geographic region (Australia/New Zealand versus Europe versus North America)<br>• Number of prior lines of therapy (0 versus 1-3 versus ≥ 3 and relapsed/refractory versus treatment naïve)<br>• Baseline ECOG Performance Status (0 versus ≥1)<br>• Baseline *CXCR4* mutation status by Sanger method (WHIM versus WT/missing)<br>• Baseline IgM level (≤40 g/L versus >40 g/L)<br>• Baseline β-2 microglobulin level (≤3 mg/L versus >3 mg/L)<br>• Baseline haemoglobin concentration (≤110 g/L versus >110 g/L)<br>• Baseline platelet count (≤100 x $10^9$/L versus >100 x $10^9$/L)<br>• Baseline presence of extramedullary disease (yes versus no)<br>• WM IPSS (low versus intermediate versus high) |

Abbreviations: AEs = adverse event; ALT = alanine transaminase; AST = aspartate aminotransferase; BID = twice daily; CR = complete response; CYP3A= cytochrome P4503A; DOR = duration of response; ECOG = Eastern Cooperative Oncology Group; EORTC = European Organisation for Research and Treatment of Cancer; EQ-5D-5L = EuroQol 5-dimensions 5-level; IPSS = international prognostic scoring system; IRC = independent review committee; IWWM = International Workshop on Waldenström's Macroglobulinemia; MRR = major response rate; n = number of patients in the category; NYHA = New York Heart Association; ORR = objective response rate; OS = overall survival; PFS = progression-free survival; PR = partial response; QLQ-C30 = Quality of Life Questionnaire-Core 30; QTcF = T interval corrected for heart rate using Fridericia's formula; UK = United Kingdom; ULN = upper limit of normal; VGPR = very good partial response; WM = Waldenström's macroglobulinaemia
Source: BeiGene, 2020[46]

### B.2.3.2 Eligibility criteria

ASPEN included patients with WM who were relapsed/refractory after ≥1 prior line of therapy, or treatment naïve and unsuitable for standard immuno-chemotherapy.[2] Key inclusion and exclusion criteria are presented in Table B.2.3.

**Table B.2.3. Key inclusion and exclusion criteria**

| | |
|---|---|
| **Inclusion criteria** | • Men and women aged ≥18 years<br>• Clinical and definitive histologic diagnosis of WM that is either treatment naïve or relapsed/refractory<br>• If treatment naïve, unsuitable for standard chemoimmunotherapy due to comorbidities and risk factors as determined by the treating physician<br>• Meet at least one criterion from the Seventh IWWM<br>• Measurable disease<br>• ECOG Performance Status 0–2<br>• Adequate bone marrow function<br>• Creatinine clearance ≥30 mL/min, AST and ALT ≤3×ULN, and bilirubin ≤2×ULN<br>• Relapse after autologous stem cell transplant if they were at least 3 months after transplant, and after allogeneic transplant if they are at least 6 months post-transplant |
| **Exclusion criteria** | • Prior exposure to a BTK inhibitor<br>• Clinically active cardiovascular disease (i.e., uncontrolled arrhythmia, class 3/4 congestive heart failure as defined by the NYHA)<br>• QTcF prolongation >480 ms<br>• Major surgery within 4 weeks<br>• WM central nervous system involvement<br>• Concomitant warfarin, vitamin K antagonist or strong CYP3A inhibitors or strong CYP3A inducers |

Abbreviations: ALT = alanine transaminase; AST = aspartate aminotransferase; BTK = Bruton tyrosine kinase; CYP = cytochrome P450; ECOG = Eastern Cooperative Oncology Group; IWWM-7 = Seventh International Workshop on Waldenström's Macroglobulinaemia; NYHA = New York Heart Association; QTcF, QT interval corrected for heart rate using Fridericia's formula; ULN = upper limit of normal; WM = Waldenström's macroglobulinaemia
Source: BeiGene, 2020[46]

### B.2.3.3 Study treatments

#### *B.2.3.3.1 Allocation to treatment*

Patients in Cohort 1 were randomised in a 1:1 ratio to receive either zanubrutinib (Arm A) or ibrutinib (Arm B). Randomisation was stratified by *CXCR4* mutational status (*CXCR4*$^{WHIM}$ versus *CXCR4*$^{WT}$ versus missing) and number of prior therapies for WM (0 versus 1–3 versus >3).[2] Patients with *MYD88*$^{WT}$ disease or with undetermined *MYD88* mutation status were enrolled in Cohort 2 (Arm C).[46]

#### *B.2.3.3.2 Treatments administered*

Patients in Arm A or Arm C received zanubrutinib 160 mg (80 mg x 2) orally BID, with at least 8 hours between doses.[46]

Patients in Arm B received ibrutinib 420 mg (140 mg x 3, or in other applicable dose forms) orally OD as per the prescribing information.[46]

Zanubrutinib or ibrutinib were taken as prescribed from Cycle 1 Day 1 until disease progression, unacceptable toxicity or death, withdrawal of consent, loss to follow-up or termination of the study by the sponsor.[46]

#### *B.2.3.3.3 Dose modification*

Treatment interruption for ≤2 consecutive cycles and ≤2 dose reductions were permitted for management of recurring, Grade 3/4 treatment-related toxicities.[2] Permitted zanubrutinib and ibrutinib treatment modifications are presented in Table B.2.4.

**Table B.2.4. Zanubrutinib and ibrutinib dose reductions**

| Toxicity occurrence | Dose level | Zanubrutinib dose | Ibrutinib dose |
|---|---|---|---|
| First | 0 (starting dose) | Restart at 160 mg BID | Restart at 420 mg OD |
| Second | -1 dose level | Restart at 80 mg BID | Restart at 280 mg OD |
| Third | -2 dose level | Restart at 80 mg OD | Restart at 140 mg OD |
| Fourth | Discontinue | Discontinue | Discontinue |

Abbreviations: BID = twice daily; OD = once daily
Source: BeiGene, 2020[46]

### B.2.3.4 Assessments and outcomes

#### *B.2.3.4.1 Response assessment*

Response was evaluated using an adaptation of the response criteria updated at the Sixth IWWM (IWWM-6). Two different response criteria were used for this study: overall combined response and overall IgM response. Overall combined response considers the presence of extramedullary disease while the overall IgM response solely uses IgM reduction and immunofixation for response assessment. Overall combined response was used to assess the primary efficacy endpoint, whereas overall IgM response was used in a post-hoc analysis. Response categories included CR, VGPR, PR, minor response, stable disease, and progressive disease. Alternatively, an assessment of IgM flare could be assigned instead of progressive disease during periods of study drug withholding of at least seven consecutive days.[46]

Response assessments were performed every 4 weeks (every cycle) starting from Cycle 2, Day 1 for the first 48 weeks (12 cycles) then every 12 weeks (every 3 cycles) thereafter, based on physical examination (in cases in which organomegaly is present), laboratory evaluations, quantitative serum immunoglobulins and serum immuno-electrophoresis with M-protein quantitation by densitometry (serum protein electrophoresis [SPEP]), radiologic assessment and bone marrow studies.[2, 46]

#### *B.2.3.4.2 Efficacy outcomes*

The primary efficacy endpoint was the proportion of patients in the relapsed/refractory arm of Cohort 1 achieving either CR or VGPR, as determined by the independent review committee (IRC) using an adaptation of the response criteria updated at the IWWM-6.[2]

Secondary efficacy endpoints for Cohort 1 included:

- MRR assessed by the IRC, defined as the proportion of patients achieving CR, VGPR, or PR
- Duration of response (DOR) assessed by the IRC and the investigator, defined as the time from first determination of response (CR, VGPR, or PR; according to modified IWWM criteria) until first documentation of progression (according to modified IWWM criteria) or death, whichever comes first
- Rate of CR or VGPR assessed by the investigator

- PFS assessed by the IRC and the investigator, defined as the time from randomisation to the first documentation of progression (according to modified IWWM criteria) or death, whichever occurs first
- Resolution of treatment-precipitating symptoms, defined as the absence of the symptoms that triggered initiation of study treatment (according to IWWM treatment guidelines) at any point during study treatment
- Anti-lymphoma effect, defined as any reduction in bone marrow involvement by lymphoplasmacytoid lymphocytes and/or size of lymphadenopathy and/or splenomegaly by CT scan, at any time during study treatment[2]

#### *B.2.3.4.3 Safety outcomes*

The secondary safety endpoint for Cohort 1 was the incidence, timing and severity of adverse events (AEs).[2]

#### *B.2.3.4.4 Exploratory outcomes*

Exploratory endpoints included:

- Anti-cancer activity of zanubrutinib (i.e. CR/VGPR rate, major response rate, overall response rate, PFS, duration of response, and OS, as assessed by the IRC and by the investigator) in patients with *MYD88*$^{WT}$ WM (Cohort 2)
- Safety of zanubrutinib (i.e. incidence, severity, timing, and causation of adverse events) in patients with *MYD88*$^{WT}$ WM (Cohort 2)
- MRR according to *CXCR4* mutation status (*CXCR4*$^{WHIM}$ versus *CXCR4*$^{WT}$) in patients with *MYD88*$^{MUT}$ WM (Cohort 1)
- Time to response, defined as the time from cohort assignment until the date of first documentation of a PR or better, according to *CXCR4* mutation status (*CXCR4*$^{WHIM}$ versus *CXCR4*$^{WT}$) in patients with *MYD88*$^{MUT}$ WM (Cohort 1)
- OS, defined as the time from the date of randomisation until the date of death from any cause in patients with *MYD88*$^{MUT}$ WM (Cohort 1)
- Time to next treatment, defined as the time from the date of randomisation until the start date of a new anti-cancer therapy other than study medications in patients with *MYD88*$^{MUT}$ WM (Cohort 1)
- Reduction in IgM level from baseline over time in patients with *MYD88*$^{MUT}$ and *MYD88*$^{WT}$
- Change in quality of life as assessed by European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire-Core 30 (QLQ-C30) and EuroQol 5-dimensions 5-level (EQ-5D-5L) in patients with *MYD88*$^{MUT}$ WM (Cohort 1)
- Medical resource utilisation as assessed by the number of hospitalisations, length of hospital stays, and supportive care in patients with *MYD88*$^{MUT}$ WM (Cohort 1)[2, 47]

### B.2.3.5 Study population

#### *B.2.3.5.1 Analysis sets*

The following data sets were analysed:

- Cohort 1:
  - Intention-to-treat (ITT) Analysis Set: all randomised patients assigned to a treatment arm
  - Relapsed/Refractory Analysis Set: patients in the ITT Analysis Set with at least one prior line of therapy. This was the primary analysis set used for efficacy analyses
  - Per-protocol (PP) Analysis Set: patients in the ITT Analysis Set who received any dose of randomised treatment regimen, had a valid postbaseline measurement for either IgM (central or local) or M-protein by serum protein electrophoresis assessment (central or local), and did not have any important protocol deviation[46]
- Cohort 2:
  - Efficacy Analysis Set: all patients who received any dose of zanubrutinib and were centrally confirmed to have *MYD88*$^{WT}$[46]
  - Safety Analysis Set: all patients who received any dose of zanubrutinib or ibrutinib (Cohorts 1 and 2)[46]

The number and percentage of patients included in each analysis set are summarised in Table B.2.5.

**Table B.2.5. Analysis data sets**

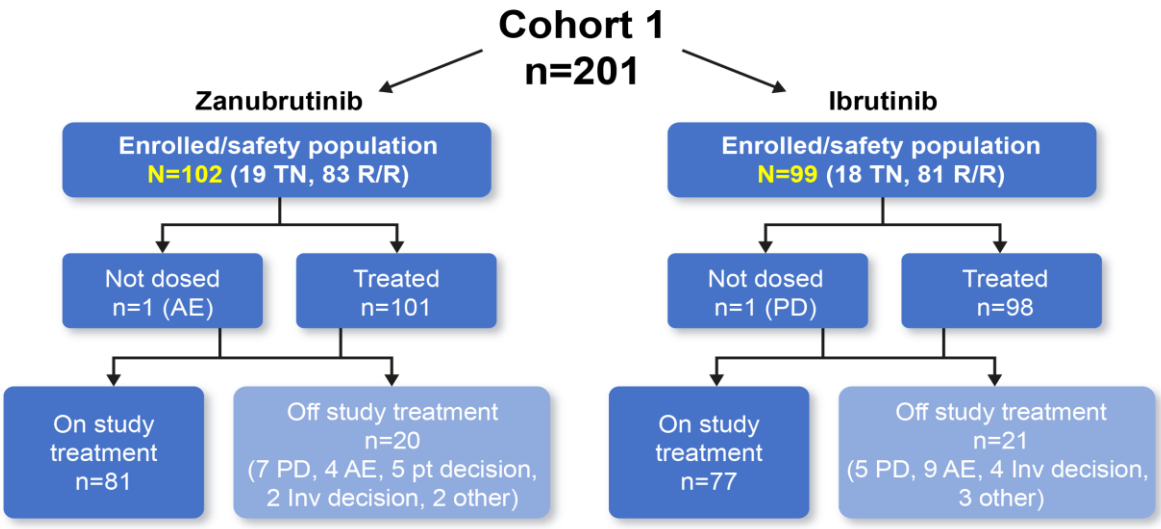
| Analysis set | Cohort 1 | | | | | | | | | Cohort 2 |
|---|---|---|---|---|---|---|---|---|---|---|
| | Treatment naïve | | | Relapsed/refractory | | | Overall | | | |
| | Zanubrutinib | Ibrutinib | Total | Zanubrutinib | Ibrutinib | Total | Zanubrutinib | Ibrutinib | Total | Zanubrutinib |
| ITT, n (%) | 19 (100.0) | 18 (100.0) | 37 (100.0) | 83 (100.0) | 81 (100.0) | 164 (100.0) | 102 (100.0) | 99 (100.0) | 201 (100.0) | - |
| Relapsed/Refractory, n (%) | - | - | - | 83 (100.0) | 81 (100.0) | 164 (100.0) | - | - | - | - |
| PP, n (%) | 19 (100.0) | 18 (100.0) | 37 (100.0) | 82 (98.8) | 79 (97.5) | 161 (98.2) | 101 (99.0) | 97 (98.0) | 198 (98.5) | - |
| PP Relapsed/ Refractory, n (%) | - | - | - | 82 (98.8) | 79 (97.5) | 161 (98.2) | - | - | - | - |
| Safety, n (%) | 19 (100.0) | 18 (100.0) | 37 (100.0) | 82 (98.8) | 80 (98.8) | 162 (98.8) | 101 (99.0) | 98 (99.0) | 199 (99.0) | 28 (100.0) |
| Efficacy, n (%) | - | - | - | - | - | - | - | - | - | 26 (92.9) |

Abbreviations: ITT = intention-to-treat; n = number of patients in the category; PP = per-protocol
Source: Tam et al., 2020[2]; BeiGene, 2020[46]; Dimopoulos et al., 2020[47]

### *B.2.3.5.2 Patient disposition*

In Cohort 1, the median follow-up time as of the data cut-off date was 19.5 months for zanubrutinib-treated patients and 19.4 months for ibrutinib-treated patients. A total of 201 patients were randomised (102 in the zanubrutinib arm and 99 in the ibrutinib arm) with 164 (81.6%) patients having relapsed/refractory disease (83 in the zanubrutinib treatment arm and 81 in the ibrutinib treatment arm). Two relapsed/refractory patients were randomised but not treated, 1 in the zanubrutinib treatment arm due to an adverse event (AE; unrelated to screening procedures) and 1 in the ibrutinib treatment arm due to progressive disease (central nervous system). As of the data cut-off date (31 August 2019), a total of 158 patients (78.6%) were continuing study treatment (81 patients [79.4%] in the zanubrutinib treatment arm and 77 patients [77.8%] in the ibrutinib treatment arm). The most common reason for discontinuing study treatment was progressive disease (7 [6.9%] zanubrutinib versus 5 [5.1%] ibrutinib - treated patients) and AE (4 [3.9%] zanubrutinib treated patients versus 9 [9.1%] ibrutinib-treated patients). A total of 158 (78.6%) patients were continuing to participate in the study and 41 (20.4%) discontinued from the study. Patient disposition for Cohort 1 is shown in Figure B.2.1.[46]

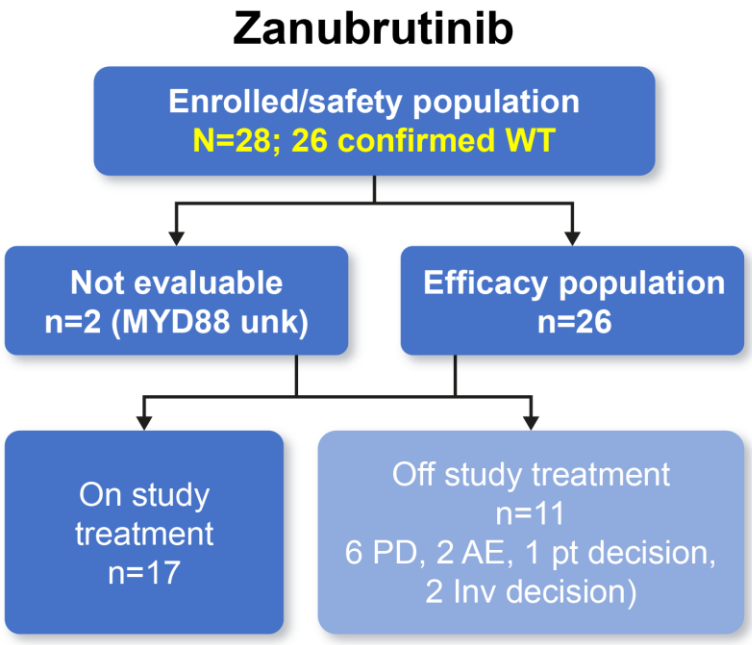
**Figure B.2.2. Patient disposition of Cohort 1**




Abbreviations: AE = adverse event; Inv = investigator; n = number of patients in the category; PD = progressive disease; pt = patient; R/R = relapsed/refractory; TN =treatment-naive
Source: BeiGene, 2020[46]

In Cohort 2, a total of 28 patients (23 relapsed/refractory) were enrolled and received zanubrutinib. Overall, the median follow-up time on study was 17.87 months (17.15 months for relapsed/refractory patients). As of the data cut-off date (31 August 2019), 17 (60.7%) patients were continuing to receive study treatment and 11 (39.3%) had discontinued study treatment. The most common reasons for discontinuing study treatment were progressive disease (6 [21.4%] patients) and AE (2 [7.1%] patients). Patient disposition for Cohort 2 is shown in Figure B.2.3.[46]

**Figure B.2.3. Patient disposition of Cohort 2**




Abbreviations: AE = adverse event; Inv = investigator; n = number of patients in the category; N = number of patients evaluable; PD = progressive disease; pt = patient; unk = unknown; WT = wild-type
Source: BeiGene, 2020[46]

Patient disposition for Cohort 1 (ITT Analysis Set) and Cohort 2 (Safety Analysis Set) are summarised in Table B.2.6.

**Table B.2.6. Patient disposition: Cohort 1 (ITT Analysis Set) and Cohort 2 (Safety Analysis Set)**

| Category | Cohort 1 | | | Cohort 2 |
|---|---|---|---|---|
| | Zanubrutinib (N=102) | Ibrutinib (N=99) | Total (N=201) | Zanubrutinib (N=28) |
| Randomised, not treated, n (%) | 1 (1.0) | 1 (1.0) | 2 (1.0) | 0 (0.0) |
| AE | 1 (1.0) | 0 (0.0) | 1 (0.5) | 0 (0.0) |
| Progressive disease | 0 (0.0) | 1 (1.0) | 1 (0.5) | 0 (0.0) |
| Treated, n (%) | 101 (99.0) | 98 (99.0) | 199 (99.0) | 28 (100.0) |
| On treatment, n (%) | 81 (79.4) | 77 (77.8) | 158 (78.6) | 17 (60.7) |
| Discontinued, n (%) | 20 (19.6) | 21 (21.2) | 41 (20.4) | 11 (39.3) |
| AE | 4 (3.9) | 9 (9.1) | 13 (6.5) | 2 (7.1) |
| Progressive disease | 7 (6.9) | 5 (5.1) | 12 (6.0) | 6 (21.4) |
| Investigator's discretion | 2 (2.0) | 4 (4.0) | 6 (3.0) | 2 (7.1) |
| Withdrawal by patient | 5 (4.9) | 0 (0.0) | 5 (2.5) | 1 (3.6) |
| Other | 2 (2.0) | 3 (3.0) | 5 (2.5) | 0 (0.0) |
| Median study follow-up (months) | 19.47 | 19.38 | 19.45 | 17.87 |
| Min, Max | 0.4, 31.2 | 0.5, 31.1 | 0.4, 31.2 | 2.3, 27.8 |

Abbreviations: AE = adverse event; ITT = intention-to-treat; n = number of patients in the category; N = number of patients evaluable
Source: Tam et al., 2020[2]; Dimopoulos et al., 2020[47]

### *B.2.3.5.3 Demographics and baseline characteristics*

The median age of all patients in the ITT Analysis Set (Cohort 1) was 70.0 years. The majority of patients were male (66.7%), white (91.0%), had an ECOG performance status of 0 or 1 and

were enrolled in sites in Europe (59.7%), Australia/New Zealand (30.8%) or North America (9.5%). The demographics and baseline characteristics were generally similar across treatment arms, however, more patients randomised to zanubrutinib than ibrutinib were >75 years old (33.3% and 22.2%, respectively) and more were anaemic (haemoglobin ≤110 g/L in 65.7% and 53.5% of patients, respectively).[46]

The median age of patients in Cohort 2 was 72.0 years with 42.9% >75 years of age. There was an equal number of male and female patients overall. The majority of patients were white (96.4%) and enrolled at sites in Europe (71.4%), Australia/New Zealand (21.4%) or North America (7.1%).[46]

A summary of baseline characteristics and demographics for Cohort 1 and Cohort 2 are shown in Table B.2.7.

**Table B.2.7. Demographics and baseline characteristics: Cohort 1 (ITT Analysis Set) and Cohort 2 (Safety Analysis Set)**

| Demographic/baseline characteristic | Cohort 1 | | | Cohort 2 |
|---|---|---|---|---|
| | Zanubrutinib (N=102) | Ibrutinib (N=99) | Total (N=201) | Zanubrutinib (N=28) |
| Median age (min, max), years | 70.0 (45, 87) | 70.0 (38, 90) | 70.0 | 72.0 (39, 87) |
| >75 years, n (%) | 34 (33.3) | 22 (22.2) | 56 (27.9) | 12 (42.9) |
| Gender, n (%) | | | | |
| Male | 69 (67.6) | 65 (65.7) | 134 (66.7) | 14 (50.0) |
| Race, n (%) | | | | |
| White | 88 (86.3) | 95 (96.0) | 183 (91.0) | 27 (96.4) |
| Asian | 4 (3.9) | 0 | 4 (2.0) | 0 |
| Unknown | 10 (9.8) | 4 (4.0) | 14 (7.0) | 1 (3.6) |
| ECOG PS | | | | |
| 0 | 46 (45.1) | 42 (42.4) | 88 (43.8) | 9 (32.1) |
| 1 | 50 (49.0) | 50 (50.5) | 100 (49.8) | 15 (53.6) |
| 2 | 6 (5.9) | 7 (7.1) | 13 (6.5) | 4 (14.3) |
| Prior lines of therapy, n (%) | | | | |
| 0 | 19 (18.6) | 18 (18.2) | 37 (18.4) | 5 (17.9) |
| 1-3 | 76 (74.5) | 74 (74.7) | 150 (74.6) | 20 (71.4) |
| >3 | 7 (6.9) | 7 (7.1) | 14 (7.0) | 3 (10.7) |
| Genotype | | | | |
| *MYD88*$^{MUT}$/*CXCR4*$^{WT}$ | 91 (89.2) | 90 (90.9) | 181 (90.0) | 23 (82.1) |
| *MYD88*$^{MUT}$/*CXCR4*$^{WHIM}$ | 11 (10.8) | 8 (8.1) | 19 (9.5) | 1 (3.6) |
| IPSS WM, n (%) | | | | |
| Low | 17 (16.7) | 13 (13.1) | 30 (14.9) | 5 (17.9) |
| Intermediate | 38 (37.3) | 42 (42.4) | 80 (39.8) | 11 (39.3) |
| High | 47 (46.1) | 44 (44.4) | 91 (45.3) | 12 (42.9) |
| Haemoglobin ≤110 g/L, n (%) | 67 (65.7) | 53 (53.5) | 120 (59.7) | 15 (53.6) |

Abbreviations: *CXCR4* = C-X-C Motif Chemokine Receptor 4; ECOG PS = Eastern Cooperative Oncology Group performance status; IPSS WM = International Prognostic Scoring System for Waldenström's Macroglobulinemia; ITT = intention-to-treat; *MYD88* = myeloid differentiation primary response gene 88; n = number of patients in the category; N = number of patients evaluable; WHIM = warts, hypogammaglobulinemia, infections, myelokathexis; WT = wild-type
Source: Tam et al., 2020[2]; BeiGene, 2020[46]

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

### B.2.4.1 Primary efficacy analysis

Two hypotheses were tested using a hierarchical fixed-sequence procedure to adjust for multiplicity – superiority of zanubrutinib compared with ibrutinib in CR/VGPR rate in 1) the Relapsed/Refractory Analysis Set, and 2) the ITT Analysis Set.[2, 46] A Cochran-Mantel-Haenszel (CMH) test for difference in CR/VGPR rates was performed for both comparisons, stratified by the *CXCR4* status (WHIM versus WT/missing), prior line of therapy (1–3 versus >3 for the Relapsed/Refractory Analysis Set; 0 versus 1–3 versus >3 in the ITT Analysis Set) and age group (≤65 years versus >65 years) at a two-sided significance level of 0.025.[2] The primary objective was met if the two-sided p-value was <0.05 and the estimated difference was positive.[2]

The primary analysis of superiority in the primary endpoint was performed in the Relapsed/Refractory Analysis Set first, at least 15 months after 90% enrolment in this analysis set was completed. If superiority was demonstrated in the Relapsed/Refractory Analysis Set, superiority was further tested in the ITT Analysis Set.[46]

### B.2.4.2 Secondary efficacy analyses

#### *B.2.4.2.1 Major response rate*

Statistical significance for the first or both response comparisons triggered a test of non-inferiority in MRRs between zanubrutinib and ibrutinib, at a one-sided significance level of 0.025.[2, 46] The null and alternative hypotheses were:

- $H_0$: $MRR_{Arm\ A}$ - $MRR_{Arm\ B}$ ≤-12%
- Ha: $MRR_{Arm\ A}$ - $MRR_{Arm\ B}$ >-12%.[46]

The 95% CI for the Mantel-Haenszel common risk difference was constructed as for the primary endpoint. If the lower bound of the CI was greater than the non-inferiority margin of -12%, non-inferiority in MRR would be demonstrated.[46] If the lower limit of the 95% CI was >0%, superiority of zanubrutinib in MRR would be demonstrated.[2]

#### *B.2.4.2.2 Progression-free survival*

PFS by treatment arm was estimated at the time of primary efficacy analysis by Kaplan-Meier (KM) methodology.[2] PFS was right-censored for patients who met one of the following conditions:

- No baseline disease assessments
- Starting a new anti-cancer therapy before disease progression or death
- Disease progression or death immediately after >6 months since the last disease assessment (>12 months if a patient was on the response assessment schedule of every 24 weeks)
- Alive without documentation of disease progression.[46]

Two-sided, 95% CIs for median PFS were estimated with the Brookmeyer and Crowley method. KM methodology was used to estimate PFS at selected time points, with

corresponding 95% CIs estimated using the Greenwood's formula. Duration of follow-up for PFS was estimated using the reverse KM method.[2]

The HR (Arm A/Arm B) for PFS and its two-sided 95% CI were estimated from a stratified Cox regression model, stratified by *CXCR4* status, prior lines of therapy, and age group performed only at the final analysis of PFS. An unstratified Cox regression model was used to estimate the HR of PFS in zanubrutinib compared with ibrutinib and the corresponding 95% CI at the final analysis of PFS.[46]

#### *B.2.4.2.3 Duration of response*

Censoring conventions for duration of response were as described for PFS. DOR was not compared between the two arms.[46]

### B.2.4.3 Safety analyses

Safety data were summarised by treatment arm and by combining Arms A and C in the Safety Analysis Set using descriptive statistics.[46] The distribution of times to first occurrence of AESIs was summarised using KM methodology.[2]

## B.2.5 Quality assessment of the relevant clinical effectiveness evidence

**Table B.2.8. Quality assessment of ASPEN**

| | |
|---|---|
| **Was randomisation carried out appropriately?** | Yes |
| **Was the concealment of treatment allocation adequate** | N/A |
| **Were the groups similar at the outset of the study in terms of prognostic factors?** | Yes, patients were well balanced between study arms for key characteristics, Randomisation was stratified by *CXCR4* mutation status and number of prior lines of therapy. |
| **Were the care providers, participants and outcome assessors blind to treatment allocation?** | No formal blinding was used as the study used an open-label design. However, potential bias was mitigated by determination of the primary endpoint by an IRC |
| **Were there any unexpected imbalances in drop-outs between groups?** | No, study treatment discontinuation was similar between groups |
| **Is there any evidence to suggest that the authors measured more outcomes than they reported?** | No |
| **Did the analysis include an analysis? If so, was this appropriate and were appropriate methods used to account for missing data?** | Yes |

Abbreviations: *CXCR4* = C-X-C Motif Chemokine Receptor 4; IRC = independent review committee; ITT = intention-to-treat; N/A = not applicable

## B.2.6 Clinical effectiveness results of the relevant trials

### B.2.6.1 ASPEN

Efficacy data from ASPEN are presented for the ITT Analysis Set and the Cohort 2 Efficacy Analysis Set (including both treatment-naïve and relapsed/refractory patients), representing the proposed licensed indication.

#### *B.2.6.1.1 Cohort 1*

##### *B.2.6.1.1.1 IRC-assessed VGPR/CR rate (primary endpoint)*

In Cohort 1, the rate of IRC-assessed CR and VGPR was 28.4% in all patients treated with zanubrutinib and 19.2% in patients treated with ibrutinib (95% CI, -1.5–22.0; p=0.09). The estimated difference between the two arms adjusted for the stratification factors and age group was 10.2% (Table B.2.9).[2]

**Table B.2.9. IRC-assessed response in Cohort 1 (ITT Analysis Set)**

| Assessment | Zanubrutinib (N=102) | Ibrutinib (N=99) |
|---|---|---|
| CR + VGPR, n (%) | 29 (28.4) | 19 (19.2) |
| CR + VGPR risk difference (95% CI) | 10.2 (-1.5–22.0) p=0.09 | |
| OR, n (%) | 96 (94.1) | 92 (92.9) |
| MRR, n (%) | 79 (77.5) | 77 (77.8) |
| Best overall response, n (%) | | |
| CR | 0 (0.0) | 0 (0.0) |
| VGPR | 29 (28.4) | 19 (19.2) |
| PR | 50 (49.0) | 58 (58.6) |
| Minor response | 17 (16.7) | 15 (15.2) |
| Stable disease | 3 (2.9) | 3 (3.0) |
| Progressive disease | 2 (2.0) | 2 (2.0) |
| Not evaluable | 0 (0.0) | 2 (2.0) |

Abbreviations: CI = confidence interval; CR = complete response; IRC = independent review committee; ITT = intention-to-treat; MRR = major response rate; n = number of patients in the category; N = number of patients evaluable; OR = overall response; PR = partial response; VGPR = very good partial response rate
Source: Tam et al., 2020[2]; BeiGene, 2020[46]

In the relapsed/refractory population, 28.9% of patients treated with zanubrutinib and 19.8% treated with ibrutinib achieved VGPR or CR (with estimated difference of 10.7% [95% CI, -2.5–23.9; p=0.116).[2]

The testing for the primary endpoint of VGPR or CR rate superiority required testing in the Relapsed/Refractory Analysis Set prior to testing in the ITT Analysis Set. While numerically higher rates of VGPR or CR were seen across analysis sets, the primary efficacy endpoint was not significant in the Relapsed/Refractory Analysis Set (p=0.116), thus the study did not meet the primary efficacy endpoint and testing for other endpoints and resulting p-values in the following sections are descriptive.[46]

*B.2.6.1.1.2 IRC-assessed duration of response (secondary endpoint)*

In Cohort 1, the median durations of VGPR or CR and major response according to overall combined assessment were not reached in either treatment arm who achieved a response to the study treatment, as shown in Table B.2.10.[2]

**Table B.2.10. IRC-assessed duration of response in Cohort 1 (ITT Analysis Set)**

| Assessment | Zanubrutinib (N=102) | Ibrutinib (N=99) |
|---|---|---|
| Duration of CR or VGPR | | |
| Median follow-up, months (95% CI) | 13.6 (9.7–16.6) | 7.7 (2.8–12.9) |
| Median DOR, months (95% CI) | NE (NE-NE) | NE (8.0-NE) |
| Event-free rate at, % (95% CI) | | |
| 12 months | 100.0 (NE–NE) | 64.2 (28.8–85.4) |
| 18 months | 92.9 (59.1–99.0) | 64.2 (28.8–85.4) |
| 24 months | NE (NE–NE) | NE (NE–NE) |
| Duration of Major Response | | |
| Median follow-up, months (95% CI) | 14.8 (13.8–16.8) | 13.9 (12.3–15.7) |
| Median DOR, months (95% CI) | NE (NE-NE) | NE (NE-NE) |
| Event-free rate at, % (95% CI) | | |
| 12 months | 94.4 (85.8–97.9) | 87.9 (77.0–93.8) |
| 18 months | 85.2 (71.7–92.6) | 87.9 (77.0–93.8) |
| 24 months | 85.2 (71.7–92.6) | 81.6 (62.4–91.6) |

Abbreviations: CI = confidence interval; CR = complete response; DOR = duration of response; IRC, independent review committee; ITT = intention-to-treat; N = number of patients evaluable; NE = not evaluable; VGPR = very good partial response
Source: Tam et al., 2020[2]; BeiGene, 2020[46]

*B.2.6.1.1.3 IRC-assessed progression-free survival (secondary endpoint)*

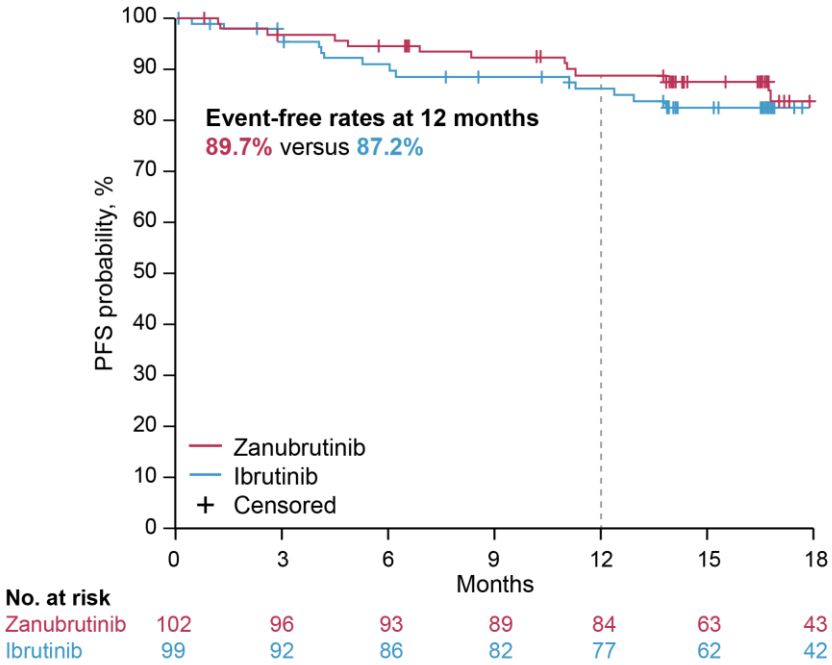
At the time of the data cut-off date, median PFS was not reached in either treatment arm of Cohort 1. The event-free rates at 12 months for patients treated with zanubrutinib or ibrutinib were 89.7% and 87.2%, respectively,[46] and 85.0% and 83.8% at 18 months[2] (Table B.2.11 and Figure B.2.4).

**Table B.2.11. IRC-assessed progression-free survival in Cohort 1 (ITT Analysis Set)**

| Assessment | Zanubrutinib (N=102) | Ibrutinib (N=99) |
|---|---|---|
| Median follow-up, months (95% CI) | 18.0 (16.7–19.4) | 18.5 (16.7–19.3) |
| Median PFS, months (95% CI) | NE (NE–NE) | NE (NE–NE) |
| Events, n (%) | | |
| Progressive disease | 13 (12.7) | 10 (10.1) |
| Death | 2 (2.0) | 6 (6.1) |
| Event-free rate at, % (95% CI) | | |
| 6 months | 95.0 (88.4–97.9) | 91.6 (83.9–95.7) |
| 9 months | 92.9 (85.7–96.5) | 89.5 (81.3–94.2) |
| 12 months | 89.7 (81.7–94.3) | 87.2 (78.6–92.5) |
| 18 months | 85.0 (75.2–91.2) | 83.8 (74.5–89.9) |
| 24 months | 79.4 (66.2–88.0) | 81.5 (71.1–88.5) |

Abbreviations: CI = confidence interval; IRC = independent review committee; ITT = intention-to-treat; n = number of patients in the category; N = number of patients evaluable; NE = not evaluable; PFS = progression-free survival
Source: Tam et al., 2020[2]; BeiGene, 2020[46]

**Figure B.2.4. IRC-assessed progression-free survival in Cohort 1 (ITT analysis set)**




Abbreviations: No. = number; PFS = progression free survival
Source: Tam et al., 2020[2]

*B.2.6.1.1.4 IRC-assessed time to response (secondary endpoint)*

The median time to VGPR or CR according to overall combined IRC assessment was shorter in the zanubrutinib arm than the ibrutinib arm (4.8 versus 7.4 months).[46] Time to major response (2.8 versus 2.8 months)[2] and overall response (1.0 versus 1.0 months)[46] were the same between the treatment groups.

*B.2.6.1.1.5 Overall survival (exploratory endpoint)*

At the time of the data cut-off date, OS had not been reached in either treatment arm (Table B.2.12).[46]

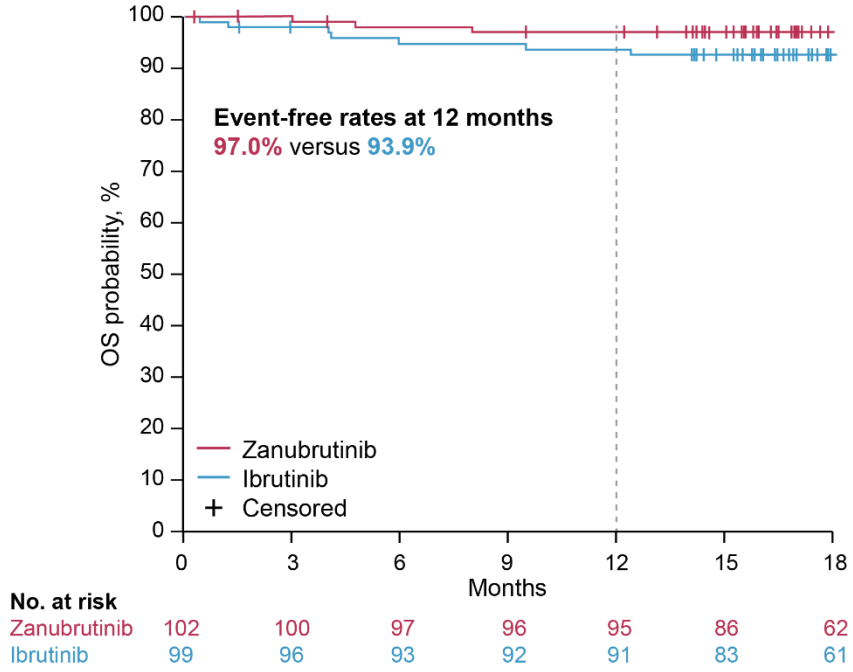
**Table B.2.12. OS in Cohort 1 (ITT Analysis Set)**

| Assessment | Zanubrutinib (N=102) | Ibrutinib (N=99) |
|---|---|---|
| Median follow-up, months (95% CI) | 19.5 (18.1–20.8) | 19.7 (18.7–20.9) |
| Median OS, months (95% CI) | NE (NE–NE) | NE (NE–NE) |
| Event-free rate at, % (95% CI) | | |
| 12 months | 97.0 (90.9–99.0) | 93.9 (86.8–97.2) |
| 18 months | 97.0 (90.9–99.0) | 92.8 (85.5–96.5) |
| 24 months | 89.5 (76.4–95.5) | 91.0 (82.5–95.5) |

Abbreviations: CI = confidence interval; ITT = intention-to-treat; N = number of patients evaluable; NE = not evaluable; OS = overall survival
Source: BeiGene, 2020[46]

OS at 12 months was 97.0% among patients treated with zanubrutinib and 93.9% among patients treated with ibrutinib (Figure B.2.5).[46]

**Figure B.2.5. OS in Cohort 1 (ITT Analysis Set)**

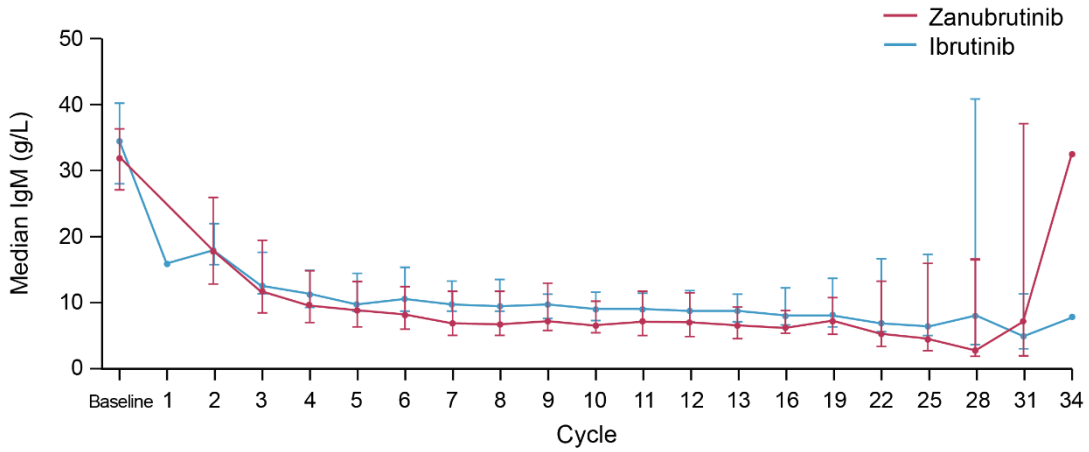



Abbreviations: No. = number; OS = overall survival
Source: BeiGene, 2020[46]

*B.2.6.1.1.6 Serum IgM improvement over time (exploratory endpoint)*

Serum IgM levels decreased over time for patients in both the zanubrutinib and ibrutinib treatment arms (79%, interquartile range [IQR] 88–63 versus 72%, IQR 86–58).[2] Zanubrutinib demonstrated greater and more sustained reductions in IgM by both the repeated-measured mixed-effect model (p=0.0314) and AUC (p=0.0370) compared with ibrutinib (Figure B.2.6).[2, 46]

**Figure B.2.6. Changes in serum IgM levels over time in Cohort 1 (ITT Analysis Set)**

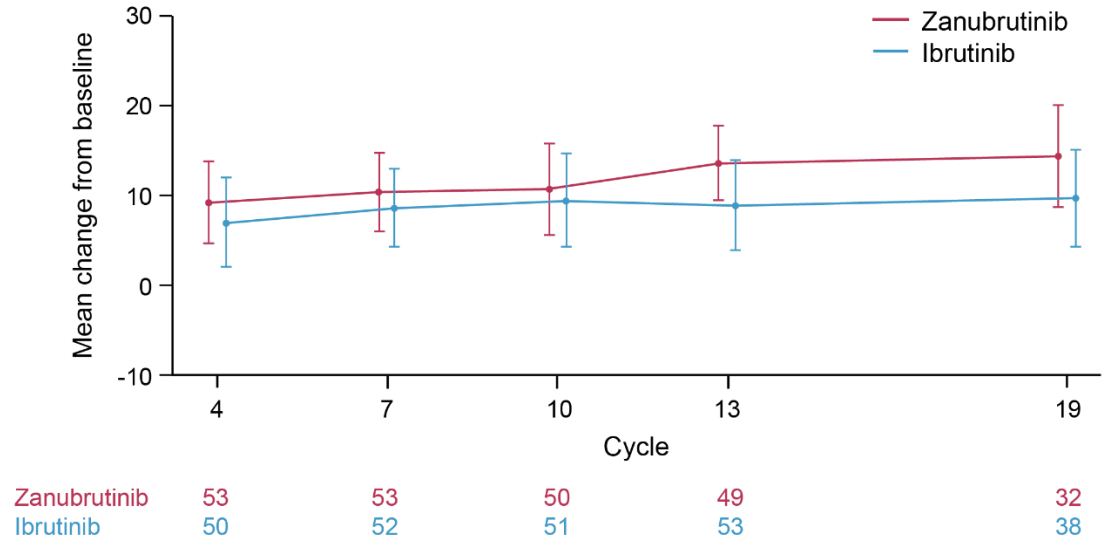



Abbreviations: IgM = immunoglobulin M; ITT = intention-to-treat
Source: BeiGene, 2020[46]

*B.2.6.1.1.7 Patient-reported outcomes*

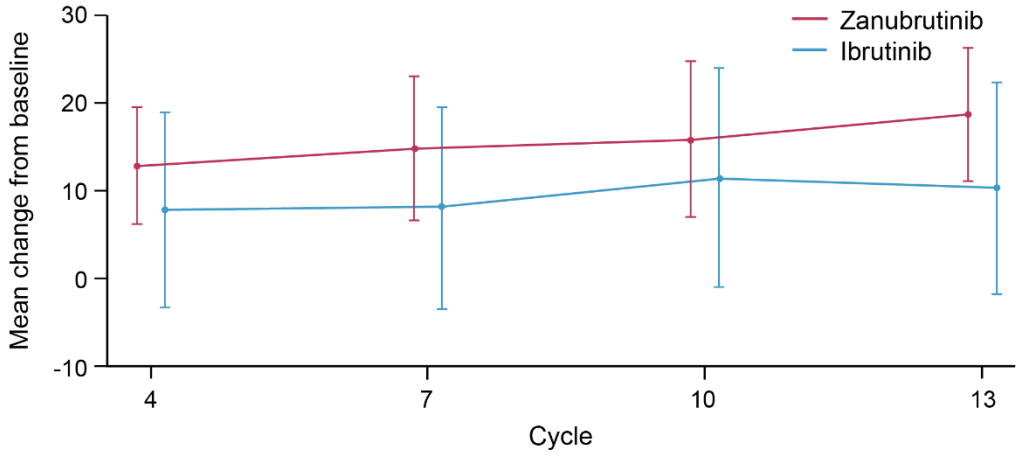
Zanubrutinib demonstrated a similar improvement to ibrutinib in QoL from baseline, with notable improvements in EQ-5D-5L and EORTC QLQ-C30 seen for loss of appetite, fatigue (mean decrease ~30%), physical (mean change from baseline >10%) and role functioning (mean increase from baseline ~20%), and dyspnoea (mean decrease >30%; Figure B.2.7; Figure B.2.9).[46] Zanubrutinib trended towards a greater improvement than ibrutinib, particularly when analysed over the first year on treatment in patients who achieved a deeper response (i.e. a response assessment of VGPR; Figure B.2.8; Figure B.2.10 ).[2]

**Figure B.2.7. EQ-5D-5L score: change from baseline over time in all patients in Cohort 1 (ITT Analysis Set)**




Abbreviations: EQ-5D-5L = EuroQol 5-Dimensions 5-Level; ITT = intention-to-treat
Source: BeiGene, 2020[46]

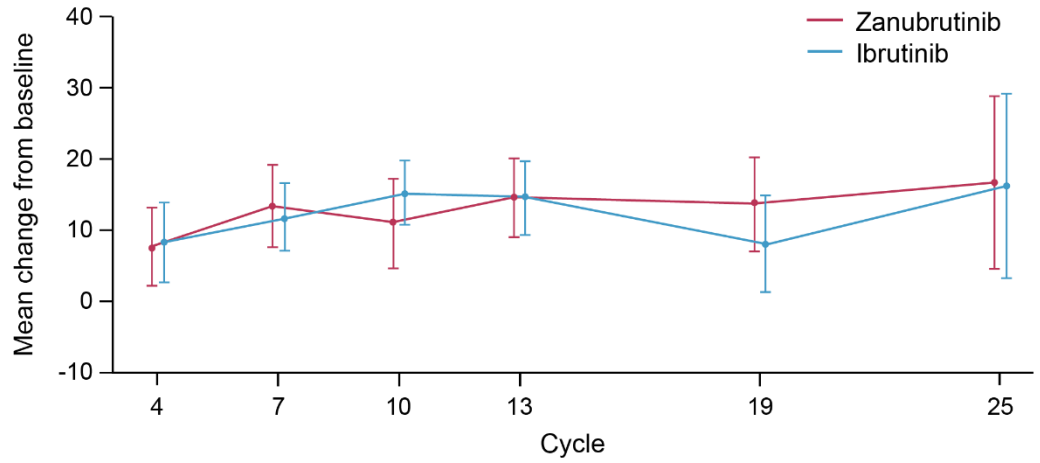
**Figure B.2.8. EQ-5D-5L score: change from baseline over time in patients achieving VGPR in Cohort 1 (ITT Analysis Set)**




| | | | | |
|---|---|---|---|---|
| Zanubrutinib | 14 | 15 | 15 | 15 |
| Ibrutinib | 12 | 12 | 12 | 13 |

Abbreviations: EQ-5D-5L = EuroQol 5-Dimensions 5-Level; ITT = intention-to-treat; VGPR = very good partial response
Source: BeiGene, 2020[46]

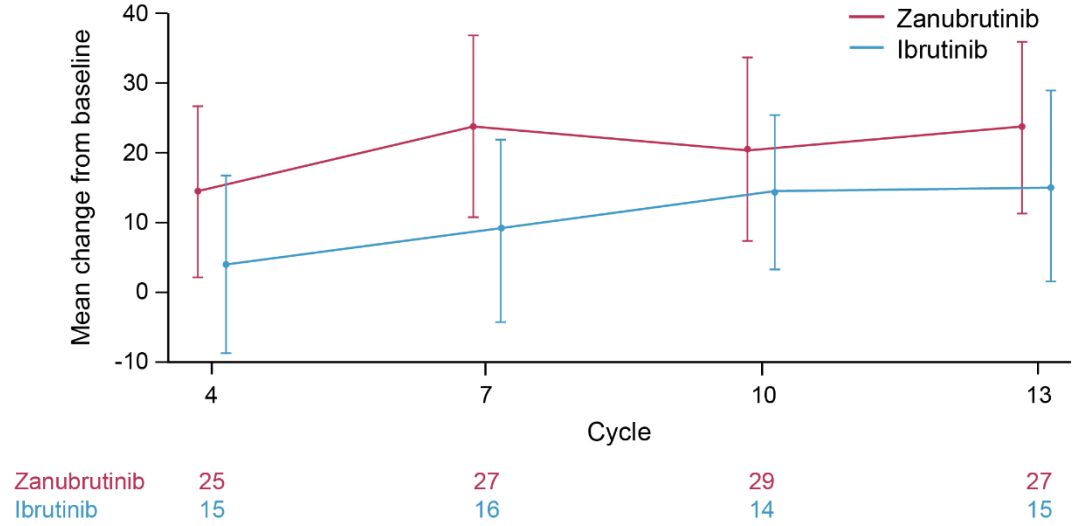
**Figure B.2.9. EORTC QLQ-C30 Global Health Status: change from baseline over time in all patients in Cohort 1 (ITT Analysis Set)**




| | | | | | | |
|---|---|---|---|---|---|---|
| Zanubrutinib | 88 | 87 | 83 | 81 | 66 | 21 |
| Ibrutinib | 70 | 70 | 69 | 69 | 52 | 15 |

Abbreviations: EORTC = European Organisation for Research and Treatment of Cancer; QLQ-C30 = Quality of Life Questionnaire-Core 30
Source: BeiGene, 2020[46]

**Figure B.2.10. EORTC QLQ-C30 Global Health Status: change from baseline over time in patients achieving VGPR in Cohort 1 (ITT Analysis Set)**




Abbreviations: EORTC = European Organisation for Research and Treatment of Cancer; QLQ-C30 = Quality of Life Questionnaire-Core 30; VGPR = very good partial response
Source: BeiGene, 2020[46]

*B.2.6.1.1.8 Time to next treatment*

The median times to initiation of non-protocol anti-cancer therapy were 6.83 months in the zanubrutinib treatment arm and 6.44 months in the ibrutinib treatment arm.[46]

### *B.2.6.1.2 Cohort 2*

All outcomes in Cohort 2 (patients with *MYD88*$^{WT}$) were exploratory.

*B.2.6.1.2.1 IRC-assessed CR/VGPR rate*

In Cohort 2, 26.9% of patients achieved CR or VGPR (Table B.2.13).[43]

**Table B.2.13. IRC-assessed disease response in Cohort 2 (Efficacy Analysis Set)**

| Assessment | Zanubrutinib (N=26) |
|---|---|
| Median follow-up, months | 17.9 |
| CR + VGPR, n (%) | 7 (26.9) |
| Overall response, n (%) | 21 (80.8) |
| Major response, n (%) | 13 (50.0) |
| Best overall response, n (%) | |
| CR | 0 (0.0) |
| VGPR | 7 (26.9) |
| PR | 6 (23.1) |
| Minor response | 8 (30.8) |
| Stable disease | 4 (15.4) |
| Progressive disease | 1 (3.8) |

Abbreviations: CR = complete response; IRC = independent review committee; n = number of patients in the category; N = number of patients evaluable; PR = partial response; VGPR = very good partial response
Source: Dimopoulos et al., 2020;[47] Garcia Sanz et al. 2020[43]

#### *B.2.6.1.2.2 IRC-assessed duration of response*

In patients who achieved a response to zanubrutinib in Cohort 2, the median duration of VGPR or CR and major response was not reached as of the data cut-off date (Table B.2.14).[46]

**Table B.2.14. IRC-assessed duration of response in Cohort 2 (Efficacy Analysis Set)**

| Assessment | Zanubrutinib (N=26) |
|---|---|
| Duration of CR or VGPR | |
| Median follow-up, months (95% CI) | 8.5 (0.0–19.3) |
| Median DOR, months (95% CI) | NE (8.1–NE) |
| Event-free rate at, % (95% CI) | |
| 12 months | 75.0 (12.8–96.1) |
| 18 months | 75.0 (12.8–96.1) |
| 24 months | NE (NE–NE) |
| Duration of major response | |
| Median follow-up, months (95% CI) | 12.0 (8.5–17.0) |
| Median DOR, months (95% CI) | NE (6.3–NE) |
| Event-free rate at, % (95% CI) | |
| 12 months | 62.3 (27.7–84.0) |
| 18 months | 62.3 (27.7–84.0) |
| 24 months | NE (NE–NE) |

Abbreviations: CI = confidence interval; CR = complete response; DOR = duration of response; IRC = independent review committee; N = number of patients evaluable; NE = not evaluable; VGPR = very good partial response
Source: BeiGene, 2020[46]

#### *B.2.6.1.2.3 IRC-assessed progression-free survival*

Median PFS in Cohort 2 was 27.5 months with an event-free rate at 12 months of 72.4% (Table B.2.15).[43]

**Table B.2.15. IRC-assessed progression-free survival in Cohort 2 (Efficacy Analysis Set)**

| Assessment | Zanubrutinib (N=26) |
|---|---|
| Median follow-up, months (95% CI) | 17.5 (13.9–19.4) |
| Median PFS, months (95% CI) | 27.5 (13.7–27.5) |
| Event-free rate at, % (95% CI) | |
| 6 months | 88.5 (68.4–96.1) |
| 9 months | 80.4 (59.1–91.4) |
| 12 months | 72.4 (50.6–85.8) |
| 18 months | 68.1 (46.2–82.6) |
| 24 months | 68.1 (46.2–82.6) |

Abbreviations: CI = confidence interval; IRC = independent review committee; N = number of patients evaluable; NE = not evaluable; PFS = progression-free survival
Source: BeiGene, 2020[46]; Garcia Sanz et al. 2020[43]

*B.2.6.1.2.4 IRC-assessed time to response*

Median times to VGPR or CR, major response and overall response were 5.65, 2.89 and 0.99 months, respectively. Time to response results for Cohort 2 were similar to those in Cohort 1 for patients treated with zanubrutinib.[46]

*B.2.6.1.2.5 Overall survival*

At the time of the data cut-off date, median OS was not reached in Cohort 2 (Table B.2.16).[46]

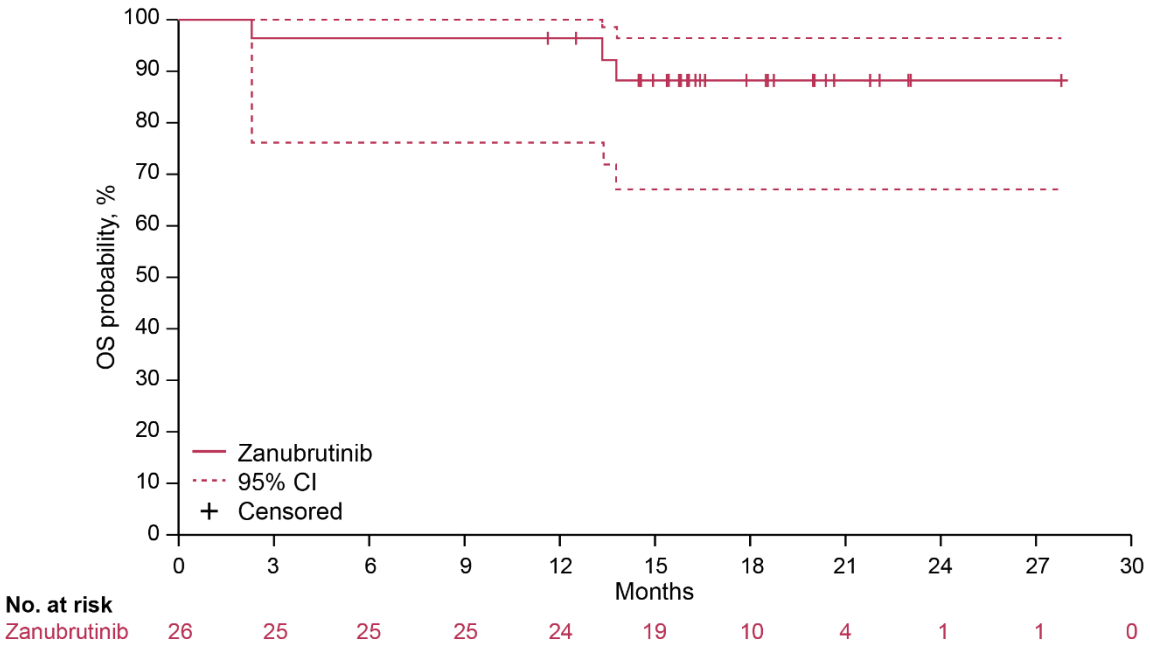
**Table B.2.16. OS in Cohort 2 (Efficacy Analysis Set)**

| Assessment | Zanubrutinib (N=26) |
|---|---|
| Median follow-up, months (95% CI) | 16.5 (15.7–18.7) |
| Median OS, months (95% CI) | NE (NE–NE) |
| Event-free rate at, % (95% CI) | |
| 12 months | 96.2 (75.7–99.4) |
| 18 months | 87.8 (66.7–95.9) |
| 24 months | 87.8 (66.7–95.9) |

Abbreviations: CI = confidence interval; N = number of patients evaluable; NE = not evaluable; OS = overall survival
Source: BeiGene, 2020[46]

OS at 12 months was 96.2% (95% CI, 75.7–99.4) in all patients treated with zanubrutinib (Figure B.2.11).[46]

**Figure B.2.11. OS in Cohort 2 (Efficacy Analysis Set)**

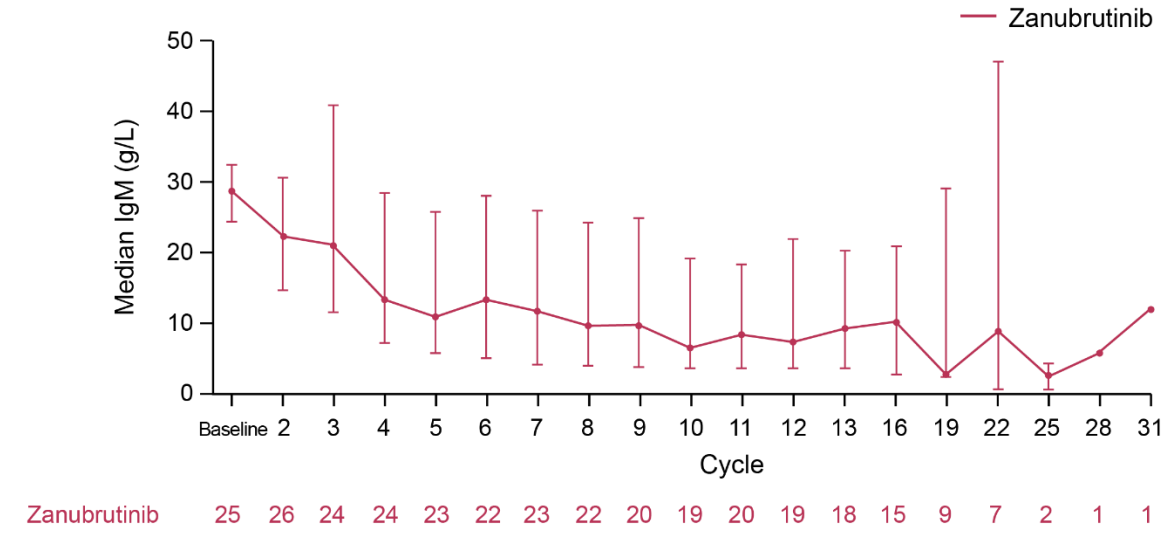



Abbreviations: CI = confidence interval; No. = number
Source: BeiGene, 2020[46]

*B.2.6.1.2.6 Serum IgM improvement over time*

Similar to results in Cohort 1, serum IgM levels decreased over time for patients treated with zanubrutinib in Cohort 2 (Figure B.2.12).[46]

**Figure B.2.12. Changes in serum IgM levels over time in Cohort 2 (Efficacy Analysis Set)**




Abbreviations: IgM = immunoglobulin M
Source: BeiGene, 2020[46]

### B.2.6.2 BGB-3111-AU-003

Long-term follow-up data from the BGB-3111-AU-003 Phase 1/2 study (N=78) of zanubrutinib have demonstrated deep and durable responses in patients with treatment-naïve or relapsed/refractory WM.[48] At a median follow-up of 30.3 months, overall response rate was 95.9% and rates of VGPR/CR increased with prolonged treatment from 20.5% at 6 months, to 32.9% at 12 months and 43.8% at 24 months.[44, 48]

Median OS was not reached at the time of data cut-off. The OS rate for the overall study population at 36 months was 84.8% and PFS was 80.5%.[44]

**Table B.2.17. Efficacy endpoints in patients with WM (Efficacy Evaluable Set)**

| Assessment | Zanubrutinib (N=73) |
|---|---|
| Overall response, n (%) | 70 (95.9) |
| VGPR + CR, n (%) | 33 (45.2) |
| 36-month PFS, % (95% CI) | 80.5 (68.5–88.3) |
| 36-month OS, % (95% CI) | 84.8 (71.3–92.3) |

Abbreviations: CI = confidence interval; CR = complete response; N = number of patients evaluable; OS = overall survival; PFS = progression-free survival; VGPR = very good partial response
Source: Trotman et al., 2020[44]; BeiGene, 2020[48]

### B.2.6.3 Efficacy conclusion

Data from the pivotal Phase 3 ASPEN (BGB-3111-302) study and supportive BGB-3111-AU-003 study have demonstrated that zanubrutinib has a clear efficacy benefit in the treatment of patients with WM, independent of line of therapy and independent of *MYD88* mutational status.

Zanubrutinib treatment resulted in rapid, deep, and sustained reduction in IgM, high overall and major responses, and high rates of VGPR, along with improvements in QoL measures and symptom resolution, with equivalent response quality and PFS to ibrutinib.

In ASPEN, zanubrutinib demonstrated an equivalent VGPR rate to ibrutinib in patients with relapsed/refractory or treatment-naïve WM (28.4% versus 19.2%) and was achieved at an earlier median time of 4.8 months after zanubrutinib treatment compared with 7.4 months after ibrutinib treatment. Similarly, 12-month PFS and OS data were comparable between zanubrutinib and ibrutinib (89.7% versus 87.2%, and 97.0% versus 93.9%, respectively). Compared with ibrutinib, zanubrutinib demonstrated significantly greater and more sustained reductions in IgM, a marker of WM disease control, by both the repeated-measured mixed-effect model (p=0.0314) and AUC (p=0.037).[2]

In addition to equivalent clinical outcomes, zanubrutinib also demonstrated similar improvements in QoL compared to ibrutinib, particularly when analysed over the first year of treatment in patients who achieved a deeper response. Notable improvements in EQ-5D-5L and EORTC QLQ-C30 scores were seen for loss of appetite, fatigue (mean decrease ~30%), physical (mean change from baseline >10%) and role functioning (mean increase from baseline ~20%), and dyspnoea (mean decrease >30%).[46]

Long-term follow-up data from the Phase 1/2 BGB-3111-AU-003 study of zanubrutinib showed deep and durable responses in patients with relapsed/refractory or treatment-naïve WM. At a median follow-up of 32.7 months, overall response rate was 95.9% and rates of VGPR/CR increased with prolonged treatment from 20.5% at 6 months, 32.9% at 12 months and 43.8% at 24 months.[44]

Zanubrutinib has demonstrated major responses in patients with *MYD88*$^{WT}$ WM. ASPEN included the largest cohort of patients with WM with confirmed *MYD88*$^{WT}$ (n=26) studied in terms of efficacy of a BTK inhibitor. Zanubrutinib showed clinically meaningful antitumor activity in patients with *MYD88*$^{WT}$ WM, with a major response rate of 50.0% including 26.9% with VGPR.[43] Taken together, these data demonstrate that treatment with zanubrutinib is an effective strategy to improve clinical and QoL outcomes in patients with WM, regardless of mutation status.[2, 43, 44, 46]

## B.2.7 Subgroup analysis

Subgroup analyses of the primary and selected secondary endpoints were conducted by:

- Gender
- Age (≤65 years versus >65 years; >75 years versus ≤75 years)
- Geographic region (Australia/New Zealand versus Europe versus North America)
- Number of prior lines of therapy (0 versus 1-3 versus ≥ 3 and relapsed/refractory versus treatment naïve)
- Baseline ECOG Performance Status (0 versus ≥1)
- Baseline *CXCR4* mutation status by Sanger method (WHIM versus WT/missing)
- Baseline IgM level (≤40 g/L versus >40 g/L)
- Baseline β-2 microglobulin level (≤3 mg/L versus >3 mg/L)
- Baseline haemoglobin concentration (≤110 g/L versus > 10 g/L)
- Baseline platelet count (≤100 x $10^9$/L versus > 100 x $10^9$/L)
- Baseline presence of extramedullary disease (yes versus no)

- WM IPSS (low versus intermediate versus high).[46]

#### B.2.7.1 Summary of results

Results of the subgroup analyses are presented in Appendix E. The proportions of patients in Cohort 1 and Cohort 2 who achieved a VGPR or CR were generally consistent for subgroups of interest. In Cohort 1, zanubrutinib treatment was favoured in patients ≤75 years and in prognostically more difficult-to-treat patients, such as those with higher IgM (≥40 g/L), cytopenias (e.g., haemoglobin concentration ≤110 g/L; baseline platelet count ≤100 × $10^9$/L), extramedullary disease, and medium/high IPSS scores.[46]

## B.2.8 Meta-analysis

Efficacy data supporting the use of zanubrutinib for the treatment of WM are primarily provided by a single Phase 3 study (ASPEN). Therefore, a meta-analysis was not conducted.

## B.2.9 Indirect and mixed treatment comparisons

Although direct head-to-head data comparing zanubrutinib to ibrutinib are available from the Phase 3 ASPEN RCT, ibrutinib was not included in the final scope issued for this appraisal (see Table B.1.1), and there was a lack of randomised trials identified by the SLR (see Appendix D) directly comparing zanubrutinib to the comparators of interest listed in the final scope. Therefore, an indirect treatment comparison was necessary. As described in Appendix D, two studies identified in the SLR were included in the indirect treatment comparison – one single-arm trial, Tedeschi et al. 2015,[49] for BR and another single-arm trial, Dimopoulos et al. 2007/Kastritis et al. 2015[50, 51] for DRC.

Other than BR and DRC, it was not possible to conduct further indirect comparisons, due to a lack of data in the literature to enable the comparison of zanubrutinib with the comparators of interest. However, BR and DRC currently represent the two most common regimens for the first-line treatment of WM in patients considered fit enough to tolerate them (13.1% and 16.2%, respectively [see Section B.1.3.5.2]), and the third- and second- most common second-line regimens, respectively, behind ibrutinib (18.2%).[1]

Given the lack of a common comparator linking zanubrutinib to the comparators of interest, traditional indirect treatment comparison (ITC) methods using anchored comparators were not feasible. As such, matching adjusted indirect comparisons (MAIC) were conducted to reweight individual patient-level data (IPD) of the zanubrutinib arm in ASPEN to match the populations treated with BR and DRC separately.

#### B.2.9.1 Data sources

IPD for zanubrutinib were available from ASPEN and published summary data for BR and DRC were available from the associated publications (see Appendix D).

Tedeschi et al. 2015 was considered the most suitable study of BR for inclusion in the MAIC as it reported evaluable PFS and OS KM curves, and baseline patient characteristics; had the largest sample size; and was the only EU-based study.[49] Similarly, Dimopoulos et al. 20074/Kastritis et al. 2015 was selected as it was the only prospective study of DRC; reported an evaluable OS KM curve and baseline patient characteristics; had the largest sample size; and was the only EU-based study.[50, 51]

#### *B.2.9.1.1 Individual patient-level data*

IPD for zanubrutinib were available from ASPEN, a randomised, Phase 3 study comparing zanubrutinib and ibrutinib in patients with WM who have received at least one prior therapy or as first-line treatment for patients unsuitable for chemo-immunotherapy with the *MYD88*$^{MUT}$ mutation.

#### *B.2.9.1.2 Comparator data*

For baseline patient characteristics and AE incidence, summary mean estimates were extracted from comparator trial publications whenever available.

Individual patient-level event and censoring times for survival were derived via a 2-step process for OS and PFS KM curves. First, the numerical value of the curves (i.e., time on the x-axis and proportion of patients alive on the y-axis) were obtained through graphical digitisation, using WebPlotDigitizer (http://rapps.pharmerit.com/km-curve-digitization-tool/). Second, the number of events and censoring at each time point was manually calibrated to create a "simulated" trial population that would reproduce the KM curves presented in trial publications, based on the reported number of patients at risk and/or the marker for censoring on the KM curves.

### B.2.9.2 Methodology

Three sets of pairwise MAIC were conducted, as summarised in Table B.2.18. In addition to the two pairwise comparisons that matched the overall zanubrutinib population (N=102) to the BR (N=71) and DRC (N=72) populations separately, a subgroup analysis was conducted matching zanubrutinib patients with relapsed/refractory disease (n=83) to the BR population (N=71), considering that the BR population consisted of relapsed/refractory patients only.

As described in Appendix D, no subgroup analysis was conducted for the comparison with DRC (including 72 patients for whom chemo-immunotherapy is suitable), given the small sample size of treatment-naïve patients (unsuitable for chemo-immunotherapy) in the zanubrutinib arm in ASPEN (N=19).

**Table B.2.18. Pairwise MAIC**

| Pairwise comparison | Zanubrutinib population | Comparator population |
|---|---|---|
| 1 | 102 patients in the zanubrutinib arm in the ASPEN ITT Analysis Set | 71 relapsed/refractory patients in the trial for BR[49] |
| 2 | 83 patients in the relapsed/refractory set of zanubrutinib arm in the ASPEN ITT Analysis Set | 71 relapsed/refractory patients in the trial for BR[49] |
| 3 | 102 patients in the zanubrutinib arm in the ASPEN ITT Analysis Set | 72 treatment-naïve patients in the trial for DRC[50, 51] |

Abbreviations: BR = bendamustine rituximab; DRC = dexamethasone, rituximab and cyclophosphamide; ITT = intention-to-treat; MAIC = matching-adjusted indirect comparison; OS = overall survival; PFS = progression-free survival

#### *B.2.9.2.1 Matching variables*

Matching adjustment in an MAIC ensures that treatment outcomes are comparable across trial populations to the extent of the considered baseline characteristics. Ideally, matching should be based on clinically relevant risk factors that can modify relative treatment effects. According

to the NICE Decision Support Unit (DSU) technical support document 18,[52] effect modifier status should be justified prior to analysis. For unanchored comparisons, every prognostic variable as well as effect modifier should be included.

Based on ASPEN and other published literature,[46, 50, 51, 53] a range of baseline patient characteristics were considered to be potential prognostic factors or effect modifiers, and therefore, considered for inclusion as matching variables in the MAICs regardless of data availability in the comparator trials. Matching variables included were:

- Age (≤75 versus >75 years;[46] ≤65, 66–75 versus >75 years)[51]
- Number of prior therapies (0–3 versus >3 lines;[46] 1–3 versus 3 lines)[53]
- ECOG performance status (0–1 versus >1)[46]
- *MYD88*/*CXCR4* mutation status[46]
- IgM concentration (≤40 versus >40 g/L)[46]
- β2-microglobulin concentration (≤3 versus >3 mg/L)[46]
- Platelet count (≤100 versus >100×$10^9$/L)[46]
- Haemoglobin concentration (≤110 versus >110 g/L)[46]
- Presence of extramedullary disease[46, 50]
- WM IPSS.[46]

#### *B.2.9.2.2 Matching adjustment: propensity score weighting*

To adjust for differences in baseline patient and disease characteristics among the trials, the matching algorithm proposed by Signorovitch et al. 2012 was used.[54] Specifically, IPD for zanubrutinib obtained from ASPEN were reweighted such that the weighted mean baseline characteristics matched those reported in the comparator trial publications separately. In the process of matching adjustment, each patient was assigned a weight representing the inverse of the odds of being in the ASPEN zanubrutinib arm versus being in a specific comparator trial.[54] Patients in the zanubrutinib arm who were more likely to be in the comparator trial population (based on characteristics) were assigned a higher weight in the analysis and vice versa. By assigning a weight to each patient based on baseline characteristics, each patient has more or less influence on the analysis depending on that patient's likelihood of being in the comparator trial.

Analyses were performed using the *sandwich* package in R.

#### *B.2.9.2.3 Outcome comparison*

Outcome comparisons were conducted both before and after matching adjustment, with investigator-assessed PFS and OS KM curves as the primary outcomes of interest.

Survival was compared by estimating hazard ratios (HRs) using Cox proportional hazard (PH) models when reconstructed patient data from reported KM curves were available for BR and DRC.

### B.2.9.3 Results

After matching adjustment, the reported baseline characteristics were well matched for each of the comparisons (see Appendix D). The PFS and OS KM curves before and after matching adjustment are presented in Figure B.2.13 to Figure B.2.16 for the comparison with BR, and Figure B.2.17 and Figure B.2.18 for the comparison with DRC.

**Figure B.2.13. KM curves of PFS – zanubrutinib (ITT population) vs BR**

Abbreviations: BR = bendamustine and rituximab; ITT = intention-to-treat; KM = Kaplan-Meier; PFS = progression-free survival; vs = versus

**Figure B.2.14. KM curves of PFS – zanubrutinib (relapsed/refractory subgroup) vs BR**

Abbreviations: BR = bendamustine and rituximab; KM = Kaplan-Meier; PFS = progression-free survival; vs = versus

**Figure B.2.15. KM curves of OS – zanubrutinib (ITT population) vs BR**

Abbreviations: BR = bendamustine and rituximab; ITT = intention-to-treat; KM = Kaplan-Meier; OS = overall survival; vs = versus

**Figure B.2.16. KM curves of OS – zanubrutinib (relapsed/refractory subgroup) vs BR**

Abbreviations: BR = bendamustine and rituximab; KM = Kaplan-Meier; OS = overall survival; vs = versus

**Figure B.2.17. KM curves of PFS – zanubrutinib (ITT population) vs DRC**

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; ITT = intention-to-treat; KM = Kaplan-Meier; PFS = progression-free survival; vs = versus

**Figure B.2.18. KM curves of OS – zanubrutinib (ITT population) vs DRC**

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; ITT = intention-to-treat; KM = Kaplan-Meier; PFS = progression-free survival; vs = versus

Compared with DRC, zanubrutinib was associated PFS (HR [95% CI]: and before and after matching adjustment, respectively) and (HR [95% CI]: and before and after matching, respectively). Compared with BR, zanubrutinib was associated with (HR [95% CI]: and before and after matching, respectively) and (HR [95% CI]: and before and after matching, respectively). It should be noted that the (see Section B.3.3.2, Figure B.3.15, Figure B.3.16, Figure B.3.27 and Figure B.3.28), which suggested that the validity of the PH assumption was ambiguous. For additional assumptions that should be considered when interpreting the results, refer to Section B.2.9.4.

For more comprehensive MAIC outputs applied in the cost-effectiveness analysis, including OS, PFS and time to discontinuation (TTD) KM curves, and incidence of AEs of Grade ≥3, see Section B.3.3.

### B.2.9.4 Uncertainties in the indirect and mixed treatment comparisons

Table B.2.19 presents the key assumptions of the MAIC.

**Table B.2.19. Key assumptions of the MAIC**

| Category | Assumption | Justification |
|---|---|---|
| Population (zanubrutinib matched to DRC vs DRC) | The zanubrutinib arm in ASPEN (N=102), including a mix of relapsed/refractory population (N=83) and treatment-naïve population unsuitable for chemo-immunotherapy (N=19), was adjusted to match the DRC | This assumption was necessitated by the limited availability of clinical evidence (see Appendix D).<br>It should be acknowledged that such differences in patient populations might have led to an underestimation of the |

| | | |
|---|---|---|
| | population which included only treatment-naïve population suitable for chemo-immunotherapy (N=72), assuming that the discrepancies in patient populations had limited impact on the MAIC results. | relative survival benefit of zanubrutinib compared with DRC, assuming that PFS and OS outcomes are more favourable in the treatment-naïve population (suitable for chemo-immunotherapy) compared with both treatment-naïve population (unsuitable for chemo-immunotherapy) and relapsed/refractory population. |
| Matching variables | It was assumed that any unobserved key prognostic factors were well balanced between the zanubrutinib arm and comparator arms, such that the MAIC results were robust with limited biases. | It is rarely possible to completely adjust for all unobserved or unreported baseline patent characteristics, which is a general limitation of a MAIC. Despite that, the outcome comparison was conducted before and after matching adjustment, which consistently showed survival benefit of zanubrutinib compared with the comparators. |

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; MAIC = matching-adjusted indirect comparison; N = number of patients evaluable; OS = overall survival; PFS = progression-free survival

## B.2.10 Adverse reactions

### B.2.10.1 ASPEN

#### *B.2.10.1.1 Extent of exposure*

The safety population included all patients who received any dose of zanubrutinib (N=129 [Cohort 1 N=101; Cohort 2 N=28]) or ibrutinib (N=98 [Cohort 1]).[2, 46]

The overall median duration of zanubrutinib treatment was 18.7 months in Cohort 1 and 16.4 months in Cohort 2 (see Table B.2.20).[2, 46] In Cohort 1, the median duration of treatment was comparable between the zanubrutinib and ibrutinib treatment arms (18.7 months and 18.6 months, respectively); 89% and 84% of patients had minimal exposures of 12 months.[2] In Cohort 1, the median relative dose intensities were 97.6% for zanubrutinib and 98.2% for ibrutinib. The median relative zanubrutinib dose intensity in Cohort 2 was 96.9%.[46]

**Table B.2.20. Exposure to Study Drug (Safety Analysis Set)**

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| Duration of exposure, months | | | | |
| Mean (SD) | 18.18 (6.305) | 15.11 (6.761) | 17.52 (6.505) | 17.41 (7.056) |
| Median (min, max) | 18.73 (0.8, 31.2) | 16.39 (1.4, 27.8) | 18.37 (0.8, 31.2) | 18.55 (0.3, 30.9) |
| Relative dose intensity, % | | | | |
| Mean (SD) | 91.67 (14.372) | 93.52 (10.860) | 92.07 (13.669) | 92.44 (11.295) |
| Median (min, max) | 97.64 (29.0, 100.0) | 96.92 (51.0, 100.0) | 97.51 (29.0, 100.0) | 98.18 (51.6, 100.0) |

Abbreviations: N = number of patients evaluable; SD = standard deviation
Source: BeiGene, 2020[46]

### *B.2.10.1.2 Adverse events*

An overview of AEs is presented in Table B.2.21. Higher proportions of ibrutinib-treated patients had at least one AE, Grade ≥3 AEs; SAEs, AEs leading to death, AEs leading to treatment discontinuation, and TRAEs compared with zanubrutinib-treated patients.[46]

**Table B.2.21. Overview of AEs (Safety Analysis Set)**

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| AEs, n (%) | 98 (97.0) | 24 (85.7) | 122 (94.6) | 97 (99.0) |
| Grade ≥3 | 59 (58.4) | 18 (64.3) | 77 (59.7) | 62 (63.3) |
| SAEs | 40 (39.6) | 11 (39.3) | 51 (39.5) | 40 (40.8) |
| AEs leading to death | 1 (1.0) | 0 (0.0) | 1 (0.8) | 4 (4.1) |
| AEs leading to discontinuation | 4 (4.0) | 2 (7.1) | 6 (4.7) | 9 (9.2) |
| TRAEs, n (%) | 80 (79.2) | 22 (78.6) | 102 (79.1) | 84 (85.7) |
| Grade ≥3 | 33 (32.7) | 13 (46.4) | 46 (35.7) | 42 (42.9) |
| AESIs, n (%) | 86 (85.1) | 23 (82.1) | 109 (84.5) | 81 (82.7) |

Abbreviations: AE = adverse event; AESI = adverse event of special interest; n = number of patients in the category; N = number of patients evaluable; SAE = serious adverse event; TRAE = treatment-related adverse event
Source: BeiGene, 2020[46]

AEs are summarised by system organ class (SOC) and preferred term (PT) in Table B.2.22. In Cohort 1, the most common AEs (reported in ≥20% of patients) among zanubrutinib-treated patients were neutropenia (24.8%), upper respiratory tract infection (23.8%), and diarrhoea (20.8%).[2] Numerous AEs were ≥10% more prevalent in the ibrutinib arm compared with the zanubrutinib arm, including muscle spasms (23.5% versus 9.9%), atrial fibrillation (14.3% versus 2.0%), diarrhoea (31.6% versus 20.8%), contusion (23.5% versus 12.9%), peripheral oedema (19.4% versus 8.9%), and pneumonia (12.2% versus 2.0%).[46] The only AE more prevalent (>10% higher) in the zanubrutinib arm compared with the ibrutinib treatment arm was neutropenia (12.2% and 24.8%, respectively).[46] In all zanubrutinib treated patients, the incidences of AEs were generally comparable between Cohorts 1 and 2, except for neutropenia (24.8% versus 14.3%), nausea (14.9% versus 3.6%), and dyspnoea (13.9% versus 3.6%), which were more prevalent (>10% difference) in Cohort 1; and pneumonia (2.0% versus 14.3%), respiratory tract infection (5.9% versus 17.9%), and decreased appetite (4.0% versus 14.3%), which were more prevalent (>10% difference) in Cohort 2.[46]

**Table B.2.22. AEs by SOC and PT reported in >10% of patients (Safety Analysis Set)**

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| AEs, n (%) | 98 (97.0) | 24 (85.7) | 122 (94.6) | 97 (99.0) |
| Infections and infestations | | | | |
| Upper respiratory tract infection | 24 (23.8) | 6 (21.4) | 30 (23.3) | 28 (28.6) |
| Urinary tract infection | 10 (9.9) | 4 (14.3) | 14 (10.9) | 10 (10.2) |
| Nasopharyngitis | 11 (10.9) | 2 (7.1) | 13 (10.1) | 7 (7.1) |
| Pneumonia | 2 (2.0) | 4 (14.3) | 6 (4.7) | 12 (12.2) |
| Gastrointestinal disorders | | | | |
| Diarrhoea | 21 (20.8) | 8 (28.6) | 29 (22.5) | 31 (31.6) |
| Constipation | 16 (15.8) | 4 (14.3) | 20 (15.5) | 7 (7.1) |
| Nausea | 15 (14.9) | 1 (3.6) | 16 (12.4) | 13 (13.3) |
| Vomiting | 9 (8.9) | 2 (7.1) | 11 (8.5) | 13 (13.3) |
| Blood and lymphatic system disorders | | | | |

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| Neutropenia | 25 (24.8) | 4 (14.3) | 29 (22.5) | 12 (12.2) |
| Anaemia | 12 (11.9) | 6 (21.4) | 18 (14.0) | 10 (10.2) |
| Thrombocytopenia | 10 (9.9) | 3 (10.7) | 13 (10.1) | 10 (10.2) |
| General disorders and administration site conditions | | | | |
| Fatigue | 19 (18.8) | 4 (14.3) | 23 (17.8) | 15 (15.3) |
| Pyrexia | 13 (12.9) | 6 (21.4) | 19 (14.7) | 12 (12.2) |
| Oedema peripheral | 9 (8.9) | 4 (14.3) | 13 (10.1) | 19 (19.4) |
| Injury, poisoning and procedural complications | | | | |
| Contusion | 13 (12.9) | 6 (21.4) | 19 (14.7) | 23 (23.5) |
| Musculoskeletal and connective tissue disorders | | | | |
| Back pain | 14 (13.9) | 4 (14.3) | 18 (14.0) | 6 (6.1) |
| Arthralgia | 13 (12.9) | 3 (10.7) | 16 (12.4) | 16 (16.3) |
| Pain in extremity | 11 (10.9) | 1 (3.6) | 12 (9.3) | 7 (7.1) |
| Muscle spasms | 10 (9.9) | 4 (14.3) | 14 (10.9) | 23 (23.5) |
| Respiratory, thoracic and mediastinal disorders | | | | |
| Cough | 13 (12.9) | 5 (17.9) | 18 (14.0) | 17 (17.3) |
| Dyspnoea | 14 (13.9) | 1 (3.6) | 15 (11.6) | 6 (6.1) |
| Epistaxis | 13 (12.9) | 1 (3.6) | 14 (10.9) | 19 (19.4) |
| Nervous system disorders | | | | |
| Headache | 15 (14.9) | 3 (10.7) | 18 (14.0) | 11 (11.2) |
| Dizziness | 13 (12.9) | 1 (3.6) | 14 (10.9) | 9 (9.2) |
| Skin and subcutaneous tissue disorders | | | | |
| Rash | 13 (12.9) | 3 (10.7) | 16 (12.4) | 16 (16.3) |
| Pruritus | 9 (8.9) | 4 (14.3) | 13 (10.1) | 5 (5.1) |
| Vascular disorders | | | | |
| Hypertension | 11 (10.9) | 3 (10.7) | 14 (10.9) | 16 (16.3) |
| Renal and urinary disorders | | | | |
| Haematuria | 7 (6.9) | 1 (3.6) | 8 (6.2) | 10 (10.2) |
| Cardiac disorders | | | | |
| Atrial fibrillation | 2 (2.0) | 1 (3.6) | 3 (2.3) | 14 (14.3) |

Abbreviations: AE = adverse event; n = number of patients in the category; N = number of patients evaluable; PT = preferred term; SOC = system organ class
Source: Tam et al., 2020[2]; BeiGene, 2020[46]

Grade ≥3 AEs are summarised by SOC and PT in Table B.2.23. In Cohort 1, the most common Grade ≥3 AEs (reported in ≥5% of patients) among zanubrutinib-treated patients were neutropenia (15.8%), hypertension (5.9%), thrombocytopenia (5.9%) and anaemia (5.0%). As with all Grade AEs, the only Grade ≥3 AE more prevalent (>5% higher) in the zanubrutinib arm compared with the ibrutinib treatment arm was neutropenia (15.8% and 8.2%, respectively). Grade ≥3 AEs more prevalent in the ibrutinib arm compared with the zanubrutinib arm were pneumonia (7.1% versus 1.0%) and hypertension (11.2% versus 5.9%). In all zanubrutinib treated patients, the incidences of Grade ≥3 AEs were generally comparable between Cohorts 1 and 2.[46]

**Table B.2.23. Grade ≥3 AEs by SOC and PT reported in >2% of patients (Safety Analysis Set)**

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| AEs Grade ≥3, n (%) | 59 (58.4) | 18 (64.3) | 77 (59.7) | 62 (63.3) |
| Blood and lymphatic system disorders | | | | |
| Neutropenia | 16 (15.8) | 3 (10.7) | 19 (14.7) | 8 (8.2) |
| Anaemia | 5 (5.0) | 3 (10.7) | 8 (6.2) | 5 (5.1) |
| Thrombocytopenia | 6 (5.9) | 2 (7.1) | 8 (6.2) | 3 (3.1) |
| Febrile neutropenia | 4 (4.0) | 0 (0.0) | 4 (3.1) | 0 (0.0) |
| Vascular disorders | | | | |
| Hypertension | 6 (5.9) | 3 (10.7) | 9 (7.0) | 11 (11.2) |
| Gastrointestinal disorders | | | | |
| Diarrhoea | 3 (3.0) | 2 (7.1) | 5 (3.9) | 1 (1.0) |
| Investigations | | | | |
| Neutrophil count decreased | 4 (4.0) | 0 (0.0) | 4 (3.1) | 1 (1.0) |
| Musculoskeletal and connective tissue disorders | | | | |
| Back pain | 4 (4.0) | 0 (0.0) | 4 (3.1) | 0 (0.0) |
| Arthralgia | 3 (3.0) | 0 (0.0) | 3 (2.3) | 0 (0.0) |
| Nervous system disorders | | | | |
| Syncope | 4 (4.0) | 0 (0.0) | 4 (3.1) | 2 (2.0) |
| Infections and infestations | | | | |
| Pneumonia | 1 (1.0) | 2 (7.1) | 3 (2.3) | 7 (7.1) |
| Sepsis | 2 (2.0) | 0 (0.0) | 2 (1.6) | 3 (3.1) |
| Metabolism and nutrition disorders | | | | |
| Hyponatraemia | 1 (1.0) | 2 (7.1) | 3 (2.3) | 0 (0.0) |
| Respiratory, thoracic and mediastinal disorders | | | | |
| Pleural effusion | 2 (2.0) | 1 (3.6) | 3 (2.3) | 1 (1.0) |
| Cardiac disorders | | | | |
| Atrial fibrillation | 0 (0.0 | 0 (0.0) | 0 (0.0) | 3 (3.1) |

Abbreviations: AE = adverse event; n = number of patients in the category; N = number of patients evaluable; NR = not reported; PT = preferred term; SOC = system organ class
Source: Tam et al., 2020[2]; BeiGene, 2020[46]

TRAEs are summarised by SOC and PT in Table B.2.24. In Cohort 1, the most common TRAEs (reported in ≥10% of patients) among zanubrutinib-treated patients were neutropenia (21.8%), diarrhoea (10.9%), fatigue (10.9%) and contusion (9.9%). As with AEs and Grade ≥3 AEs, the only TRAE more prevalent (>10% higher) in the zanubrutinib arm compared with the ibrutinib treatment arm was neutropenia (21.8% and 11.2%, respectively). TRAEs more prevalent in the ibrutinib arm compared with the zanubrutinib arm were diarrhoea (23.5% versus 10.9%), contusion (22.4% versus 9.9%) and atrial fibrillation (13.3% versus 1.0%). In all zanubrutinib treated patients, the incidences of AEs were generally comparable between Cohorts 1 and 2, except for neutropenia (21.8% versus 10.7%; more prevalent [>10% difference] in Cohort 1).[46]

**Table B.2.24. TRAEs by SOC and PT reported in >5% of patients (Safety Analysis Set)**

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| TRAEs, n (%) | 80 (79.2) | 22 (78.6) | 102 (79.1) | 84 (85.7) |
| Blood and lymphatic system disorders | | | | |
| Neutropenia | 22 (21.8) | 3 (10.7) | 25 (19.4) | 11 (11.2) |
| Thrombocytopenia | 9 (8.9) | 2 (7.1) | 11 (8.5) | 8 (8.2) |
| Anaemia | 6 (5.9) | 1 (3.6) | 7 (5.4) | 4 (4.1) |
| Gastrointestinal disorders | | | | |

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| Diarrhoea | 11 (10.9) | 5 (17.9) | 16 (12.4) | 23 (23.5) |
| Nausea | 7 (6.9) | 0 (0.0) | 7 (5.4) | 7 (7.1) |
| Constipation | 6 (5.9) | 0 (0.0) | 6 (4.7) | 0 (0.0) |
| Injury, poisoning and procedural complications | | | | |
| Contusion | 10 (9.9) | 4 (14.3) | 14 (10.9) | 22 (22.4) |
| General disorders and administration site conditions | | | | |
| Fatigue | 11 (10.9) | 2 (7.1) | 13 (10.1) | 9 (9.2) |
| Skin and subcutaneous tissue disorders | | | | |
| Rash | 8 (7.9) | 2 (7.1) | 10 (7.8) | 11 (11.2) |
| Musculoskeletal and connective tissue disorders | | | | |
| Muscle spasms | 7 (6.9) | 2 (7.1) | 9 (7.0) | 11 (11.2) |
| Arthralgia | 3 (3.0) | 1 (3.6) | 4 (3.1) | 6 (6.1) |
| Respiratory, thoracic and mediastinal disorders | | | | |
| Epistaxis | 7 (6.9) | 1 (3.6) | 8 (6.2) | 14 (14.3) |
| Vascular disorders | | | | |
| Hypertension | 6 (5.9) | 2 (7.1) | 8 (6.2) | 13 (13.3) |
| Infections and infestations | | | | |
| Upper respiratory tract infection | 6 (5.9) | 1 (3.6) | 7 (5.4) | 13 (13.3) |
| Pneumonia | 1 (1.0) | 2 (7.1) | 3 (2.3) | 7 (7.1) |
| Renal and urinary disorders | | | | |
| Haematuria | 4 (4.0) | 1 (3.6) | 5 (3.9) | 8 (8.2) |
| Nervous system disorders | | | | |
| Headache | 3 (3.0) | 0 (0.0) | 3 (2.3) | 7 (7.1) |
| Cardiac disorders | | | | |
| Atrial fibrillation | 1 (1.0) | 0 (0.0) | 1 (0.8) | 13 (13.3) |

Abbreviations: n = number of patients in the category; N = number of patients evaluable; PT = preferred term; SOC = system organ class; TRAE = treatment-related adverse event
Source: BeiGene, 2020[46]

Grade ≥3 TRAEs are summarised by PT and SOC in Table B.2.25. In Cohort 1, Grade ≥3 TRAEs occurred in 32.7% and 42.9% of patients in the zanubrutinib and ibrutinib treatment arms, respectively. Hypertension (9.2%), neutropenia (6.1%), and atrial fibrillation and pneumonia (each 3.1%) were the most commonly reported Grade ≥3 TRAEs in the ibrutinib treatment arm. Neutropenia (13.9%), thrombocytopenia (5.9%) and neutrophil count decreased (4.0%) were the most commonly reported AEs in the zanubrutinib treatment arm. Grade ≥3 TRAEs occurred in 46.4% of patients in Cohort 2.[46]

**Table B.2.25. Grade ≥3 TRAEs by SOC and PT reported in >1 patient (Safety Analysis Set)**

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| Grade ≥3 TRAEs, n (%) | 33 (32.7) | 13 (46.4) | 46 (35.7) | 42 (42.9) |
| Blood and lymphatic system disorders | | | | |
| Neutropenia | 14 (13.9) | 3 (10.7) | 17 (13.2) | 6(6.1) |
| Thrombocytopenia | 6 (5.9) | 1 (3.6) | 7 (5.4) | 2 (2.0) |
| Febrile neutropenia | 3 (3.0) | 0 (0.0) | 3 (2.3) | 0 (0.0) |
| Infections and infestations | | | | |
| Influenza | 2 (2.0) | 0 (0.0) | 2 (1.6) | 0 (0.0) |
| Pneumonia | 0 (0.0) | 1 (3.6) | 1 (0.8) | 3 (3.1) |
| Cardiac disorders | | | | |
| Atrial fibrillation | 0 (0.0) | 0 (0.0) | 0 (0.0) | 3 (3.1) |
| Gastrointestinal disorders | | | | |
| Diarrhoea | 2 (2.0) | 2 (7.1) | 4 (3.1) | 1 (1.0) |

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| Investigations | | | | |
| Neutrophil count decreased | 4 (4.0) | 0 (0.0) | 4 (3.1) | 0 (0.0) |
| Vascular disorders | | | | |
| Hypertension | 2 (2.0) | 2 (7.1) | 4 (3.1) | 9 (9.2) |

Abbreviations: AE = adverse event; n = number of patients in the category; N = number of patients evaluable; PT = preferred term; SOC = system organ class; TRAE = treatment-related adverse event
Source: BeiGene, 2020[46]

#### *B.2.10.1.3 Serious adverse events*

SAEs are summarised by SOC and PT in Table B.2.26. In Cohort 1, the number of SAEs were comparable across both the zanubrutinib and ibrutinib arms (39.6% versus 40.8%). The most common SAE in the ibrutinib arm was pneumonia (9.2%), followed by pyrexia and sepsis (each 3.1%), whereas febrile neutropenia, influenza and neutropenia were the most common SAEs in the zanubrutinib arm (each 3.0%).[46]

In all zanubrutinib-treated patients, the incidences of SAEs were generally comparable between Cohort 1 and Cohort 2. Pneumonia was more common in Cohort 2 compared with Cohort 1 (10.7% versus 1.0%), although Cohort 2 had an older patient population and smaller sample size than Cohort 1.[46]

**Table B.2.26. SAEs by SOC and PT reported in >1 patient (Safety Analysis Set)**

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| SAEs, n (%) | 40 (39.6) | 11 (39.3) | 51 (39.5) | 40 (40.8) |
| Blood and lymphatic system disorders | | | | |
| Febrile neutropenia | 3 (3.0) | 0 (0.0) | 3 (2.3) | 0 (0.0) |
| Neutropenia | 3 (3.0) | 0 (0.0) | 3 (2.3) | 0 (0.0) |
| Anaemia | 2 (2.0) | 0 (0.0) | 2 (1.6) | 0 (0.0) |
| Thrombocytopenia | 2 (2.0) | 0 (0.0) | 2 (1.6) | 0 (0.0) |
| Infections and infestations | | | | |
| Influenza | 3 (3.0) | 0 (0.0) | 3 (2.3) | 1 (1.0) |
| Lower respiratory tract infection | 2 (2.0) | 1 (3.6) | 3 (2.3) | 0 (0.0) |
| Sepsis | 2 (2.0) | 0 (0.0) | 2 (1.6) | 3 (3.1) |
| Pneumonia | 1 (1.0) | 3 (10.7) | 4 (3.1) | 9 (9.2) |
| Urinary tract infection | 0 (0.0) | 0 (0.0) | 1 (0.8) | 2 (2.0) |
| Cellulitis | 0 (0.0) | 2 (7.1) | 2 (1.6) | 0 (0.0) |
| General disorders and administration site conditions | | | | |
| Pyrexia | 2 (2.0) | 1 (3.6) | 3 (2.3) | 3 (3.1) |
| Drug withdrawal syndrome | 1 (1.0) | 1 (3.6) | 2 (1.6) | 0 (0.0) |
| Injury, poisoning and procedural complications | | | | |
| Periorbital haematoma | 1 (1.0) | 1 (3.6) | 2 (1.6) | 0 (0.0) |
| Subdural haemorrhage | 1 (1.0) | 1 (3.6) | 2 (1.6) | 0 (0.0) |
| Neoplasms benign, malignant and unspecified (including cysts and polyps) | | | | |
| Basal cell carcinoma | 2 (2.0) | 0 (0.0) | 2 (1.6) | 0 (0.0) |
| Respiratory, thoracic and mediastinal disorders | | | | |
| Pleural effusion | 2 (2.0) | 0 (0.0) | 2 (1.6) | 1 (1.0) |
| Respiratory failure | 1 (1.0) | 1 (3.6) | 2 (1.6) | 0 (0.0) |
| Cardiac disorders | | | | |
| Pericarditis | 0 (0.0) | 0 (0.0) | 0 (0.0) | 2 (2.0) |
| Hepatobiliary disorders | | | | |

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| Cholecystitis | 0 (0.0) | 0 (0.0) | 0 (0.0) | 2 (2.0) |
| Nervous system disorders | | | | |
| Loss of consciousness | 0 (0.0) | 0 (0.0) | 0 (0.0) | 2 (2.0) |

Abbreviations: n = number of patients in the category; N = number of patients evaluable; NR = not reported; PT = preferred term; SAE = serious adverse event; SOC = system organ class
Source: BeiGene, 2020[46]

#### *B.2.10.1.4 Deaths*

A summary of deaths is presented in Table B.2.27. In Cohort 1, the number of deaths were comparable across both the zanubrutinib and ibrutinib arms (5.9% versus 7.1%). The most common cause of death was disease progression in both treatment arms. The incidence of deaths was higher in Cohort 2 than in Cohort 1 (10.7% versus 5.9%); there was a higher proportion of older patients in Cohort 2 than in Cohort 1.[46]

**Table B.2.27. Summary of deaths (Safety Analysis Set)**

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| All deaths, n (%) | 6 (5.9) | 3 (10.7) | 9 (7.0) | 7 (7.1) |
| Death due to AE | 1 (1.0) | 1 (3.6) | 2 (1.6) | 2 (2.0)* |
| Death due to progressive disease | 3 (3.0) | 1 (3.6) | 4 (3.1) | 3 (3.1)† |
| Death unknown | 1 (1.0) | 0 (0.0) | 1 (0.8) | 2 (2.0)‡ |
| Other | 1 (1.0) | 1 (3.6) | 2 (1.6) § | 0 (0.0) |
| Deaths within 30 days of last dose date | 1 (1.0) | 1 (3.6) | 2 (1.6) | 5 (5.1) |
| Death due to AE | 1 (1.0) | 0 (0.0) | 1 (0.8) | 2 (2.0)[a] |
| Death due to progressive disease | 0 (0.0) | 1 (3.6) | 1 (0.8) | 1 (1.0)b |
| Death unknown | 0 (0.0) | 0 (0.0) | 0 (0.0) | 2 (2.0)c |
| Other | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) |
| Deaths >30 days of last dose date | 5 (5.0) | 2 (7.1) | 7 (5.4) | 2 (2.0) |
| Death due to AE | 0 (0.0) | 1 (3.6) | 1 (0.8) | 0 (0.0) |
| Death due to progressive disease | 3 (3.0) | 0 (0.0) | 3 (2.3) | 2 (2.0) |
| Death unknown | 1 (1.0) | 0 (0.0) | 1 (0.8) | 0 (0.0) |
| Other | 1 (1.0) | 1 (3.6) | 2 (1.6) | 0 (0.0) |

Abbreviations: AE = adverse event; n = number of patients in the category; N = number of patients evaluable
*Does not include two AEs leading to death (PTs of unknown and cardiac failure acute) that were instead attributed to Death unknown and Death due to progressive disease in the Death electronic case report form; †Includes one death due to an AE of cardiac failure acute attributable to cardiac amyloidosis in the context of disease progression; ‡Includes one unexplained death due to an AE of death of unknown origin; §Includes one occurrence each of community acquired pneumonia and died in sleep/sudden death
Source: BeiGene, 2020[46]

Fatal AEs are summarised by SOC and PT in Table B.2.28. Deaths due to AEs occurred in four ibrutinib-treated patients (4.1%) and one zanubrutinib-treated patient (1.0%); all five deaths due to AEs occurred within 30 days of the last dose date. The deaths due to AEs in the ibrutinib arm were due to cause unknown, acute cardia failure, bacterial sepsis and sepsis; the death due to an AE in the zanubrutinib arm was due to cardiomegaly.[46]

**Table B.2.28. AEs leading to death by SOC and PT (Safety Analysis Set)**

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| AEs leading to death, n (%) | 1 (1.0) | 0 (0.0) | 1 (0.8) | 4(4.1) |
| Cardiac disorders | 1 (1.0) | 0 (0.0) | 1 (0.8) | 1 (1.0) |
| Cardiomegaly | 1 (1.0) | 0 (0.0) | 1 (0.8) | 0 (0.0) |
| Cardiac failure acute | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |
| General disorders and administration site conditions | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |
| Death | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |
| Infections and infestations | 0 (0.0) | 0 (0.0) | 0 (0.0) | 2 (2.0) |
| Bacterial sepsis | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |
| Sepsis | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |

Abbreviations: AE = adverse event; n = number of patients in the category; N = number of patients evaluable; NR = not reported; PT = preferred term; SOC = system organ class
Source: BeiGene, 2020[46]

### *B.2.10.1.5 Adverse events of special interest*

#### *B.2.10.1.5.1 Haemorrhage*

Haemorrhage events by PT are summarised in Table B.2.29. In Cohort 1, a higher proportion of patients treated with ibrutinib experienced haemorrhage compared with zanubrutinib. In the ibrutinib treatment arm, 59.2% of patients had haemorrhage, compared with 48.5% in the zanubrutinib treatment arm. Mild or moderate mucocutaneous bleeding were the predominant events reported in the ibrutinib and zanubrutinib treatment arms. In all zanubrutinib-treated patients, the incidence of haemorrhage events was slightly higher in Cohort 1 (48.5%) compared with Cohort 2 (39.3%).[46]

**Table B.2.29. Haemorrhage events by PT (Safety Analysis Set)**

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| Haemorrhage (including minor bleeds involving mucous membranes and skin), n (%) | 49 (48.5) | 11 (39.3) | 60 (46.5) | 58 (59.2) |
| Contusion | 13 (12.9) | 6 (21.4) | 19 (14.7) | 23 (23.5) |
| Epistaxis | 13 (12.9) | 1 (3.6) | 14 (10.9) | 19 (19.4) |
| Haematuria | 7 (6.9) | 1 (3.6) | 8 (6.2) | 10 (10.2) |
| Petechiae | 7 (6.9) | 0 (0.0) | 7 (5.4) | 3 (3.1) |
| Conjunctival haemorrhage | 5 (5.0) | 1 (3.6) | 6 (4.7) | 5 (5.1) |
| Haematoma | 5 (5.0) | 0 (0.0) | 5 (3.9) | 7 (7.1) |
| Angina bullosa haemorrhagic | 3 (3.0) | 0 (0.0) | 3 (2.3) | 0 (0.0) |
| Increased tendency to bruise | 3 (3.0) | 1 (3.6) | 4 (3.1) | 5 (5.1) |
| Purpura | 3 (3.0) | 0 (0.0) | 3 (2.3) | 6 (6.1) |
| Rectal haemorrhage | 3 (3.0) | 0 (0.0) | 3 (2.3) | 1 (1.0) |
| Gingival bleeding | 2 (2.0) | 1 (3.6) | 3 (2.3) | 5 (5.1) |
| Ecchymosis | 1 (1.0) | 1 (3.6) | 2 (1.6) | 4 (4.1) |
| Retinal haemorrhage | 0 (0.0) | 0 (0.0) | 0 (0.0) | 4 (4.1) |
| Haemoptysis | 1 (1.0) | 2 (7.1) | 3 (2.3) | 1 (1.0) |

Abbreviations: n = number of patients in the category; N = number of patients evaluable; NR = not reported; PT = preferred term
Source: Tam, 2020[45]; BeiGene, 2020[46]

Major haemorrhage events (defined as serious or Grade ≥3 bleeding at any site, or central nervous system bleeding of any grade) by PT are summarised in Table B.2.30. In Cohort 1, more major haemorrhages were reported in the ibrutinib treatment arm (9.2%) compared with the zanubrutinib treatment arm (5.9%). The only major haemorrhages reported in >1 patient were haematuria and retinal haemorrhage (each 2%). The incidence of major haemorrhage was similar between Cohort 1 and Cohort 2.[46]

**Table B.2.30. Major haemorrhage events by PT (Safety Analysis Set)**

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| Major haemorrhage, n (%) | 6 (5.9) | 2 (7.1) | 8 (6.2) | 9 (9.2) |
| Eye haemorrhage | 1 (1.0) | 0 (0.0) | 1 (0.8) | 0 (0.0) |
| Gastric haemorrhage | 1 (1.0) | 0 (0.0) | 1 (0.8) | 0 (0.0) |
| Haemothorax | 1 (1.0) | 0 (0.0) | 1 (0.8) | 0 (0.0) |
| Lower gastrointestinal haemorrhage | 1 (1.0) | 0 (0.0) | 1 (0.8) | 0 (0.0) |
| Periorbital haematoma | 1 (1.0) | 1 (3.6) | 2 (1.6) | 0 (0.0) |
| Subdural haemorrhage | 1 (1.0) | 1 (1.0) | 1 (3.6) | 0 (0.0) |
| Tumour haemorrhage | 1 (1.0) | 0 (0.0) | 1 (0.8) | 0 (0.0) |
| Haematuria | 0 (0.0) | 0 (0.0) | 0 (0.0) | 2 (2.0) |
| Haemorrhagic disorder | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |
| Melaena | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |
| Post procedural haemorrhage | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |
| Retinal haemorrhage | 0 (0.0) | 0 (0.0) | 0 (0.0) | 2 (2.0) |
| Subarachnoid haemorrhage | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |
| Subdural haematoma | 0 (0.0) | 1 (3.6) | 1 (0.8) | 1 (1.0) |
| Gastric ulcer haemorrhage | 0 (0.0) | 1 (3.6) | 1 (0.8) | 0 (0.0) |

Abbreviations: n = number of patients in the category; N = number of patients evaluable; NR = not reported; PT = preferred term
Source: BeiGene, 2020[46]

### *B.2.10.1.5.2 Atrial fibrillation/flutter*

Atrial fibrillation/flutter events by PT are summarised in Table B.2.31. A higher proportion of patients treated with ibrutinib had AEs of atrial fibrillation/flutter compared with zanubrutinib. In Cohort 1, 15 (15.3%) patients (14 events [14.3%] of atrial fibrillation, 2 events [2.0%] of atrial flutter) in the ibrutinib treatment arm and 2.0% of patients (both atrial fibrillation) in the zanubrutinib treatment arm reported atrial fibrillation or flutter. The incidence of atrial fibrillation/flutter was similar between Cohort 1 and Cohort 2.[46]

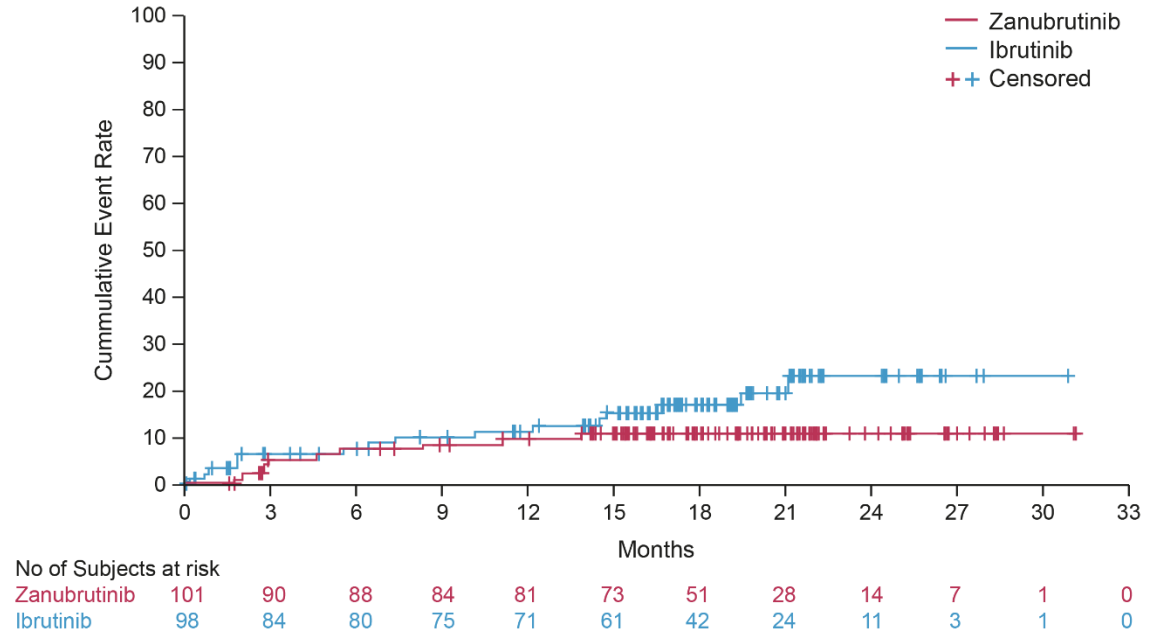
**Table B.2.31. Atrial fibrillation/flutter events by PT (Safety Analysis Set)**

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| Atrial fibrillation | | | | |
| All Grades | 2 (2.0) | 1 (3.6) | 3 (2.3) | 14 (14.3) |
| Grade ≥3 | 0 (0.0) | 0 (0.0) | 0 (0.0) | 3 (3.1) |
| Atrial flutter | | | | |
| All Grades | 0 (0.0) | 0 (0.0) | 0 (0.0) | 2 (2.0) |
| Grade ≥3 | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |

Abbreviations: n = number of patients in the category; N = number of patients evaluable; PT = preferred term
Source: BeiGene, 2020[46]

Risk factors for atrial fibrillation, such as prior history of atrial fibrillation, diabetes or hypertension, were balanced across the study arms in Cohort 1. Of the 12 zanubrutinib-treated patients (Cohort 1 and Cohort 2 combined) with a history of atrial fibrillation/flutter, none had it worsen and become an AE while on treatment. However, 3 of the 8 patients (37.5%) with a history of atrial fibrillation randomised to ibrutinib treatment developed an AE of atrial fibrillation. In addition, the risk of developing atrial fibrillation over time was lower in zanubrutinib-treated patients compared with those treated with ibrutinib (Figure B.2.19).

**Figure B.2.19. KM curve of time to atrial fibrillation or flutter in Cohort 1 (Safety Analysis Set)**




Abbreviations: KM = Kaplan Meier; No. = number
Note: Cohort 1 includes patients with activating mutations in *MYD88*
Source: BeiGene, 2020[46]

### *B.2.10.1.5.3 Hypertension*

Hypertension events by PT are summarised in Table B.2.32. The ibrutinib treatment arm had a higher proportion of patients with hypertension than the zanubrutinib treatment arm, particularly Grade ≥3 hypertension. In Cohort 1, 17.3% of patients in the ibrutinib treatment arm and 10.9% of patients in the zanubrutinib treatment arm experienced hypertension. Grade ≥3 hypertension was reported for 12.2% of ibrutinib-treated patients and 5.9% of zanubrutinib-treated patients.[2, 46]

**Table B.2.32. Hypertension events by PT (Safety Analysis Set)**

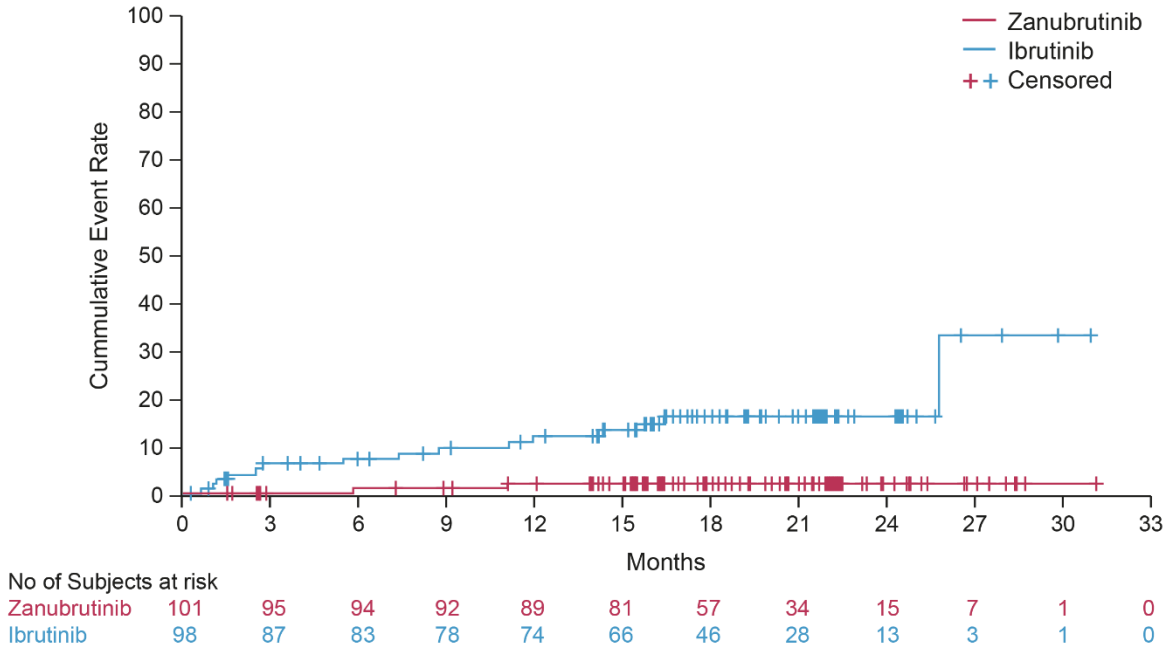
| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| Hypertension (any grade) | | | | |
| Hypertension | 11 (10.9) | 3 (10.7) | 14 (10.9) | 16 (16.3) |
| Blood pressure increased | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |
| Hypertension (Grade ≥3) | | | | |
| Hypertension | 6 (5.9) | 3 (10.7) | 9 (7.0) | 11 (11.2) |
| Blood pressure increased | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |

Abbreviations: n = number of patients in the category; N = number of patients evaluable; PT = preferred term
Source: Tam, 2020[45]; BeiGene, 2020[46]

Risk factors for hypertension were generally balanced across the study arms including prior hypertension, type 2 diabetes and hypercholesterolemia. In patients reporting hypertension, 7/11 (63.6%) patients in the zanubrutinib arm had a history of hypertension at baseline compared with 6/17 (35.3%) indicating that there were fewer new cases of hypertension in zanubrutinib-treated patients than ibrutinib-treated patients (4 cases vs 11 cases). The risk of

developing hypertension early on was comparable between the treatment arms but became higher over time in patients in the ibrutinib arm compared with the zanubrutinib arm (Figure B.2.20). The incidence of hypertension was similar between Cohort 1 and Cohort 2.[46]

**Figure B.2.20. KM curve of time to hypertension in Cohort 1 (Safety Analysis Set)**




Abbreviations: KM = Kaplan Meier
Cohort 1 includes patients with activating mutations in *MYD88*
Source: BeiGene, 2020[46]

### *B.2.10.1.5.4 Second primary malignancy*

Second primary malignancy events by PT are summarised in Table B.2.33. The rate of second primary malignancies was comparable between the ibrutinib and zanubrutinib treatment arms (11.2% and 11.9%, respectively). In both treatment arms of Cohort 1, skin cancers (basal cell carcinoma, squamous cell carcinoma of skin, Bowen's disease, skin cancer, malignant melanoma) comprised the majority of second primary malignancies (ibrutinib 9.2%; zanubrutinib 7.9%). The incidence of second primary malignancies was similar between Cohort 1 and Cohort 2.[46]

**Table B.2.33. Second primary malignancy events by PT (Safety Analysis Set)**

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| Second Primary Malignancy (Malignant tumours SMQ), n (%) | 12 (11.9) | 4 (14.3) | 16 (12.4) | 11 (11.2) |
| Basal cell carcinoma | 4 (4.0) | 3 (10.7) | 7 (5.4) | 2 (2.0) |
| Squamous cell carcinoma of skin | 2 (2.0) | 0 (0.0) | 2 (1.6) | 4 (4.1) |
| Bowen's disease | 1 (1.0) | 0 (0.0) | 1 (0.8) | 1 (1.0) |
| Chronic myelomonocytic leukaemia | 1 (1.0) | 0 (0.0) | 1 (0.8) | 0 (0.0) |
| Colorectal cancer metastatic | 1 (1.0) | 0 (0.0) | 1 (0.8) | 0 (0.0) |
| Endometrial adenocarcinoma | 1 (1.0) | 0 (0.0) | 1 (0.8) | 0 (0.0) |
| Lung neoplasm malignant | 1 (1.0) | 0 (0.0) | 1 (0.8) | 0 (0.0) |
| Malignant melanoma | 1 (1.0) | 0 (0.0) | 1 (0.8) | 0 (0.0) |

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| Malignant melanoma stage I | 1 (1.0) | 0 (0.0) | 1 (0.8) | 0 (0.0) |
| Plasma cell myeloma | 1 (1.0) | 0 (0.0) | 1 (0.8) | 0 (0.0) |
| Skin cancer | 1 (1.0) | 0 (0.0) | 1 (0.8) | 2 (2.0) |
| Bladder transitional cell carcinoma | 0 (0.0) | 0 (0.0) | 0 (0.0) | 2 (2.0) |
| Chronic myeloid leukaemia | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |
| Queyrat erythroplasia | 0 (0.0) | 1 (3.6) | 1 (0.8) | 0 (0.0) |

Abbreviations: n = number of patients in the category; N = number of patients evaluable; PT = preferred term; SMQ = Standardized Medical Dictionary for Regulatory Activities Query
Source: BeiGene, 2020[46]

#### *B.2.10.1.5.5 Tumour lysis syndrome*

No adverse events for tumour lysis syndrome were reported.[46]

#### *B.2.10.1.5.6 Infections*

Infection events by PT are summarised in Table B.2.34. In Cohort 1, infections were among the most common AEs reported in both treatment arms, with similar incidences (67.3% and 66.3% in the ibrutinib and zanubrutinib arms, respectively). The majority of these AEs were mucosal infections involving the sinopulmonary and urinary tracts. The incidence of infections was slightly higher in Cohort 2 compared with Cohort 1 (75.0% and 66.3%, respectively).[46]

**Table B.2.34. Infection events by PT in ≥5% of patients (Safety Analysis Set)**

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| Infections (any grade), n (%) | 67 (66.3) | 21 (75.0) | 88 (68.2) | 66 (67.3) |
| Upper respiratory tract infection | 24 (23.8) | 6 (21.4) | 30 (23.3) | 28 (28.6) |
| Nasopharyngitis | 11 (10.9) | 2 (7.1) | 13 (10.1) | 7 (7.1) |
| Urinary tract infection | 10 (9.9) | 4 (14.3) | 14 (10.9) | 10 (10.2) |
| Lower respiratory tract infection | 8 (7.9) | 2 (7.1) | 10 (7.8) | 9 (9.2) |
| Respiratory tract infection | 6 (5.9) | 5 (17.9) | 11 (8.5) | 2 (2.0) |
| Influenza | 5 (5.0) | 0 (0.0) | 5 (3.9) | 1 (1.0) |
| Rhinitis | 5 (5.0) | 1 (3.6) | 6 (4.7) | 4 (4.1) |
| Sinusitis | 5 (5.0) | 0 (0.0) | 5 (3.9) | 7 (7.1) |
| Cellulitis | 4 (4.0) | 2 (7.1) | 6 (4.7) | 6 (6.1) |
| Gastroenteritis | 2 (2.0) | 2 (7.1) | 4 (3.1) | 5 (5.1) |
| Pneumonia | 2 (2.0) | 4 (14.3) | 6 (4.7) | 12 (12.2) |
| Conjunctivitis | 1 (1.0) | 1 (3.6) | 2 (1.6) | 6 (6.1) |
| Localised infection | 1 (1.0) | 1 (3.6) | 2 (1.6) | 7 (7.1) |
| Oral herpes | 1 (1.0) | 1 (3.6) | 2 (1.6) | 5 (5.1) |
| Herpes zoster | 4 (4.0) | 3 (10.7) | 7 (5.4) | 1 (1.0) |
| Infections (Grade ≥3) | 18 (17.8) | 8 (28.6) | 26 (20.2) | 19 (19.4) |
| Pneumonia | 1 (1.0) | 2 (7.1) | 3 (2.3) | 7 (7.1) |

Abbreviations: n = number of patients in the category; N = number of patients evaluable; PT = preferred term
Source: Tam, 2020[45]; BeiGene, 2020[46]

#### *B.2.10.1.5.7 Cytopenias*

Cytopenia events by PT are summarised in Table B.2.35. In Cohort 1, anaemia was reported in 10.2% of patients in the ibrutinib treatment arm and 11.9% in the zanubrutinib arm.[45]

However, the zanubrutinib treatment arm overall had a higher incidence of anaemia at baseline compared with ibrutinib (haemoglobin ≤110 g/L 65.7% and 53.5%, respectively).[45] In all zanubrutinib-treated patients, the incidence of anaemia was higher in Cohort 2 compared with Cohort 1 (21.4% and 11.9%, respectively).[45, 46]

A higher proportion of patients in the zanubrutinib treatment arm reported neutropenia compared with ibrutinib. In Cohort 1, neutropenia was reported in 13.3% of patients in the ibrutinib treatment arm and 29.7% in the zanubrutinib arm.[45] Despite the higher frequency of neutropenia reported with zanubrutinib treatment, the incidence of serious infections was similar between the ibrutinib and zanubrutinib treatment arms (19.4% and 14.9%, respectively) as was the incidence of Grade ≥3 infections (19.4% and 17.8%, respectively). The incidence of neutropenia was higher in Cohort 1 compared with Cohort 2 (29.7% and 17.9%, respectively).[45, 46]

In Cohort 1, thrombocytopenia was reported in 12.2% of patients in the ibrutinib treatment arm and 9.9% in the zanubrutinib treatment arm. The incidence of thrombocytopenia was similar between Cohort 1 and Cohort 2.[46]

**Table B.2.35. Cytopenia events by PT (Safety Analysis Set)**

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| Cytopenias (any grade), n (%) | | | | |
| Anaemia | 12 (11.9) | 6 (21.4) | 18 (14.0) | 10 (10.2) |
| Neutropoenia | 30 (29.7) | 5 (17.9) | 35 (27.1) | 13 (13.3) |
| Thrombocytopenia | 10 (9.9) | 3 (10.7) | 13 (10.1) | 12 (12.2) |
| Cytopenias (Grade ≥3), n (%) | | | | |
| Anaemia | 5 (5.1) | 3 (10.7) | 8 (6.2) | 5 (5.0) |
| Neutropoenia | 20 (19.8) | 3 (10.7) | 23 (17.8) | 8 (8.2) |
| Thrombocytopenia | 6 (5.9%) | 2 (7.1) | 8 (6.2) | 3 (3.1) |

Abbreviations: n = number of patients in the category; N = number of patients evaluable; PT = preferred term
Source: Tam, 2020[45]; BeiGene, 2020[46]

### *B.2.10.1.6 Adverse events leading to treatment discontinuation*

AEs leading to treatment discontinuation by SOC and PT are summarised in Table B.2.36. In Cohort 1, more AEs leading to discontinuation of study treatment were reported in the ibrutinib treatment arm compared with the zanubrutinib treatment arm (9.2% and 4.0%, respectively). Five patients in the ibrutinib treatment arm had AEs leading to study treatment discontinuation that were assessed as related to ibrutinib (drug-induced liver injury, hepatitis, interstitial lung disease, pneumonia, and pneumonitis). Two patients in the zanubrutinib treatment arm had AEs leading to study treatment discontinuation that were assessed as related to zanubrutinib (neutropenia and cardiomegaly). The incidences of AEs leading to discontinuation of study drug was similar between Cohort 1 and Cohort 2.[46]

**Table B.2.36. AEs leading to treatment discontinuation by SOC and PT reported in ≥1% of patients (Safety Analysis Set)**

| Event | Zanubrutinib | | | Ibrutinib (N=98) |
|---|---|---|---|---|
| | Cohort 1 (N=101) | Cohort 2 (N=28) | Total (N=129) | |
| AEs leading to treatment discontinuation, n (%) | 4 (4.0) | 2 (7.1) | 6 (4.7) | 9 (9.2) |
| Blood and lymphatic system disorders | | | | |
| Neutropenia | 1 (1.0) | 0 (0.0) | 1 (0.8) | 0 (0.0) |
| Cardiac disorders | | | | |
| Cardiomegaly | 1 (1.0) | 0 (0.0) | 1 (0.8) | 0 (0.0) |
| Acute myocardial infarction | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |
| Injury, poisoning and procedural complications | | | | |
| Subdural haemorrhage | 1 (1.0) | 1 (3.6) | 2 (1.6) | 0 (0.0) |
| Neoplasms benign, malignant and unspecified (including cysts and polyps) | | | | |
| Plasma cell myeloma | 1 (1.0) | 0 (0.0) | 1 (0.8) | 0 (0.0) |
| General disorders and administration site conditions | | | | |
| Death | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |
| Hepatobiliary disorders | | | | |
| Drug-induced liver injury | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |
| Hepatitis | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |
| Infections and infestations | | | | |
| Bacterial sepsis | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |
| Pneumonia | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |
| Sepsis | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |
| Respiratory, thoracic and mediastinal disorders | | | | |
| Interstitial lung disease | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |
| Pneumonitis | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (1.0) |
| Gastrointestinal disorders | | | | |
| Diarrhoea | 0 (0.0) | 1 (3.6) | 1 (0.8) | 0 (0.0) |

Abbreviations: AE = adverse event; n = number of patients in the category; N = number of patients evaluable; PT = preferred term; SOC = system organ class
Source: BeiGene, 2020[46]

### B.2.10.2 All Waldenström's macroglobulinaemia patients

Pooled safety data for all WM patients (N=253) comprised all patients with WM treated with zanubrutinib from the following studies:

- BGB-3111-302 (ASPEN; N=129): Phase 3, open-label study comparing the efficacy and safety of zanubrutinib versus ibrutinib in patients with relapsed/refractory or treatment-naïve WM (see Section B.2.6.1)
- BGB-3111-210 (N=44): Phase 2, single-arm study in Chinese patients with relapsed/refractory WM
- BGB-3111-AU-003 (N=78): Phase 1/2 study in patients with various B-cell malignancies, including patients with relapsed/refractory and treatment-naïve WM (see Appendix L)
- BGB-3111-1002 (N=2): Phase 1 dose comparison study in Chinese patients with B-cell malignancies, including 2 with WM.[55]

A summary of methodology is provided in Appendix F.

#### *B.2.10.2.1 Extent of exposure*

The median duration of exposure in all WM patients was 19.68 months (range: 0.6, 57.2 months; mean [SD] 20.56 [11.008]); 83% of patients had at least 12 months of exposure and 12.6% had ≥36 months of exposure. The total exposure was 5,202.46 patient-months. The median relative dose intensity was 98.19% (range: 17.9% to 316.7%).[55]

#### *B.2.10.2.2 Adverse events*

An overview of AEs in all zanubrutinib-treated patients with WM is presented in Table B.2.37. At least one AE of any grade was reported by 97.2% of patients (61.7% Grade ≥3 AE; 44.3% SAEs). Both Grade ≥3 and SAE frequencies were similar to those reported in ASPEN. AEs leading to death and treatment discontinuation were reported in 3.2% and 8.7%, of patients, respectively.[55]

**Table B.2.37. Overview of AEs**

| Event | All WM (N=253) |
|---|---|
| AEs, n (%) | 246 (97.2) |
| Grade ≥3 | 156 (61.7) |
| SAEs | 112 (44.3) |
| AEs leading to death | 8 (3.2) |
| AEs leading to discontinuation | 22 (8.7) |
| TRAEs, n (%) | 203 (80.2) |
| AESIs, n (%) | 230 (90.9) |
| Grade ≥3 | 134 (53.0) |
| Serious | 83 (32.8) |

Abbreviations: AE = adverse event; AESI = adverse event of special interest; n = number of patients in the category; N = number of patients evaluable; SAE = serious adverse event; TRAE = treatment-related adverse event; WM = Waldenström's macroglobulinaemia
Source: BeiGene, 2020[55]

AEs are summarised by SOC and PT in Table B.2.38. The most frequent AEs in the All WM group were upper respiratory tract infection (32.4%) and diarrhoea (21.7%).[55]

**Table B.2.38. AEs by SOC and PT reported in >10% of patients (Safety Analysis Set)**

| Event | All WM (N=253) |
|---|---|
| AEs, n (%) | 246 (97.2) |
| Infections and infestations | 194 (76.7) |
| Upper respiratory tract infection | 82 (32.4) |
| Urinary tract infection | 37 (14.6) |
| Pneumonia | 15 (5.9) |
| Nasopharyngitis | 25 (9.9) |
| Lung infection | 10 (4.0) |
| Gastrointestinal disorders | 148 (58.5) |
| Diarrhoea | 55 (21.7) |
| Constipation | 33 (13.0) |
| Nausea | 27 (10.7) |
| Vomiting | 19 (7.5) |
| Skin and subcutaneous tissue disorders | 136 (53.8) |
| Rash | 36 (14.2) |
| Purpura | 16 (6.3) |
| Investigations | 77 (30.4) |
| Neutrophil count decreased | 36 (14.2) |

| Event | All WM (N=253) |
|---|---|
| Platelet count decreased | 14 (5.5) |
| White blood cell count decreased | 13 (5.1) |
| Alanine aminotransferase increased | 7 (2.8) |
| Respiratory, thoracic and mediastinal disorders | 109 (43.1) |
| Cough | 42 (16.6) |
| Epistaxis | 28 (11.1) |
| Dyspnoea | 18 (7.1) |
| General disorders and administration site conditions | 106 (41.9) |
| Fatigue | 38 (15.0) |
| Pyrexia | 32 (12.6) |
| Oedema peripheral | 20 (7.9) |
| Musculoskeletal and connective tissue disorders | 113 (44.7) |
| Arthralgia | 34 (13.4) |
| Back pain | 34 (13.4) |
| Muscle spasms | 18 (7.1) |
| Pain in extremity | 18 (7.1) |
| Metabolism and nutrition disorders | 75 (29.6) |
| Hypokalaemia | 8 (3.2) |
| Hyperglycaemia | 8 (3.2) |
| Hyperuricaemia | 10 (4.0) |
| Injury, poisoning and procedural complications | 100 (39.5) |
| Contusion | 45 (17.8) |
| Blood and lymphatic system disorders | 91 (36.0) |
| Anaemia | 37 (14.6) |
| Neutropenia | 42 (16.6) |
| Thrombocytopenia | 20 (7.9) |
| Nervous system disorders | 90 (35.6) |
| Headache | 35 (13.8) |
| Dizziness | 24 (9.5) |
| Renal and urinary disorders | 43 (17.0) |
| Haematuria | 18 (7.1) |
| Vascular disorders | 49 (19.4) |
| Hypertension | 30 (11.9) |
| Cardiac disorders | 45 (17.8) |
| Atrial fibrillation | 7 (2.8) |

Abbreviations: AE = adverse event; n = number of patients in the category; N = number of patients evaluable; PT = preferred term; SOC = system organ class; WM = Waldenström's macroglobulinaemia
Source: BeiGene, 2020[55]

Grade ≥3 AEs are summarised by SOC and PT in Table B.2.39. The most frequent Grade ≥3 AEs were neutropenia (11.1%), decreased neutrophil count (8.7%), anaemia (7.1%), and hypertension (5.1%).[55]

**Table B.2.39. Grade ≥3 AEs by SOC and PT reported in ≥3% patients (Safety Analysis Set)**

| Event | All WM (N=253) |
|---|---|
| AEs Grade ≥3, n (%) | 156 (61.7) |
| Infections and infestations | 66 (26.1) |
| Pneumonia | 10 (4.0) |
| Lung infection | 8 (3.2) |
| Upper respiratory tract infection | 3 (1.2) |
| Sepsis | 3 (1.2) |
| Blood and lymphatic system disorders | 52 (20.6) |
| Neutropenia | 28 (11.1) |
| Anaemia | 18 (7.1) |

| Event | All WM (N=253) |
|---|---|
| Thrombocytopenia | 10 (4.0) |
| Febrile neutropenia | 6 (2.4) |
| Investigations | 29 (11.5) |
| Neutrophil count decreased | 22 (8.7) |
| Platelet count decreased | 9 (3.6) |
| White blood cell count decreased | 5 (2.0) |
| Metabolism and nutrition disorders | 12 (4.7) |
| Hypokalaemia | 2 (0.8) |
| Gastrointestinal disorders | 18 (7.1) |
| Diarrhoea | 7 (2.8) |
| Vascular disorders | 14 (5.5) |
| Hypertension | 13 (5.1) |
| Musculoskeletal and connective tissue disorders | 14 (5.5) |
| Arthralgia | 5 (2.0) |
| Back pain | 4 (1.6) |
| Nervous system disorders | 14 (5.5) |
| Syncope | 7 (2.8) |
| Cardiac disorders | 8 (3.2) |
| Atrial fibrillation | 1 (0.4) |

Abbreviations: AE = adverse event; n = number of patients in the category; N = number of patients evaluable; PT = preferred term; SOC = system organ class; WM = Waldenström's macroglobulinaemia
Source: BeiGene, 2020[55]

TRAEs are presented by SOC and PT in Table B.2.40. The most common TRAEs of any grade were neutropenia (13.8%), decreased neutrophil count (13.4%), and contusion (13.0%).[55]

**Table B.2.40. TRAEs by SOC and PT reported in ≥10% of patients (Safety Analysis Set)**

| Event | All WM (N=253) |
|---|---|
| TRAEs, n (%) | 203 (80.2) |
| Skin and subcutaneous tissue disorders | 72 (28.5) |
| Rash | 21 (8.3) |
| Purpura | 15 (5.9) |
| Investigations | 47 (18.6) |
| Neutrophil count decreased | 34 (13.4) |
| Platelet count decreased | 13 (5.1) |
| White blood cell count decreased | 12 (4.7) |
| Alanine aminotransferase increased | 4 (1.6) |
| Infections and infestations | 67 (26.5) |
| Upper respiratory tract infection | 14 (5.5) |
| Lung infection | 5 (2.0) |
| Gastrointestinal disorders | 62 (24.5) |
| Diarrhoea | 25 (9.9) |
| Blood and lymphatic system disorders | 61 (24.1) |
| Neutropenia | 35 (13.8) |
| Anaemia | 12 (4.7) |
| Injury, poisoning and procedural complications | 38 (15.0) |
| Contusion | 33 (13.0) |
| General disorders and administration site conditions | 40 (15.8) |
| Fatigue | 20 (7.9) |
| Respiratory, thoracic and mediastinal disorders | 30 (11.9) |
| Epistaxis | 16 (6.3) |
| Renal and urinary disorders | 13 (5.1) |
| Haematuria | 11 (4.3) |

| Event | All WM (N=253) |
|---|---|
| Musculoskeletal and connective tissue disorders | 25 (9.9) |
| Muscle spasms | 10 (4.0) |
| Vascular disorders | 22 (8.7) |
| Hypertension | 16 (6.3) |
| Cardiac disorders | 19 (7.5) |
| Atrial fibrillation | 3 (1.2) |

Abbreviations: n = number of patients in the category; N = number of patients evaluable; PT = preferred term; SOC = system organ class; TRAE = treatment-related adverse event; WM = Waldenström's macroglobulinaemia
Source: BeiGene, 2020[55]

#### *B.2.10.2.3 Serious adverse events*

SAEs are presented by SOC and PT in Table B.2.41. In total, 44.3% of patients reported at least one SAE. The most frequent SAEs were pneumonia (4.7%), lung infection and cellulitis (2.8% each), and febrile neutropenia and pyrexia (2.0% each).[55]

**Table B.2.41. SAEs by SOC and PT reported in ≥2 patients (Safety Analysis Set)**

| Event | All WM (N=253) |
|---|---|
| SAEs, n (%) | 112 (44.3) |
| Infections and infestations | 56 (22.1) |
| Pneumonia | 12 (4.7) |
| Lung infection | 7 (2.8) |
| Cellulitis | 7 (2.8) |
| Urinary tract infection | 1 (0.4) |
| Lower respiratory tract infection | 3 (1.2) |
| Upper respiratory tract infection | 3 (1.2) |
| Influenza | 3 (1.2) |
| Sepsis | 3 (1.2) |
| Gastrointestinal disorders | 11 (4.3) |
| Respiratory, thoracic and mediastinal disorders | 13 (5.1) |
| Pleural effusion | 4 (1.6) |
| Blood and lymphatic system disorders | 14 (5.5) |
| Anaemia | 4 (1.6) |
| Febrile neutropenia | 5 (2.0) |
| Neutropenia | 4 (1.6) |
| Thrombocytopenia | 2 (0.8) |
| General disorders and conditions | 10 (4.0) |
| Pyrexia | 5 (2.0) |
| Neoplasms benign, malignant (incl cysts and polyps) | 15 (5.9) |
| Basal cell carcinoma | 3 (1.2) |
| Cardiac disorders | 10 (4.0) |
| Atrial fibrillation | 2 (0.8) |
| Nervous system disorders | 9 (3.6) |

Abbreviations: n = number of patients in the category; N = number of patients evaluable; PT = preferred term; SAE = serious adverse event; SOC = system organ class; WM = Waldenström's macroglobulinaemia
Source: BeiGene, 2020[55]

### *B.2.10.2.4 Deaths*

A summary of deaths is presented in Table B.2.42. Overall, 9.1% of patients died during the studies, most commonly from AEs (4.0%) and progressive disease (3.2%). Six deaths occurred within 30 days of last treatment; (2.0% due to AEs).[55]

**Table B.2.42. Summary of deaths (Safety Analysis Set)**

| Event | All WM (N=253) |
|---|---|
| All deaths, n (%) | 23 (9.1) |
| Progressive disease | 8 (3.2) |
| AE | 10 (4.0) |
| Unknown | 3 (1.2) |
| Other | 2 (0.8) |
| Deaths within 30 days of last dose date | 6 (2.4) |
| AE | 5 (2.0) |
| Progressive disease | 1 (0.4) |
| Unknown | 0 (0.0) |
| Other | 0 (0.0) |
| Deaths >30 days of last dose date | 17 (6.7) |
| Progressive disease | 7 (2.8) |
| AE | 5 (2.0) |
| Unknown | 3 (1.2) |
| Other | 2 (0.8) |

Abbreviations: AE = adverse event; n = number of patients in the category; N = number of patients evaluable; WM = Waldenström's macroglobulinaemia
Source: BeiGene, 2020[55]

Fatal AEs are summarised by PT in Table B.2.43. Grade 5 AEs were reported for 8 patients (3.2%); no events were reported in >1 subject each.[55]

**Table B.2.43. AEs leading to death by PT (Safety Analysis Set)**

| Event | All WM (N=253) |
|---|---|
| AEs leading to death, n (%) | 8 (3.2) |
| Death | 1 (0.4) |
| Multiple organ dysfunction syndrome | 1 (0.4) |
| Abdominal sepsis | 1 (0.4) |
| Acute hepatitis B | 1 (0.4) |
| Adenocarcinoma gastric | 1 (0.4) |
| Arthritis bacterial | 1 (0.4) |
| Bronchiectasis | 1 (0.4) |
| Cardiomegaly | 1 (0.4) |
| Scedosporium infection | 1 (0.4) |

Abbreviations: AE = adverse event; n = number of patients in the category; N = number of patients evaluable; PT = preferred term; SOC = system organ class; WM = Waldenström's macroglobulinaemia
Source: BeiGene, 2020[55]

### *B.2.10.2.5 Adverse events of special interest*

AESIs are summarised by category in Table B.2.44. In total, 90.9% of zanubrutinib-treated patients reported at least one AESI. AEs within the categories of infections (76.7%), haemorrhage (52.2%), and neutropenia (30.0%) were reported most frequently. Events that

met the criteria for seriousness and/or were Grade ≥3 were reported in 32.8% and 53.0% of patients, respectively.[55]

**Table B.2.44. AESI by category (Safety Analysis Set)**

| Event | All WM (N=253) |
|---|---|
| AESI, n (%) | 230 (90.9) |
| Serious AESI, n (%) | 83 (32.8) |
| Grade ≥3 AESI, n (%) | 134 (53.0) |
| Anaemia | 37 (14.6) |
| Serious | 4 (1.6) |
| Grade ≥3 | 18 (7.1) |
| Atrial fibrillation and flutter | 7 (2.8) |
| Serious | 2 (0.8) |
| Grade ≥3 | 1 (0.4) |
| Haemorrhage (inclusive of major haemorrhage) | 132 (52.2) |
| Major haemorrhage | 14 (5.5) |
| Serious | 11 (4.3) |
| Grade ≥3 | 14 (5.5) |
| Hypertension | 31 (12.3) |
| Serious | 0 (0.0) |
| Grade ≥3 | 13 (5.1) |
| Infections | 194 (76.7) |
| Serious | 56 (22.1) |
| Grade ≥3 | 66 (26.1) |
| Opportunistic infections | 5 (2.0) |
| Serious | 3 (1.2) |
| Grade ≥3 | 3 (1.2) |
| Neutropenia | 76 (30.0) |
| Serious | 9 (3.6) |
| Grade ≥3 | 50 (19.8) |
| Second primary malignancies (inclusive of skin cancers) | 38 (15.0) |
| Serious | 14 (5.5) |
| Grade ≥3 | 15 (5.9) |
| Skin cancers | 25 (9.9) |
| Serious | 4 (1.6) |
| Grade ≥3 | 4 (1.6) |
| Thrombocytopenia | 33 (13.0) |
| Serious | 3 (1.2) |
| Grade ≥3 | 18 (7.1) |
| Tumour lysis syndrome | 0 (0.0) |
| Serious | 0 (0.0) |
| Grade ≥3 | 0 (0.0) |

Abbreviations: AESI = adverse event of special interest; n = number of patients in the category; N = number of patients evaluable; WM = Waldenström's macroglobulinaemia
Source: BeiGene, 2020[55]

#### *B.2.10.2.5.1 Haemorrhage*

In the All WM group, 132 patients (52.2%) reported at least one haemorrhage event. The most frequently reported events were petechiae/purpura/contusion (29.2%) and epistaxis (11.1%). Serious and Grade ≥3 events were reported in 11 (4.3%) and 14 (5.5%) patients, respectively. Haemorrhage events leading to treatment discontinuation were reported in four patients.[55]

Major haemorrhage occurred in 14 patients (5.5%); the most frequently reported events were haemothorax, periorbital hematoma, and subdural haemorrhage (n=2 each). Events that led

to treatment discontinuation were reported in four patients (subdural haemorrhage in two patients and haematuria and purpura in one patient each).[55]

#### *B.2.10.2.5.2 Atrial fibrillation/flutter*

In the All WM group, seven patients (2.8%) reported at least one occurrence of atrial fibrillation. Serious and Grade ≥3 events were reported in two patients and one patient, respectively.[55]

#### *B.2.10.2.5.3 Hypertension*

In the All WM group, 31 (12.3%) patients reported treatment-emergent hypertension. Grade ≥3 events were reported in 13 patients (5.1%), none of which were serious.[55]

#### *B.2.10.2.5.4 Second primary malignancy*

Second primary malignancy events are summarised in Table B.2.45. Overall, 15.0% of patients reported second primary malignancies, most of which (9.9%) were skin cancers. The most frequently reported events were basal cell carcinoma (6.3%), squamous carcinoma of the skin (3.6%), and Bowen's disease (1.6%). Serious and Grade ≥3 events were reported in 14 (5.5%) and 15 (5.9%) patients, respectively. One patient in this group died from complications of gastric adenocarcinoma.[55]

**Table B.2.45. Second primary malignancy events reported in >1 patient (Safety Analysis Set)**

| Event | All WM (N=253) |
|---|---|
| AESI of second primary malignancies, n (%) | 38 (15.0) |
| Basal cell carcinoma | 16 (6.3) |
| Squamous cell carcinoma of skin | 9 (3.6) |
| Malignant melanoma | 1 (0.4) |
| Squamous cell carcinoma of head and neck | 2 (0.8) |
| Bowen's disease | 4 (1.6) |
| Prostate cancer | 2 (0.8) |
| Skin cancer | 2 (0.8) |
| Adenocarcinoma gastric | 1 (0.4) |
| Breast cancer | 1 (0.4) |
| External ear neoplasm malignant | 1 (0.4) |
| Lung neoplasm malignant | 2 (0.8) |
| Squamous cell carcinoma | 1 (0.4) |

Abbreviations: AESI = adverse event of special interest; n = number of patients in the category; N = number of patients evaluable; WM = Waldenström's macroglobulinaemia
Source: BeiGene, 2020[55]

#### *B.2.10.2.5.5 Tumour lysis syndrome*

No adverse events for tumour lysis syndrome were reported.[55]

#### *B.2.10.2.5.6 Infections*

Infection events are presented in Table B.2.46. In the All WM group, 76.7% of patients reported at least one infection. The most frequent infections were upper respiratory tract infection (32.4%), urinary tract infection (14.6%), nasopharyngitis (9.9%), lower respiratory tract infection (7.1%), cellulitis (6.3%), pneumonia (5.9%), and sinusitis (5.1%).

Serious and Grade ≥3 infections were reported in 56 (22.1%) and 66 (26.1%) patients, respectively. Four patients died from infectious complications (abdominal sepsis, bacterial arthritis, scedosporium infection, and acute hepatitis B). Infections led to treatment discontinuation in five patients (2.0%).[55]

**Table B.2.46. Infection events reported in ≥5% of patients (Safety Analysis Set)**

| Event | All WM (N=253) |
|---|---|
| Infections, n (%) | 194 (76.7) |
| Upper respiratory tract infection | 82 (32.4) |
| Urinary tract infection | 37 (14.6) |
| Pneumonia | 15 (5.9) |
| Nasopharyngitis | 25 (9.9) |
| Lung infection | 10 (4.0) |
| Sinusitis | 13 (5.1) |
| Lower respiratory tract infection | 18 (7.1) |
| Cellulitis | 16 (6.3) |
| Skin infection | 7 (2.8) |
| Oral herpes | 5 (2.0) |
| Localised infection | 8 (3.2) |
| Conjunctivitis | 6 (2.4) |
| Influenza | 8 (3.2) |
| Gastroenteritis | 6 (2.4) |
| Pharyngitis | 3 (1.2) |
| Respiratory tract infection | 12 (4.7) |
| Rhinitis | 6 (2.4) |

Abbreviations: n = number of patients in the category; N = number of patients evaluable; WM = Waldenström's macroglobulinaemia
Source: BeiGene, 2020[55]

#### *B.2.10.2.5.7 Cytopenias*

In the All WM group, 64.0% of patients were anaemic at baseline. A total of 37 patients (14.6%) reported at least one occurrence of treatment-emergent anaemia; Grade ≥3 events were reported in 7.1% of patients. A total of 17 of 37 (45.9%) patients with treatment-emergent anaemia received red blood cell transfusion within 30 days of onset.[55]

At baseline, 13.0% of patients were neutropenic. A total of 76 patients (30.0%) reported at least one occurrence of treatment-emergent neutropenia. Grade ≥3 and serious events were reported in 19.8% and 3.6% of patients, respectively. AEs leading to treatment discontinuation were reported in one patient. A total of 36 of 76 (47.4%) neutropenic patients received granulocyte colony stimulating factor within 30 days of onset.[55]

At baseline, 13.4% of patients were thrombocytopenic. Thirty-three (13.0%) patients reported at least one occurrence of treatment-emergent thrombocytopenia; Grade ≥3 AEs were reported in 7.1% of patients and SAEs were reported in 3 patients (1.2%). Four of 33 (12.1%) thrombocytopenic patients received platelet transfusions within 30 days of onset.[55]

### *B.2.10.2.6 Adverse events leading to treatment discontinuation*

In total, 22 (8.7%) patients reported events leading to treatment discontinuation; the only event that led to treatment discontinuation in >1 patient was subdural haemorrhage (n=2).[55]

### B.2.10.3 Safety conclusions

Zanubrutinib has a favourable safety and tolerability profile compared with ibrutinib, with a numerically lower rate of several AEs, such as atrial fibrillation (2.0% versus 15.3%), major haemorrhage (5.9% versus 9.2%) and hypertension (10.9% versus 16.3%).[46] Additionally, there was no difference in the rate of infection despite higher rates of neutropenia with zanubrutinib compared with ibrutinib.[2] There were also fewer AEs leading to death (1.0 versus 4.1%), a lower rate of discontinuation due to AEs (4.0 versus 9.2%) and AEs leading to dose reduction (13.9 versus. 23.5%) with zanubrutinib compared with ibrutinib. In ASPEN, the primary reason for treatment discontinuation was disease progression in both treatment arms.[2]

In all zanubrutinib-treated patients in ASPEN, the incidences of AEs, Grade ≥3 AEs; SAEs; AEs leading to death or treatment discontinuation; and TRAEs were generally comparable between Cohort 1 and Cohort 2.[2, 47]

In a pooled analysis of 253 patients with WM, zanubrutinib demonstrated a tolerable safety profile. The AEs observed were consistent with those seen in ASPEN and the known toxicity profile for the BTK inhibitor class.[55]

## B.2.11 Ongoing studies

There are no additional ongoing studies due to provide additional evidence in the next 12 months for relapsed/refractory or treatment-naïve WM.

## B.2.12 Innovation

Treatment options for WM are limited across all lines of treatment and patients can cycle through and exhaust all available therapies.[16] No established treatment approach for WM has curative potential,[36] and once immuno-chemotherapy (e.g. rituximab combinations such as BR and DRC) and ibrutinib have been exhausted, there are no additional treatment options for relapsed/refractory patients. For treatment-naïve patients with WM unsuitable for chemo-immunotherapy, options are currently limited to best supportive care and ibrutinib.[42] There is a particular unmet need in those with *MYD88*$^{WT}$ WM, where ibrutinib has been found to demonstrate a shorter median survival and a lower probability of response than in those with *MYD88*$^{MUT}$.[16]

Ibrutinib is the only currently available treatment specifically developed for WM. However, real-world studies have shown that ibrutinib is discontinued by 22% of patients within one year of initiation due to unacceptable toxicity, disease progression and non-response.[41] Discontinuation combined with suboptimal adherence and treatment holds are associated with negative clinical outcomes for patients.[37]

Zanubrutinib, a potent and selective next-generation BTK inhibitor, is a new, potential treatment choice for those with WM unsuitable for chemo-immunotherapy or with relapsed/refractory WM, irrespective of *MYD88* status.[2, 47] Zanubrutinib binds to BTK, preventing the activation of the BCR signalling pathway; this inhibits the growth of malignant B-cells and leads to cell death. Zanubrutinib has high selectivity for BTK and so does not interact with other kinases. As zanubrutinib is highly specific and selective for BTK, and was designed to minimise off-target inhibition of other kinases, it has the potential to significantly improve outcomes and reduce side effects compared with existing therapies for WM.[2]

## B.2.13 Interpretation of clinical effectiveness and safety evidence

### B.2.13.1 Interim findings from the clinical evidence

The efficacy and safety of zanubrutinib in WM is supported by a comprehensive clinical trial programme, including the largest Phase 3 trial of BTK inhibitors in WM and the first head-to-head comparison of BTK inhibitors in any disease.[2] The benefit risk profile overall supports the use of zanubrutinib over ibrutinib in patients with WM based on comparable efficacy, with QoL improvements and superior safety.[2] Patients with WM treated with zanubrutinib demonstrated high rates of durable response that were rapid in onset with earlier achievement of VGPR and reduction in IgM compared with ibrutinib. Zanubrutinib has a favourable safety and tolerability profile compared with ibrutinib, with a numerically lower rate of several AEs, such as atrial fibrillation, bleeding and hypertension, and a lower rate of discontinuation due to AEs.[2] Zanubrutinib has also shown equivalent improvement in QoL from baseline compared with ibrutinib, with notable improvements in EQ-5D-5L score and EORTC QLQ-C30 subscales, including fatigue, physical functioning, and dyspnoea.[2]

In addition to efficacy in patients with *MYD88*$^{MUT}$ WM, zanubrutinib is an effective, well tolerated treatment in patients with *MYD88*$^{WT}$ WM.[47] As a selective BTK inhibitor, zanubrutinib offers improved safety and tolerability, and comparable efficacy over existing treatment options and therefore provides a new treatment choice for patients with WM, regardless of *MYD88* status, and regardless of line of therapy.[2, 43]

Due to the lack of head-to-head data comparing zanubrutinib to BR and DRC, an indirect treatment comparison was necessary. In a MAIC (see Section B.2.9), zanubrutinib treatment was associated with [redacted] compared with BR (HR [95% CI]: [redacted] and [redacted], respectively). Similarly, treatment with zanubrutinib was associated with significantly longer [redacted] (HR [95% CI]: [redacted]) and [redacted] (HR [95% CI]: [redacted]) compared with DRC.

### B.2.13.2 Strengths and limitations of the clinical evidence base

Overall, clinical data for zanubrutinib provide an appropriate evidence base for assessment of its clinical and cost-effectiveness for the treatment of WM.

The strengths of the clinical evidence base are:

- ASPEN was a robust, multicentre, head-to-head RCT, which randomised 201 patients with WM who were relapsed/refractory or treatment naïve and not suitable for chemotherapy[2]
- The trial included 33 patients in the UK, and enrolled patients representative of those who would receive treatment with zanubrutinib[46]
- ASPEN assessed the efficacy and safety of zanubrutinib in patients with *MYD88*$^{WT}$ WM and showed clinically meaningful anti-tumour activity[43]
- The study also included an assessment of HRQoL, as measured by the EQ-5D-5L and EORTC QLQ-C30 instruments[46]
- Safety data from four studies was pooled to assess the tolerability of zanubrutinib in 253 patients with WM[55]

The limitations of the clinical evidence base include:

- ASPEN was limited to open-label treatment masking due to differences in the number of capsules administered[2]
- There were differences in the zanubrutinib and ibrutinib populations at baseline with more patients being >75 years and haemoglobin levels ≤110 g/L[2]
- The primary endpoint of the study was not met, which meant that secondary endpoints could not be tested for significance[2]
- The study used the surrogate endpoint of VGPR/CR[2]
- As the study was not powered for OS and PFS, they were not reached at the time of data cut-off.[2] [redacted]

# B.3. Cost effectiveness

## B.3.1 Published cost-effectiveness studies

An SLR of the published literature and health technology assessment (HTA) submission documents was conducted to identify previously developed economic models that evaluated the cost-effectiveness of pharmacologic therapy in patients with WM (see Appendix G). Three cost-effectiveness analyses were identified; results of the published cost-effectiveness analyses are presented in Appendix G.

## B.3.2 Economic analysis

As described in Appendix G, all previous cost-effectiveness analyses in WM adopted a five-state Markov model, based on results from an ITC of ibrutinib versus physicians' choice of therapy (i.e., a mix of various chemo-immunotherapies), relying on the IPD of both a single-arm Phase 2 trial (Study 1118E, NCT01614821) and a European chart review that was available to ibrutinib's manufacturer. However, during the appraisal of ibrutinib in WM, (NICE TA491), the evidence review group questioned whether the available data justified the use of a five-state sequence-based model.

Despite potential limitations of the data applied in the previous models for ibrutinib and the necessities of making extra assumptions, it was feasible to adopt a five-state model for ibrutinib, as the manufacturer of ibrutinib had access to the study protocol, clinical study report and IPD of the European chart review. However, because such information was not publicly available for this analysis, and given the limited clinical data from both the zanubrutinib trials and published literature (as described in Section B.2), a standard three-state model (pre-progression survival, post-progression survival, death) was developed from the perspective of the NHS and personal social services to evaluate the cost-effectiveness of zanubrutinib in the treatment of WM compared with ibrutinib, BR, and DRC.

### B.3.2.1 Patient population

The target population was adult patients with WM previously treated with at least one prior line of therapy, or who are treatment naïve and unsuitable for chemo-immunotherapy.

### B.3.2.2 Model structure

A standard three-state partitioned survival model (PSM; pre-progression survival, post-progression survival, death) was developed to project the long-term clinical and economic consequences, based on data availability (see Appendix D for results of the SLR of clinical evidence) and in line with common modelling approaches and assumptions in oncology.

A cycle length of 28 days was adopted, which provided the appropriate level of detail and was consistent with the treatment dose schedules. A lifetime horizon, assumed to be 30 years, was adopted. This is in line with the mean baseline age of patients in the ASPEN ITT population (69.5 years)[46] and the lifetime horizon assumption (30 years) adopted in the model supporting NICE TA491 (ibrutinib in WM).[3] Half-cycle correction was applied.

A discounting rate of 3.5% was applied for costs and clinical outcomes that occurred beyond the first model year as per the NICE guide to methods of technology appraisal (2013).[56]

A comparison of economic features with previous NICE TAs is provided in Table B.3.1.

**Table B.3.1. Features of the economic analysis**

| | NICE TA491[3, 57] | Current appraisal | |
|---|---|---|---|
| | | **Chosen values** | **Justification** |
| Time horizon | 30 years (lifetime) | 30 years (lifetime) | NICE reference case |
| Model structure and health state | Five-state Markov model | Three-state PSM | A five-state Markov model was feasible in TA491 for ibrutinib because the manufacturer of ibrutinib had access to the study protocols, clinical study reports, and the IPD of an unpublished European chart review study. However, such information was not publicly available for this analysis. In addition, the ERG questioned whether the available data above justified the use of a five-state sequence-based model. Given the above and the limited clinical data from both the zanubrutinib trials and published literature (as described in Section B.2), a three-state model was developed for this analysis |
| Source of utilities | RESONATE trial for ibrutinib in relapsed/refractory CLL | ASPEN | In the previous appraisal, as no utility data were collected in Study 1118E and no WM-specific data were identified in the literature, utility inputs in the model were informed by the RESONATE study of ibrutinib in relapsed/refractory CLL. In contrast, WM-specific data was available from ASPEN and therefore used to inform the current appraisal (see Section B.3.4). |
| Sources of costs | NHS reference costs; PSSRU; BNF | NHS reference costs; PSSRU; BNF | NICE reference case |

Abbreviations: BNF = British national formulary; CLL = chronic lymphocytic leukaemia; ERG = evidence review group; HRQoL = health-related quality of life; IPD = individual patient-level data; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; PSM = partitioned survival model; PSSRU = Personal Social Services Research Unit; WM = Waldenström's macroglobulinaemia

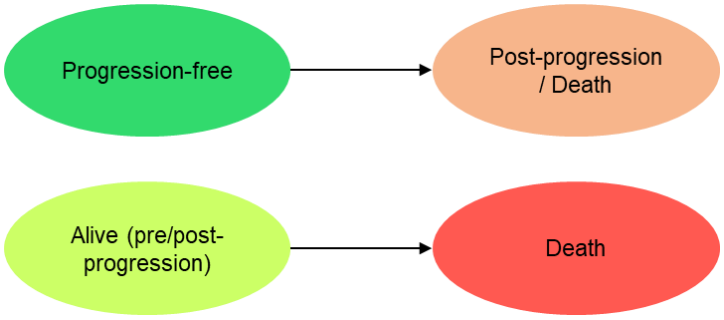
In the three-state PSM, it is assumed that any patient can be in any of the following mutually exclusive health states throughout a lifetime: pre-progression survival, post-progression survival, and death, depending on occurrences of progression and mortality events. At any time during the model, the proportion of patients in each health state always sums 100%.

The pre-progression health state includes patients who have not yet progressed on the primary treatment. The post-progression state reflects the proportion of patients who have experienced disease progression but remain alive.

Figure B.3.1 presents the modelled clinical pathway of a cohort of patients in the standard PSM. At a cohort level, at baseline, all the patients are alive without experiencing disease progression (i.e., pre-progression survival health state). All patients in the pre-progression health state are at risk of disease progression and mortality. As time goes by, an increasing proportion of patients will experience disease progression (i.e., enter post-progression survival health state) or mortality events (i.e., enter the death health state). All the patients alive (either in pre-progression survival or post-progression survival states) are at risk of mortality. That is, the proportion of patients in the pre-progression survival health state can only stay the same

or decrease over time, while the cumulative proportion of patients experiencing disease progression or a mortality event can only stay the same or increase over time.

**Figure B.3.1. Modelled patient transitions in three-state PSM**

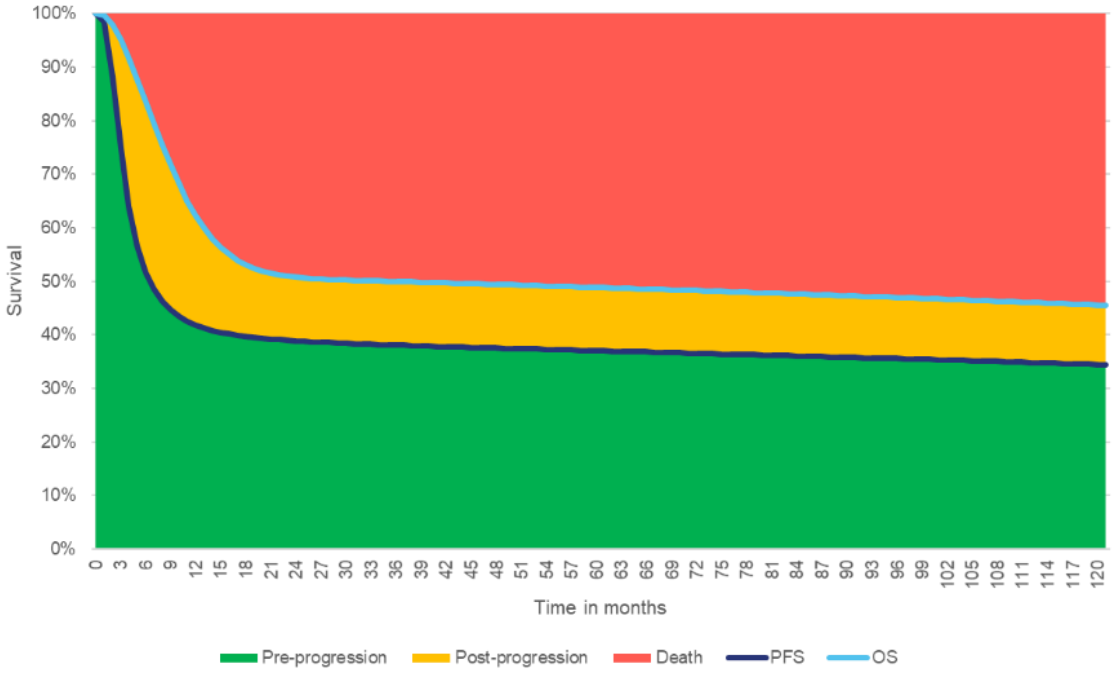



Abbreviations: PSM = partitioned survival model

As patients staying in specific health states accrue associated costs, life years (LYs) and quality-adjusted life years (QALYs), the proportion of patients in different health states in each model cycle are applied to estimate total costs, LYs, and QALYs of the entire cohort over time.

To determine the proportion of patients in each model health state over time, a standard PSM approach was adopted, in which PFS and OS curves were used together to distribute the population by health state using an area-under-the-curve (AUC) approach, as presented in Figure B.3.2.

**Figure B.3.2. Survival curves and health state distributions in the three-state PSM**




Abbreviations: OS = overall survival; PFS = progression-free survival; PSM = partitioned survival model

Using an AUC approach, the proportion of patients in the pre-progression health state was determined by the area (shown in green in Figure B.3.2) below the PFS curve (shown in dark blue). The proportion of patients in the death health state was determined by the area (shown in red in Figure B.3.2) above the OS curve (shown in light blue). The proportion of patients in

the post-progression state (shown in yellow in Figure B.3.2) was determined by the area between the PFS (dark blue) and OS (light blue) curves.

The OS and PFS curves were adjusted by general mortality such that at any time during the model, the mortality rates for the modelled population should not be lower than the mortality rates for the general population per country-specific life tables. Therefore, the hazard directly applied at any time (t) during the model was the maximum of the hazard of the parametric model (t) and hazard of background mortality (t).

### B.3.2.3 Intervention technology and comparators

The intervention of interest was zanubrutinib. Comparators were ibrutinib, BR and DRC. Other than BR and DRC, it was not possible to conduct comparisons with chemotherapy regimens or BSC, due to a lack of data in the literature to enable comparison of zanubrutinib with the comparators of interest (see Appendix D). However, BR and DRC are currently the two most common regimens for the first-line treatment of WM in patients considered fit enough to tolerate them (13.1% and 16.2%, respectively [see Section B.1.3.5.2]), and the third- and second- most common second-line regimens, respectively, behind ibrutinib (18.2%).[1]

The dosage information for zanubrutinib was obtained from ASPEN. The drug doses for the other treatment regimens were based on the SmPC where applicable or trial publications (see Table B.3.2).

**Table B.3.2. Intervention and comparators with dosage information**

| Regimen | Dosage | Stopping rule | Relative dose intensity |
|---|---|---|---|
| Zanubrutinib | 160 mg orally BID[46] | Until disease progression, or no longer tolerated by the patient[46] | 97.64%[46] |
| Ibrutinib | 420 mg orally OD[58] | Until disease progression, or no longer tolerated by the patient[58] | 98.18%[46] |
| BR | Rituximab (375 mg/m², day 1) plus bendamustine (90 mg/m², days 1 and 2) IV infused every cycle. Repeated every 4 weeks[49] | Until 6 cycles[49] or disease progression (assumption) | 100% (assumption) |
| DRC | Dexamethasone 20 mg IV on day 1, rituximab 375 mg/m² IV on day 1, and cyclophosphamide 100 mg/m² orally twice daily on days 1 through 5. Repeated every 3 weeks[50] | Until 6 cycles[50] or disease progression (assumption) | 100% (assumption) |

Abbreviations: BID = twice daily BR = bendamustine and rituximab; DRC = dexamethasone, rituximab and cyclophosphamide; OD = once daily; IV = intravenous

## B.3.3 Clinical parameters and variables

### B.3.3.1 Baseline characteristics

In the base-case analysis, baseline patient characteristics were based on the unadjusted data of the ASPEN ITT population (Table B.3.3), consistently for all three pairwise comparisons. Scenario analyses (see Section B.3.8.3) were conducted using the baseline patient

characteristics after matching adjustment (Table B.3.4). For details of the MAIC, see Section B.2.9.

**Table B.3.3. Baseline patient characteristics, base-case analysis**

| Parameter | Value (N=201) | Source |
|---|---|---|
| Female proportion, % | 33.33 | ASPEN IPD |
| Mean age, year | 69.53 | |
| Body surface area, $m^2$ | 1.86 | |

Abbreviations: IPD = individual patient-level data; N = number of patients evaluable

**Table B.3.4. Baseline patient characteristics, scenario analyses**

| Parameter | Value | Source |
|---|---|---|
| Zanubrutinib (match BR; $n_{eff}$=[redacted]) | | |
| Female proportion, % | 39.05 | ASPEN IPD |
| Mean age, year | 70.84 | |
| Body surface area, $m^2$ | 1.84 | |
| Zanubrutinib (match DRC; $n_{eff}$=[redacted]) | | |
| Female proportion, % | 39.52 | ASPEN IPD |
| Mean age, year | 69.39 | |
| Body surface area, $m^2$ | 1.87 | |

Abbreviations: BR = bendamustine and rituximab; DRC = dexamethasone, rituximab and cyclophosphamide; IPD = individual patient-level data; $n^{eff}$ = effective sample size

### B.3.3.2 PFS, OS, and TTD

As specified in Section B.3.2.2, the PSM included three mutually exclusive health states: pre-progression survival, post-progression survival, and death. To determine the time spent in each health state and the accrued costs and QALYs, the proportion of patients in each health state over time was derived from the PFS and OS curves using the AUC approach. In addition, to estimate the drug costs for the BTK inhibitors (i.e., zanubrutinib and ibrutinib), a TTD curve was applied.

To extrapolate the PFS, OS and TTD beyond the trial period, the following steps were conducted, in line with the recommendations of NICE DSU technical support document 14:[59]

- First, the PH assumption was assessed through log-cumulative hazard plots in order to determine whether it was appropriate to apply a PH modelling approach with treatment group included as a covariate, or to fit independent parametric models to each treatment group separately
- Second, six parametric models were fitted (exponential, Weibull, Gompertz, log-normal, log-logistic, and gamma)
- Third, the most plausible model was selected based on assessment of:Internal validity of OS/PFS/TTD, based on Akaike information criteria (AIC) and Bayesian information criteria (BIC) fit statistics and visual inspection. Given the uncertainty of the survival data of BTK inhibitors due to its immaturity, the structural stability of the parametric models was also assessed through visual inspection of the 95% CI of the models
    - External validity of OS, based on published estimates and clinical expert opinion on the clinical plausibility of the extrapolated survival and hazard pattern. For a more detailed description on the expert elicitation, see Section B.3.10.1

- External validity of PFS/TTD, based on the alignment between PFS and TTD in parametric distribution given that disease progression usually results in a treatment discontinuation.

Analyses were performed using the *flexsurv* package in R. Results of the model selection are summarised in Table B.3.5.

**Table B.3.5. Summary of model selection in the base-case analysis**

| Outcome | Treatment | Base-case analysis | | Scenario analysis | |
|---|---|---|---|---|---|
| | | Setting | Justification | Setting | Justification |
| **Pairwise comparison of zanubrutinib vs ibrutinib** | | | | | |
| OS | Zanubrutinib<br>Ibrutinib | Dependent exponential model | • Relatively parallel log-cumulative hazard plots between treatments<br>• Clinically plausible mean OS for both treatments<br>• Clinically plausible hazard patterns for both treatments<br>• The lowest BIC | Applying a hazard ratio of one to OS/PFS/TTD of ibrutinib to derive OS/PFS/TTD of zanubrutinib respectively beyond the trial period | Although zanubrutinib was associated with slightly better survival outcomes within the trial period, given the similar survival outcomes, a conservative approach of applying a hazard ratio of one was explored.<br>No alternative parametric distribution was assessed, given that none of the non-exponential distributions were associated with clinically plausible hazard patterns |
| PFS | Zanubrutinib<br>Ibrutinib | Dependent exponential model | • Relatively parallel log-cumulative hazard plots between treatments<br>• The lowest BIC<br>• Alignment with TTD in parametric distribution | | |
| TTD | Zanubrutinib<br>Ibrutinib | Dependent exponential model | • Relatively parallel log-cumulative hazard plots between treatments<br>• The lowest BIC<br>• Alignment with PFS in parametric distribution | | |
| **Pairwise comparison of zanubrutinib (match DRC) vs DRC** | | | | | |
| OS | Zanubrutinib (matching DRC)<br>DRC | Dependent gamma model | • Clinically plausible mean OS for both treatments<br>• Clinically plausible hazard patterns for both treatments<br>• The second lowest BIC | Dependent Weibull model; dependent Gompertz model | • Clinically plausible mean OS for both treatments<br>• Clinically plausible hazard patterns for both treatments<br>• The third and fourth lowest BIC |
| PFS | Zanubrutinib (matching DRC)<br>DRC | Dependent exponential model | • The lowest BIC<br>• Alignment with TTD in parametric distribution | None | For both PFS and TTD, the exponential distribution was consistently associated with obviously lower BIC compared with other distributions |
| TTD | Zanubrutinib (matching DRC) | Independent exponential model | • The lowest BIC<br>• Alignment with PFS in parametric distribution | | |
| | DRC | N/A | N/A | N/A | N/A |
| **Pairwise comparison of zanubrutinib (match BR) vs BR** | | | | | |
| OS | Zanubrutinib (matching BR) | Independent exponential model | | Dependent Weibull model; | • Clinically plausible mean OS and |

| Outcome | Treatment | Base-case analysis | | Scenario analysis | |
|---|---|---|---|---|---|
| | | Setting | Justification | Setting | Justification |
| | BR | Independent Weibull model | • Clinically plausible mean OS for both treatments<br>• Clinically plausible hazard patterns for BR | dependent gamma model | hazard patterns for zanubrutinib (match BR)<br>• Clinically plausible hazard pattern for both treatment arms |
| PFS | Zanubrutinib (match BR) | Dependent exponential model | • Relatively parallel log-cumulative hazard plots dependent models<br>• The lowest BIC<br>• Alignment with TTD in parametric distribution (specific for zanubrutinib) | None | For both PFS and TTD, the exponential distribution was consistently associated with obviously lower BIC compared with other distributions. |
| | BR | | | | |
| TTD | Zanubrutinib (match BR) | Independent exponential model | • The lowest BIC<br>• Alignment with PFS in parametric distribution (specific for zanubrutinib) | None | |
| | BR | N/A | N/A | N/A | N/A |

Abbreviations: BIC = Bayesian information criteria; BR = bendamustine and rituximab; DRC = dexamethasone, rituximab and cyclophosphamide; N/A = not applicable; OS = overall survival; PFS = progression-free survival; TTD = time to treatment discontinuation; vs = versus

#### *B.3.3.2.1 Zanubrutinib versus ibrutinib*

Figure B.3.3 to Figure B.3.5 present the KM curves for PFS, OS, and TTD respectively for zanubrutinib and ibrutinib, based on the head-to-head comparison for the ASPEN ITT population.

**Figure B.3.3. KM curves of PFS – zanubrutinib vs ibrutinib**

Abbreviations: KM = Kaplan-Meier; PFS = progression-free survival; vs = versus

**Figure B.3.4. KM curves of OS – zanubrutinib vs ibrutinib**

Abbreviations: KM = Kaplan-Meier; OS = overall survival; vs = versus

**Figure B.3.5. KM curves of TTD – zanubrutinib vs ibrutinib**

Abbreviations: KM = Kaplan-Meier; TTD = time to treatment discontinuation; vs = versus

#### *B.3.3.2.1.1 Assessment of PH assumption*

The PH assumption was assessed through visual inspection of the log-cumulative hazard plots for PFS, OS, and TTD (Figure B.3.6 to Figure B.3.8). These plots show relatively straight and parallel curves. Despite the two curves crossing at the end for OS and TTD (Figure B.3.7 and Figure B.3.8, respectively), the crossing occurred at the end of follow-up with a limited number of patients at risk with considerable uncertainty, which may be less informative. Hence, jointly fit models were applied to model both treatment arms in one parametric model with zanubrutinib included as a covariate. Six parametric distributions were assessed.

**Figure B.3.6. Log-cumulative hazards vs log time for PFS – zanubrutinib vs ibrutinib**

Abbreviation: PFS = progression-free survival; vs = versus

**Figure B.3.7. Log-cumulative hazards vs log time for OS – zanubrutinib vs ibrutinib**

Abbreviation: OS = overall survival; vs = versus

**Figure B.3.8. Log-cumulative hazards vs log time for TTD – zanubrutinib vs ibrutinib**

Abbreviation: TTD = time to treatment discontinuation; vs = versus

*B.3.3.2.1.2 Assessment of internal validity of OS/PFS/TTD*

Goodness-of-fit was assessed with AIC and BIC statistics (Table B.3.6) and visual comparison of the KM curves against the parametric curves was performed (Figure B.3.9 to Figure B.3.11). The fit statistics for PFS and OS indicated that the exponential model provided the best fit to both PFS and OS. For TTD, the log-normal model was associated with the lowest AIC whereas the exponential model was associated with the lowest BIC. To avoid over-fitting, goodness-of-fit was assessed based on BIC statistics, and therefore the exponential model was considered to provide better fits to the KM curves.

The results of fit statistics were validated through visual inspections.

**Table B.3.6. Fit statistics for jointly fitted PFS, OS, and TTD – zanubrutinib vs ibrutinib**

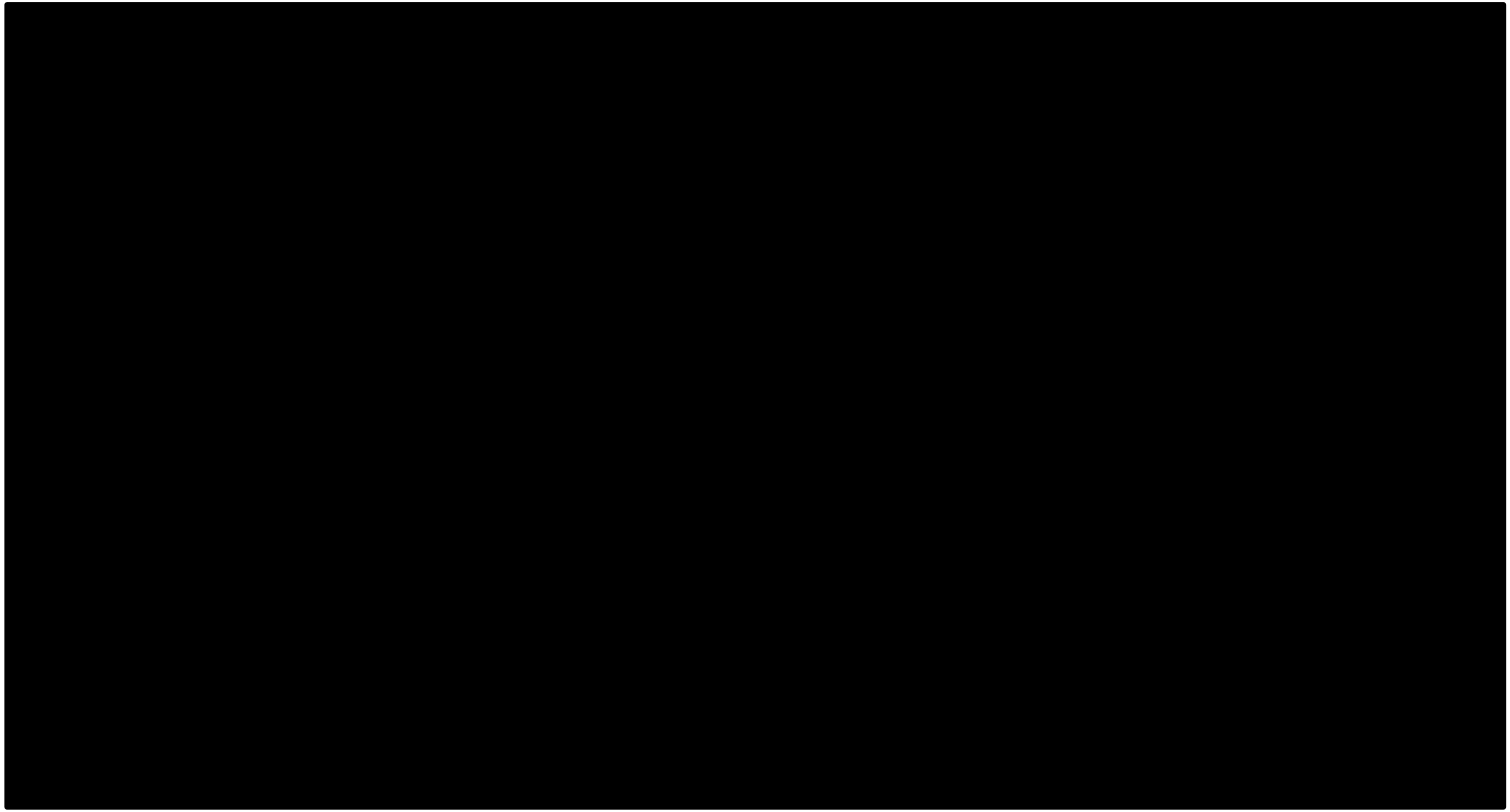
| Parametric distribution | PFS | | OS | | TTD | |
|---|---|---|---|---|---|---|
| | AIC | BIC | AIC | BIC | AIC | BIC |
| Exponential | | | | | | |
| Weibull | | | | | | |
| Gompertz | | | | | | |
| Log-normal | | | | | | |
| Log-logistic | | | | | | |
| Gamma | | | | | | |

Abbreviations: AIC = Akaike information criteria; BIC = Bayesian information criteria; OS = overall survival; PFS = progression-free survival; TTD = time to discontinuation; vs = versus
Note: The lowest AIC or BIC is in italics

**Figure B.3.9. Visual inspection of jointly fitted parametric vs KM curves for PFS – zanubrutinib vs ibrutinib**

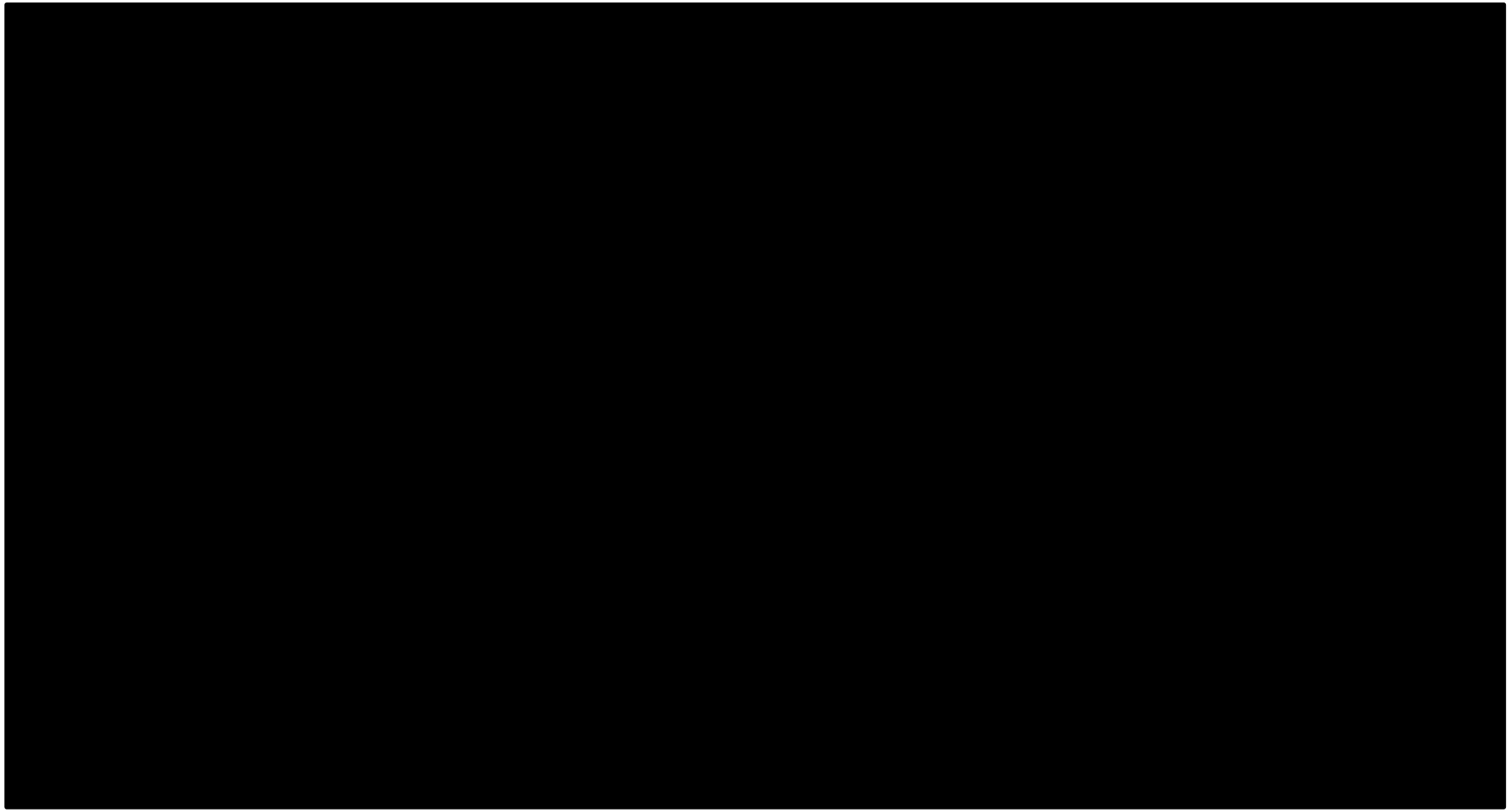
Abbreviations: KM = Kaplan-Meier; PFS = progression-free survival; vs = versus
Notes: (1) The graphs above are for assessment of internal validity only without adjusting for background mortality. (2) The 6 graphs on the left are identical to the 6 graphs on the right except for the 95% confidence intervals of the parametric curves in order for clearer visual inspection

**Figure B.3.10. Visual inspection of jointly fitted parametric vs KM curves for OS – zanubrutinib vs ibrutinib**



Abbreviations: KM = Kaplan-Meier; OS = overall survival; vs = versus
Notes: (1) The graphs above are for assessment of internal validity only without adjusting for background mortality. (2) The 6 graphs on the left are identical to the 6 graphs on the right except for the 95% confidence interval of the parametric curve in order for clearer visual inspection

**Figure B.3.11. Visual inspection of jointly fitted parametric vs KM curves for TTD – zanubrutinib vs ibrutinib**



Abbreviations: KM = Kaplan-Meier; TTD = time to treatment discontinuation; vs = versus
Notes: (1) The graphs above are for assessment of internal validity only without adjusting for background mortality. (2) The 6 graphs on the left are identical to the 6 graphs on the right except for the 95% confidence interval of the parametric curve in order for clearer visual inspection

*B.3.3.2.1.3 Assessment of external validity of OS*

Although goodness-of-fit assessment supported the use of the exponential model, given that the fit statistics were very close across distributions, and more importantly, given the immaturity of survival data for BTK inhibitors from ASPEN, the external validity was assessed through:

- comparison of modelled survival versus the observed survival in BGB-3111-AU-003, the Phase 1/2 trial for zanubrutinib with slightly longer follow-up
- review of external literature and technology appraisals (including clinical trials for other BTK inhibitors in the WM population,[53, 60] previous technology appraisals in WM[57], other published literature) and
- clinical experts' opinions on the clinical plausibility of modelled survival and hazard patterns.

As presented in Table B.3.7 and Table B.3.8, all the jointly fitted models generated similar, albeit slightly higher, OS for zanubrutinib; mean OS was between [redacted] years for zanubrutinib and between [redacted] years for ibrutinib. The exponential model generated the most conservative mean OS for both zanubrutinib and ibrutinib.

The landmark OS for zanubrutinib from all these models (5-year OS rates of [redacted]) were generally aligned with, but slightly higher than, the landmark OS observed in BGB-3111-AU-003 (48-month OS rate of [redacted] in a total of 73 patients, [redacted] in 49 relapsed/refractory patients, [redacted] in 24 treatment-naïve patients unsuitable for chemo-immunotherapy, after a median follow-up of 48 months).

In addition to the clinical trials for zanubrutinib above, survival results of the previously published clinical trial for ibrutinib (Phase 2 Study 1118E for ibrutinib monotherapy) were reviewed. However, given the immaturity of publicly available survival data (e.g., OS rate of 90% after a median follow-up of 37 months[57]), these results are not informative for the validation of long-term survival extrapolation.

Given the immaturity of survival data in the clinical trials for BTK inhibitors in WM in general, the long-term OS estimates based on less recent studies (in which BTK inhibitors were not an available treatment option then) were reviewed and suggested that patients not treated with BTK inhibitors had a median OS of approximately 10 years.[3, 12, 57] Although these studies might not be completely informative for validation of the exact OS with BTK inhibitors, given the data limitations in WM, it might still be informative to rely on all available data to inform the plausible range of OS in patients treated with BTK inhibitors. For example, it was reported in NICE TA491 (ibrutinib in WM) that the median OS in WM ranged from less than 4 to 12 years and that median OS in the European chart review study was 123 months (i.e., approximately 10 years) for patients receiving a mix of physicians' choice of therapy (second-line: 47% for BR, 31% for DRC, 11% for FCR, 0% for Clad-R, 11% for other; third- or fourth-line: 43% for BR, 15% for DRC, 9% for FCR, 30% for Clad-R, 3% for other). However, considerable country-specific OS differences were noted (e.g., UK-specific median OS was reported to be 5 years; exact estimates for other EU countries were not publicly reported).[3, 57] It was also reported in the ESMO Clinical Practice Guideline in 2018 that the median OS exceeded 10 years.[12]

In addition to the published estimates above, clinical experts were consulted as to the clinical plausibility of the modelled OS estimates and the hazard patterns (see Section B.3.10.1 for more details). Clinical experts stated that all parametric distributions generated clinically

plausible mean OS of approximately [redacted] years (Table B.3.7 and Table B.3.8). Experts also stated that patients treated with BTK inhibitors (either relapsed/refractory or treatment-naïve patients unsuitable for chemo-immunotherapy) would have monotonically increasing hazards of death, given that once these patients progressed on BTK inhibitors, they would likely quickly run out of active treatment options. Other than the exponential model with constant hazard, all other models were associated with decreasing hazards over time, before adjusting for background mortality. The comparison of hazard patterns across distributions suggested that the exponential model was associated with the most clinically plausible hazard pattern, assuming a relatively homogenous WM population. The results of hazard pattern were aligned with the results of landmark and mean OS estimates that the exponential model was associated with the most conservative OS. Given the above, the exponential model was considered to be the most clinically plausible.

**Table B.3.7. Landmark, median, mean and hazard patterns of OS – zanubrutinib[a]**

| | Exponential | Weibull | Gompertz | Log-normal | Log-logistic | Gamma |
|---|---|---|---|---|---|---|
| Landmark | | | | | | |
| 2 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 5 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 10 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 15 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Median (year) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Mean (year) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Hazard pattern | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically decreasing | Monotonically decreasing | Monotonically decreasing | Monotonically decreasing | Monotonically decreasing |
| After adjusting for background mortality | Constant in the first 8 years; then increasing | Monotonically decreasing in the first 5 years; then increasing | Monotonically decreasing in the first 5 years; then increasing | Monotonically decreasing in the first 4 years; then increasing | Monotonically decreasing in the first 5 years; then increasing | Monotonically decreasing in the first 6 years; then increasing |

Abbreviation: OS = overall survival

[a] Unless otherwise specified, the survival estimates were after adjustment by background mortality such that in any time during the model, the morality rate of modelled population would not be lower than that of general population.

**Table B.3.8. Landmark, median, mean and hazard patterns of OS – ibrutinib[a]**

| | Exponential | Weibull | Gompertz | Log-normal | Log-logistic | Gamma |
|---|---|---|---|---|---|---|
| Landmark | | | | | | |
| 2 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 5 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 10 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 15 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Median (year) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Mean (year) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Hazard pattern | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically decreasing | Monotonically decreasing | Monotonically decreasing | Monotonically decreasing | Monotonically decreasing |
| After adjusting for background mortality | Constant in the first 11 years; then increasing | Monotonically decreasing in the first 8 years; then increasing | Monotonically decreasing in the first 5 years; then increasing | Monotonically decreasing in the first 5 years; then increasing | Monotonically decreasing in the first 6 years; then increasing | Monotonically decreasing in the first 8 years; then increasing |

Abbreviation: OS = overall survival

[a] Unless otherwise specified, the survival estimates were after adjustment by background mortality such that in any time during the model, the morality rate of modelled population would not be lower than that of general population.

*B.3.3.2.1.4 Assessment of external validity of PFS/TTD*

The results of BIC statistics showed that the exponential distribution provided the best fit to both PFS and TTD, which was aligned with the clinical association between disease progression and treatment discontinuation.

*B.3.3.2.1.5 Summary of model selection*

Given all the information above, the dependent exponential model was applied for OS in the base-case analysis because (1) it was associated with clinically plausible mean OS as confirmed by clinical experts; (2) it was associated with more clinically plausible hazard pattern (which led to the most conservative mean OS); and (3) it was associated with the lowest BIC.

For PFS and TTD, the dependent exponential models were applied for both PFS and TTD in the base-case analysis because the exponential distribution was associated with the lowest BIC for both PFS and TTD, which was aligned with the clinical association between disease progression and treatment discontinuation.

Additional scenarios of survival extrapolation were also explored (e.g., applying a hazard ratio of one to survival curves of ibrutinib to derive survival curves of zanubrutinib beyond the trial period; see Section B.3.8.3).

### *B.3.3.2.2 Zanubrutinib versus dexamethasone-rituximab-cyclophosphamide*

Figure B.3.12 to Figure B.3.14 present the KM curves of PFS, OS, and TTD respectively for zanubrutinib versus DRC, based on the MAIC results (see Section B.2.9).

**Figure B.3.12. KM curves of PFS – zanubrutinib vs DRC**

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; PFS = progression-free survival; vs = versus

**Figure B.3.13. KM curves of OS – zanubrutinib vs DRC**

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; OS = overall survival; vs = versus

**Figure B.3.14. KM curves of TTD – zanubrutinib**

Abbreviation: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; TTD, time to treatment discontinuation

*B.3.3.2.2.1 Assessment of PH assumption*

The PH assumption was assessed through visual inspection of the log-cumulative hazard plots for PFS and OS (Figure B.3.15 and Figure B.3.16). The plots showed relatively straight and parallel curves overall, which supported the use of a single model for both PFS and OS with treatment group included as a covariate. However, at certain time points the plots appeared to cross or diverge.

Given the above, for both OS and PFS, both dependent models (with treatment included as a covariate) and independent models were assessed. For TTD, only independent models were fitted to the zanubrutinib arm. For each outcome and model type, six parametric distributions were assessed.

**Figure B.3.15. Log-cumulative hazards vs log time for PFS – zanubrutinib vs DRC**

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; PFS = progression-free survival; vs = versus

**Figure B.3.16. Log-cumulative hazards vs log time for OS – zanubrutinib vs DRC**

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; OS = overall survival; vs = versus

*B.3.3.2.2.2 Assessment of internal validity of OS/PFS/TTD*

Goodness-of-fit was assessed with AIC and BIC statistics (Table B.3.9) and visual comparison of the KM curves against the parametric curves was performed (Figure B.3.17 to Figure B.3.23). The fit statistics indicated that the exponential distribution was associated with the lowest BIC across all the parametric distributions for both dependent and independent models. The results of fit statistics were validated through visual inspections.

**Table B.3.9. Fit statistics for jointly fitted PFS, OS and TTD – zanubrutinib (match DRC) vs DRC**

| Parametric distribution | PFS | | OS | | TTD | |
|---|---|---|---|---|---|---|
| | AIC | BIC | AIC | BIC | AIC | BIC |
| Exponential | [redacted] | [redacted] | [redacted] | [redacted] | N/A | |
| Weibull | | | | | | |
| Gompertz | | | | | | |
| Log-normal | | | | | | |
| Log-logistic | | | | | | |
| Gamma | | | | | | |

Abbreviations: AIC = Akaike information criteria; BIC = Bayesian information criteria; DRC = dexamethasone, rituximab and cyclophosphamide; N/A = not applicable; OS = overall survival; PFS = progression-free survival; TTD = time to discontinuation
Note: The lowest AIC or BIC is in italics

**Table B.3.10. Fit statistics for independently fitted PFS, OS and TTD – zanubrutinib (match DRC)**

| Parametric distribution | PFS | | OS | | TTD | |
|---|---|---|---|---|---|---|
| | AIC | BIC | AIC | BIC | AIC | BIC |
| Exponential | [redacted] | | | | | |
| Weibull | | | | | | |
| Gompertz | | | | | | |
| Log-normal | | | | | | |
| Log-logistic | | | | | | |
| Gamma | | | | | | |

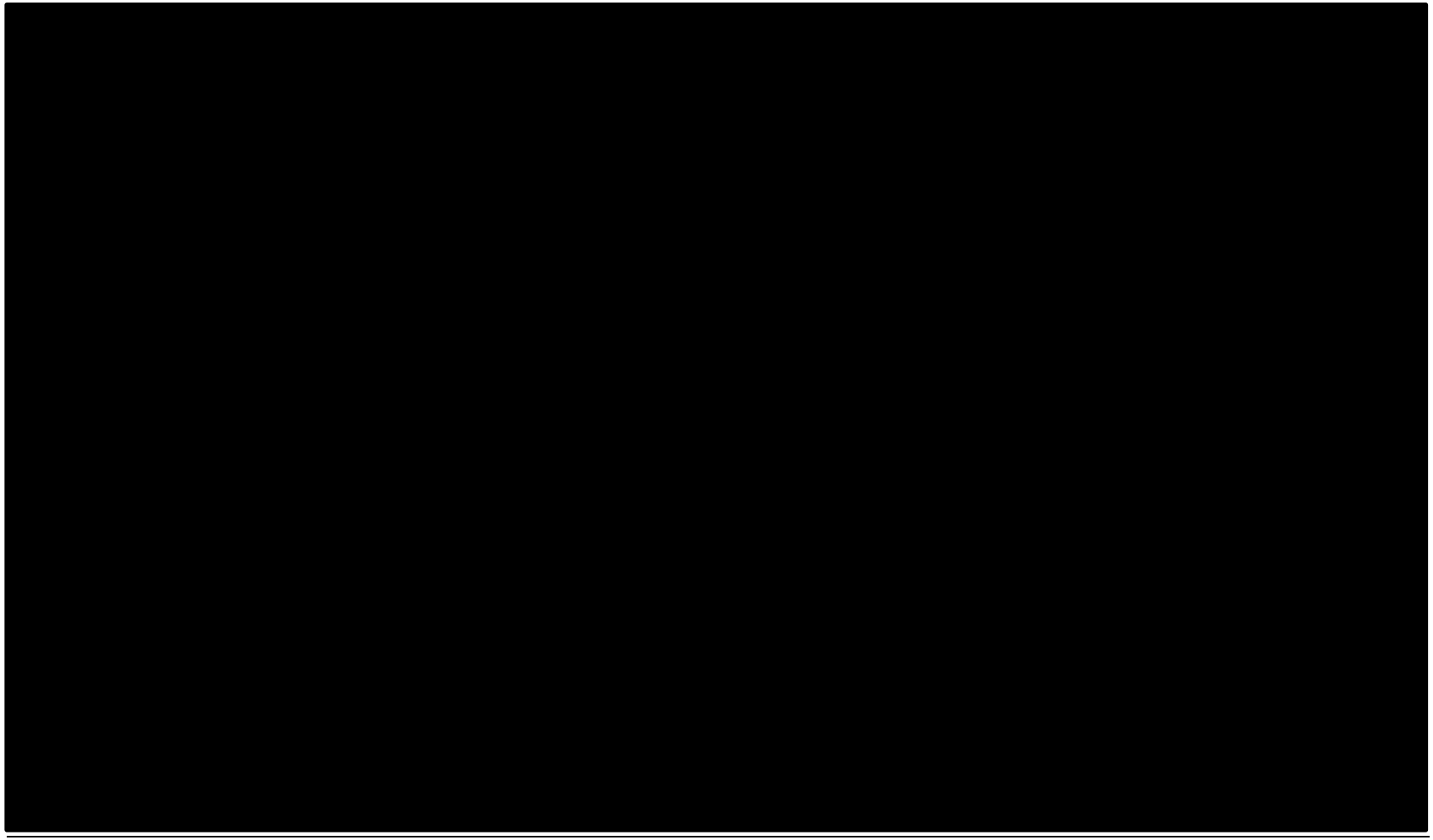
Abbreviations: AIC = Akaike information criteria; BIC = Bayesian information criteria; DRC = dexamethasone, rituximab and cyclophosphamide; OS = overall survival; PFS = progression-free survival; TTD = time to discontinuation
Note: The lowest AIC or BIC is in italics

**Table B.3.11. Fit statistics for independently fitted PFS, OS and TTD – DRC**

| Parametric distribution | PFS | | OS | | TTD | |
|---|---|---|---|---|---|---|
| | AIC | BIC | AIC | BIC | AIC | BIC |
| Exponential | [redacted] | | | | N/A | |
| Weibull | | | | | | |
| Gompertz | | | | | | |
| Log-normal | | | | | | |
| Log-logistic | | | | | | |
| Gamma | | | | | | |

Abbreviations: AIC = Akaike information criteria; BIC = Bayesian information criteria; DRC = dexamethasone, rituximab and cyclophosphamide; N/A = not applicable; OS = overall survival; PFS = progression-free survival; TTD = time to discontinuation
Note: The lowest AIC or BIC is in italics

**Figure B.3.17. Visual inspection of jointly fitted parametric vs KM curves for PFS – zanubrutinib (match DRC) vs DRC**



Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; PFS = progression-free survival; vs = versus
Notes: (1) The graphs above are for assessment of internal validity only without adjusting for background mortality. (2) The 6 graphs on the left are identical to the 6 graphs on the right except for the 95% confidence interval of the parametric curve in order for clearer visual inspection. (3) The curve of DRC on the third row and in the second column is mislabelled as log-normal which is supposed to be gamma

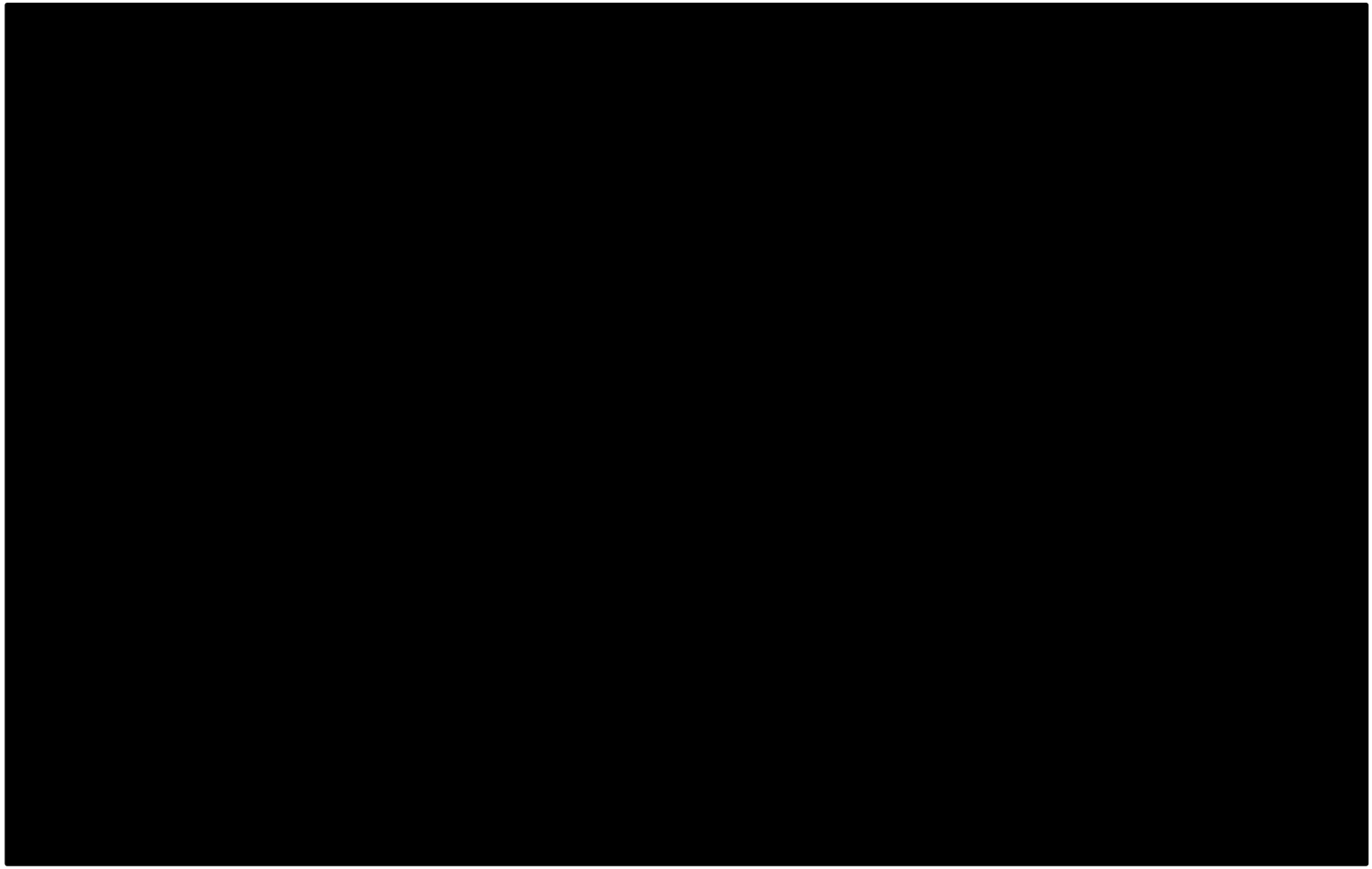
**Figure B.3.18. Visual inspection of independently fitted parametric vs KM curves for PFS – zanubrutinib (match DRC)**

Abbreviations: dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; PFS = progression-free survival
Note: The graphs above are for assessment of internal validity only without adjusting for background mortality

**Figure B.3.19. Visual inspection of independently fitted parametric vs KM curves for PFS – DRC**

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; PFS = progression-free survival
Note: The graphs above are for assessment of internal validity only without adjusting for background mortality

**Figure B.3.20. Visual inspection of jointly fitted parametric versus KM curves for OS – zanubrutinib (match DRC) vs DRC**



Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; OS = overall survival; vs = versus
Notes: (1) The graphs above are for assessment of internal validity only without adjusting for background mortality. (2) The 6 graphs on the left are identical to the 6 graphs on the right except for the 95% confidence interval of the parametric curve in order for clearer visual inspection. (3) The last graph on the third row and in the second column is supposed to be gamma for DRC

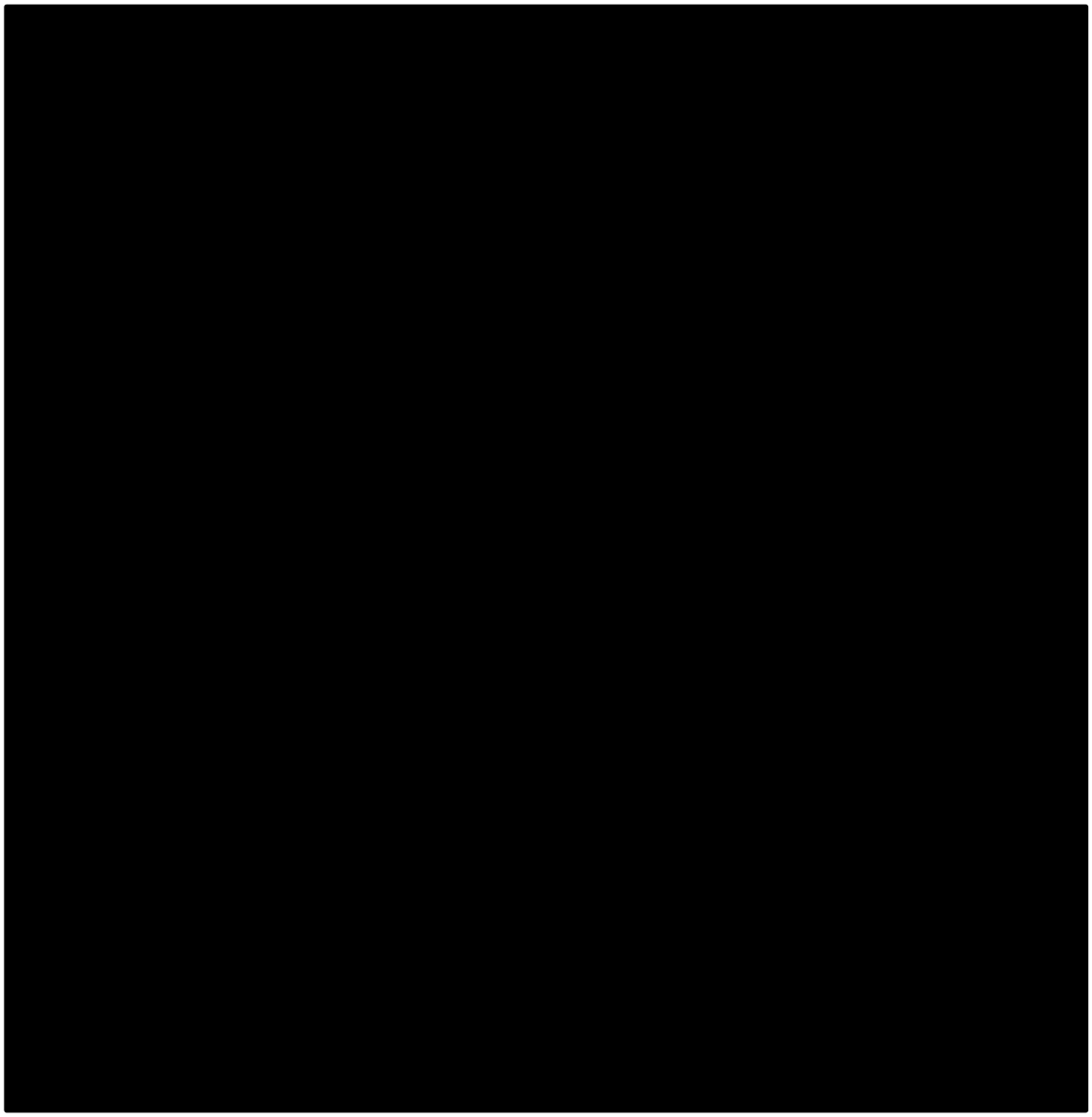
**Figure B.3.21. Visual inspection of independently fitted parametric versus KM curves for OS – zanubrutinib (match DRC)**

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; OS = overall survival
Note: The graphs above are for assessment of internal validity only without adjusting for background mortality

**Figure B.3.22. Visual inspection of independently fitted parametric versus KM curves for OS – DRC**

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; OS = overall survival
Note: The graphs above are for assessment of internal validity only without adjusting for background mortality

**Figure B.3.23. Visual inspection of independently fitted parametric versus KM curves for TTD – zanubrutinib (match DRC)**



Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; TTD = time to treatment discontinuation
Note: The graphs above are for assessment of internal validity only without adjusting for background mortality

*B.3.3.2.2.3 Assessment of external validity of OS*

Although goodness-of-fit assessment for zanubrutinib (matching DRC) supported the use of the exponential model, given that the fit statistics were close across parametric distributions, and more importantly, given the immaturity of survival data for zanubrutinib from ASPEN, external validity was assessed.

As presented in Table B.3.12, all the parametric models generated similar mean OS for the zanubrutinib arm (matching DRC; [redacted] years), which were very close to the mean OS for zanubrutinib without matching adjustment ([redacted] years in Table B.3.7). As per previous discussion based on clinical expert opinions (see Section B.3.3.2.1.3), all the parametric models of zanubrutinib (matching DRC) generated clinically plausible mean OS estimates of

approximately [redacted] years, whereas dependent Weibull (mean OS: [redacted] years), dependent Gompertz (mean OS: [redacted] years), and dependent gamma (mean OS: [redacted] years) models were associated with monotonically increasing hazards (before and after adjusting for background mortality) that were considered to be more clinically plausible, assuming a relatively homogenous WM population.

For DRC, given the relatively mature survival data, it was appropriate to rely on goodness-of-fit to inform model selection, which supported the use of the exponential model. In addition, external validity was assessed through:

- Comparison of modelled survival versus the observed survival in the Phase 1/2 BGB-3111-AU-003 study, with slightly longer follow-up
- Review of external literature and technology appraisals (including the previous technology appraisal in WM[57] and other published literature), and
- Clinical expert opinion on the clinical plausibility of modelled survival and hazard patterns.

As presented in Table B.3.13, all the parametric models generated similar OS estimates for DRC (mean: [redacted] years; median: [redacted] years). These estimates were generally aligned with previously published median estimates of approximately 10 years.[3, 12, 57] More specifically, it was reported in NICE TA491 (ibrutinib in WM) that median OS ranged from less than 4 to 12 years and that median OS in the European chart review study was 123 months (i.e., approximately 10 years) for patients receiving a mix of physicians' choice of therapy (second-line: 47% for BR, 31% for DRC, 11% for FCR, 0% for Clad-R, 11% for other; third- or fourth-line: 43% for BR, 15% for DRC, 9% for FCR, 30% for Clad-R, 3% for other). However, considerable country-specific OS differences were noted (e.g., UK-specific median OS was reported to be 5 years; exact estimates for other EU countries not publicly reported).[3, 57] It was also reported in the ESMO Clinical Practice Guideline in 2018 that median OS exceeds 10 years.[12]

Clinical experts were consulted as to clinically plausible OS estimates (see Section B.3.10.1 for more details). The experts stated that all the models generated clinically plausible OS estimates of approximately [redacted] years (Table B.3.13), based on the data of a study that was initiated about 15 years ago in treatment-naïve patients (i.e., Dimopoulos et al. 2007/Kastritis et al. 2015[50, 51]). More specifically, it was also suggested that:

- In more recent years, treatment-naïve patients treated with chemo-immunotherapy would likely live for [redacted] years on average, with approximately 3–4 additional years of life compared with 15 years ago.
- In more recent years, relapsed/refractory patients treated with second-line chemo-immunotherapy would likely have [redacted] years of life, whereas relapsed/refractory patients treated with third-line chemo-immunotherapy would likely have [redacted] years of life.

Given the clinical expert opinion above, most of the models generated a plausible mean OS within the range of [redacted] years (dependent Weibull with [redacted] years, dependent Gompertz with [redacted] years, dependent gamma with [redacted] years, independent Weibull with [redacted] years, independent Gompertz with [redacted] years, independent gamma with [redacted] years) in relapsed/refractory patients treated with second-line chemo-immunotherapy that is more comparable to the ASPEN patient population (with a mix of 85% relapsed/refractory patients and 15% treatment-naïve patients unsuitable for chemo-immunotherapy).

In addition, clinical experts were consulted as to the hazard pattern of the disease (see Section B.3.10.1). The experts stated that both relapsed/refractory and treatment-naïve patients treated with chemo-immunotherapy would have monotonically increasing hazards of death. Therefore, dependent/independent Weibull (mean OS: [redacted] years), dependent/independent Gompertz (mean OS: [redacted] years), and dependent/independent gamma (mean OS: [redacted] years) models with monotonically increasing hazards (before and after adjusting for background mortality) were considered to be more clinically plausible for DRC, assuming a relatively homogenous WM population.

**Table B.3.12. Landmark, median, mean and hazard patterns of OS – zanubrutinib (match DRC)[a]**

| | Exponential | Weibull | Gompertz | Log-normal | Log-logistic | Gamma |
|---|---|---|---|---|---|---|
| **Jointly fitted models** | | | | | | |
| Landmark | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 2 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 5 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 10 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 15 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Median (year) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Mean (year) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Hazard pattern | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically increasing | Monotonically increasing | Increasing in the first 1 year; then decreasing | Increasing in the first 5 years; then decreasing | Monotonically increasing |
| After adjusting for background mortality | Constant in the first 7 years; then increasing | Monotonically increasing | Monotonically increasing | Increasing in the first 2 years; then decreasing for 8 years; then increasing | Increasing in the first 3 years; then stable for 10 years; then increasing | Monotonically increasing |
| **Independently fitted models** | | | | | | |
| Landmark | | | | | | |
| 2 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 5 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 10 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 15 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Median (year) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Mean (year) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Hazard pattern | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically decreasing | Monotonically decreasing | Increasing in the first 3 months; then decreasing | Monotonically decreasing | Monotonically decreasing |
| After adjusting for background mortality | Constant in the first 8 years; then increasing | Stable in the first 7 years; then increasing | Decreasing in the first 2 years; then increasing | Increasing in the first 3 months; then decreasing for 4 years; then increasing | Stable in the first 6 years; then increasing | Stable in the first 8 years; then increasing |

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; OS = overall survival

[a] Unless otherwise specified, the survival estimates were after adjustment by background mortality such that in any time during the model, the morality rate of modelled population would not be lower than that of general population

**Table B.3.13. Landmark, median, mean and hazard patterns of OS – DRC[a]**

| | Exponential | Weibull | Gompertz | Log-normal | Log-logistic | Gamma |
|---|---|---|---|---|---|---|
| **Jointly fitted models** | | | | | | |
| Landmark | | | | | | |
| 2 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 5 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 10 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 15 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Median (year) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Mean (year) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Hazard pattern | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically increasing | Monotonically increasing | Increasing in the first 1 year; then decreasing | Increasing in the first 3 years; then decreasing | Monotonically increasing |
| After adjusting for background mortality | Constant in the first 15 years; then increasing | Monotonically increasing | Monotonically increasing | Increasing in the first 1 year; then decreasing for 11 years; then increasing | Increasing in the first 3 years; then decreasing for 10 years; then increasing | Monotonically increasing |
| **Independently fitted models** | | | | | | |
| Landmark | | | | | | |
| 2 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 5 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 10 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 15 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Median (year) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Mean (year) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Hazard pattern | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically increasing | Monotonically increasing | Increasing in the first 1.5 years; then decreasing | Increasing in the first 3 years; then decreasing | Monotonically increasing |
| After adjusting for background mortality | Constant in the first 15 years; then increasing | Monotonically increasing | Monotonically increasing | Increasing in the first 1 year; then decreasing for 12 years; then increasing | Increasing in the first 3 years; then decreasing for 11 years; then increasing | Monotonically increasing |

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; OS = overall survival

[a] Unless otherwise specified, the survival estimates were after adjustment by background mortality such that in any time during the model, the morality rate of modelled population would not be lower than that of general population

*B.3.3.2.2.4 Assessment of internal validity of PFS/TTD*

The results of BIC statistics showed that the exponential distribution provided the best fit consistently to the dependent models of PFS, the independent model of PFS for both treatment arms, and the independent model of TTD for zanubrutinib (matching DRC), separately, which was aligned with the clinical association between disease progression and treatment discontinuation.

*B.3.3.2.2.5 Summary of model selection*

Given the above, a dependent gamma model was applied for OS in the base-case analysis because (1) it was associated with clinically plausible hazard patterns for both treatment arms; (2) it was associated with clinically plausible mean OS for both treatment arms; (3) it was associated with the second lowest BIC.

For PFS, a dependent exponential model was applied for both treatment arms, whereas for TTD, an independent exponential model was applied for zanubrutinib (matching DRC) in the base-case analysis, given that the exponential distribution was associated with the lowest BIC for both PFS and TTD. It was also aligned with the clinical association between disease progression and treatment discontinuation.

Additional scenarios of survival extrapolation were also explored (e.g., dependent Weibull and dependent Gompertz models for OS; see Section B.3.8.3).

### *B.3.3.2.3 Zanubrutinib versus bendamustine-rituximab*

Figure B.3.24 to Figure B.3.26 present the KM curves for PFS, OS, and TTD for zanubrutinib versus BR, based on results from the MAIC (see Section B.2.9).

**Figure B.3.24. KM curves of PFS – zanubrutinib vs BR**

Abbreviations: BR = bendamustine and rituximab; KM = Kaplan-Meier; PFS = progression-free survival; vs = versus

**Figure B.3.25. KM curves of OS – zanubrutinib vs BR**

Abbreviations: BR = bendamustine and rituximab; KM = Kaplan-Meier; OS = overall survival; vs = versus

**Figure B.3.26. KM curves of TTD – zanubrutinib**

Abbreviations: BR = bendamustine and rituximab; KM = Kaplan-Meier; TTD = time to treatment discontinuation

*B.3.3.2.3.1 Assessment of PH assumption*

The PH assumption was assessed through visual inspection of the log-cumulative hazard plots for PFS and OS of zanubrutinib (matching BR) and BR (Figure B.3.27 and Figure B.3.28). The plots show relatively straight and parallel curves, which supports the use of a single model with treatment group included as a covariate, while at certain time points the plots appeared

to be unparallel or diverge. Therefore, both dependent models (with treatment included as a covariate) and independent models were assessed, each using six parametric distributions.

**Figure B.3.27. Log-cumulative hazards vs log time for PFS – zanubrutinib vs BR**

Abbreviations: BR = bendamustine and rituximab; PFS = progression-free survival; vs = versus

**Figure B.3.28. Log-cumulative hazards vs log time for OS – zanubrutinib vs BR**

Abbreviations: BR = bendamustine and rituximab; OS = overall survival

*B.3.3.2.3.2 Assessment of internal validity of OS/PFS/TTD*

Goodness-of-fit of parametric distributions were assessed with AIC and BIC statistics (Table B.3.14–Table B.3.16) and a visual comparison of the KM curves against the parametric curves was performed (Figure B.3.29 to Figure B.3.35). The fit statistics indicated that the exponential distribution was associated with the lowest BIC across all the parametric models. The results of fit statistics were validated through visual inspections.

**Table B.3.14. Fit statistics for jointly fitted PFS, OS and TTD – zanubrutinib (match BR) vs BR**

| Parametric distribution | PFS | | OS | | TTD | |
|---|---|---|---|---|---|---|
| | AIC | BIC | AIC | BIC | AIC | BIC |
| Exponential | [redacted] | [redacted] | [redacted] | [redacted] | N/A | |
| Weibull | [redacted] | [redacted] | [redacted] | [redacted] | | |
| Gompertz | [redacted] | [redacted] | [redacted] | [redacted] | | |
| Log-normal | [redacted] | [redacted] | [redacted] | [redacted] | | |
| Log-logistic | [redacted] | [redacted] | [redacted] | [redacted] | | |
| Gamma | [redacted] | [redacted] | [redacted] | [redacted] | | |

Abbreviations: AIC = Akaike information criteria; BIC = Bayesian information criteria; BR = bendamustine and rituximab; N/A = not applicable; OS = overall survival; PFS = progression-free survival; TTD = time to discontinuation; vs = versus
Note: The lowest AIC or BIC is in italics

**Table B.3.15. Fit statistics for independently fitted PFS, OS and TTD – zanubrutinib (match BR)**

| Parametric distribution | PFS | | OS | | TTD | |
|---|---|---|---|---|---|---|
| | AIC | BIC | AIC | BIC | AIC | BIC |
| Exponential | [redacted] | | | | | |
| Weibull | | | | | | |
| Gompertz | | | | | | |
| Log-normal | | | | | | |
| Log-logistic | | | | | | |
| Gamma | | | | | | |

Abbreviations: AIC = Akaike information criteria; BIC = Bayesian information criteria; BR = bendamustine and rituximab; OS = overall survival; PFS = progression-free survival; TTD = time to discontinuation
Note: The lowest AIC or BIC is in italics

**Table B.3.16. Fit statistics for independently fitted PFS, OS and TTD – BR**

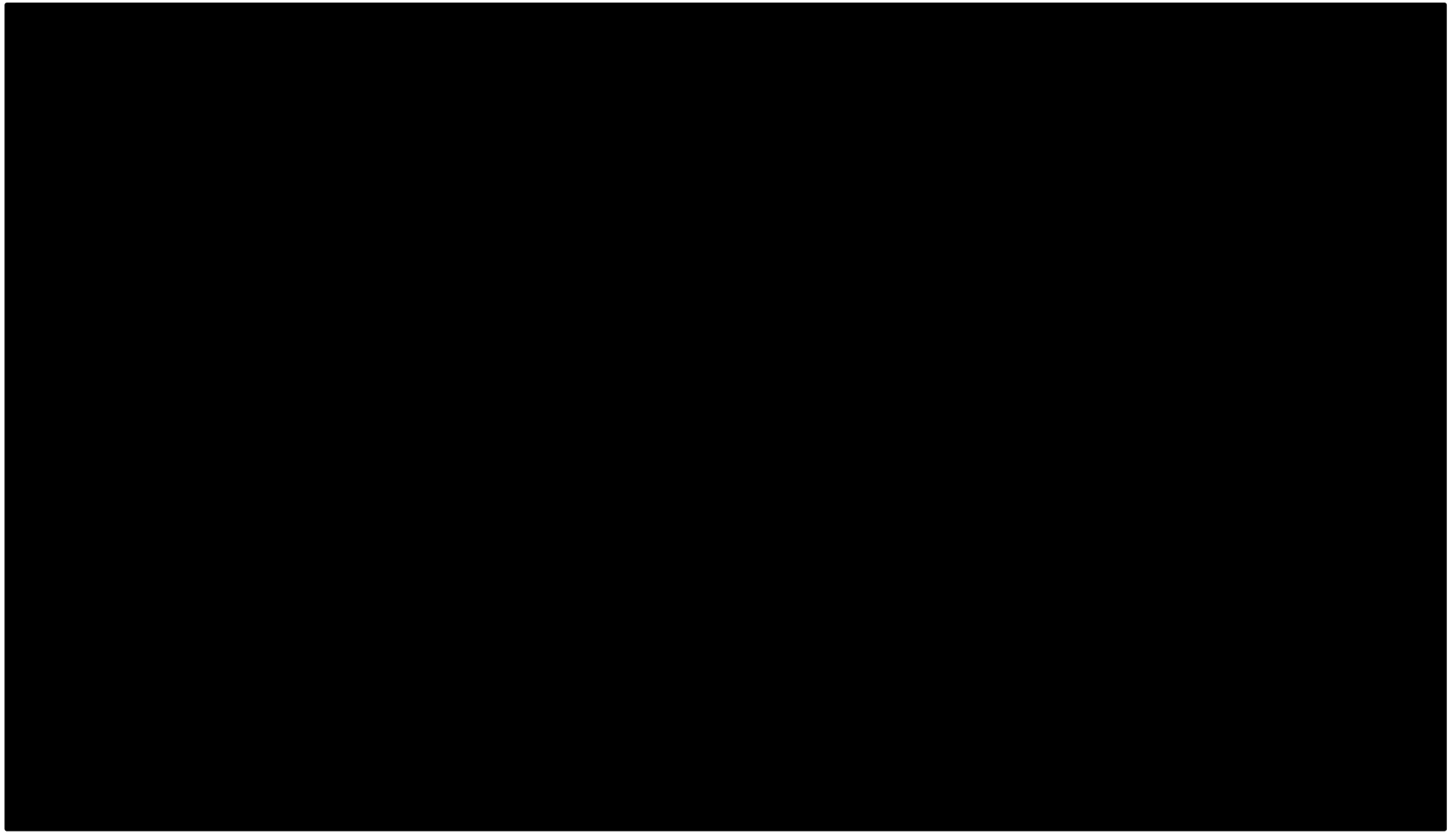
| Parametric distribution | PFS | | OS | | TTD | |
|---|---|---|---|---|---|---|
| | AIC | BIC | AIC | BIC | AIC | BIC |
| Exponential | [redacted] | | | | N/A | |
| Weibull | | | | | | |
| Gompertz | | | | | | |
| Log-normal | | | | | | |
| Log-logistic | | | | | | |
| Gamma | | | | | | |

Abbreviations: AIC = Akaike information criteria; BIC = Bayesian information criteria; BR = bendamustine and rituximab; N/A = not applicable; OS = overall survival; PFS = progression-free survival; TTD = time to discontinuation
Note: The lowest AIC or BIC is in italics

**Figure B.3.29. Visual inspection of jointly fitted parametric vs KM curves for PFS – zanubrutinib (match BR) vs BR**

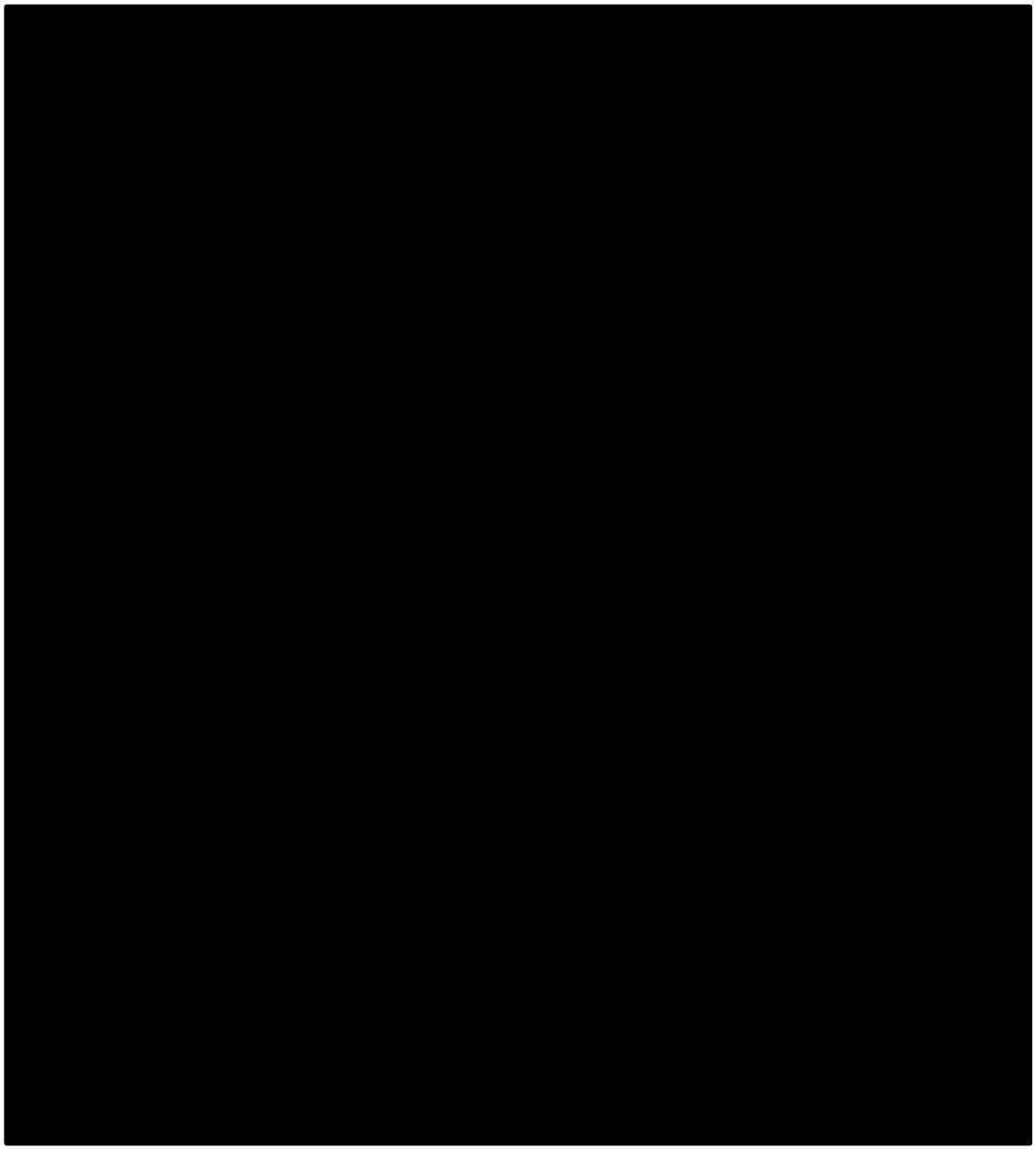
Abbreviations: BR = bendamustine and rituximab; KM = Kaplan-Meier; PFS = progression-free survival; vs = versus
Notes: (1) The graphs above are for assessment of internal validity only without adjusting for background mortality. (2) The 6 graphs on the left are identical to the 6 graphs on the right except for the 95% CI of the parametric curve in order for clearer visual inspection

**Figure B.3.30. Visual inspection of jointly fitted parametric vs KM curves for OS – zanubrutinib (match BR) vs BR**



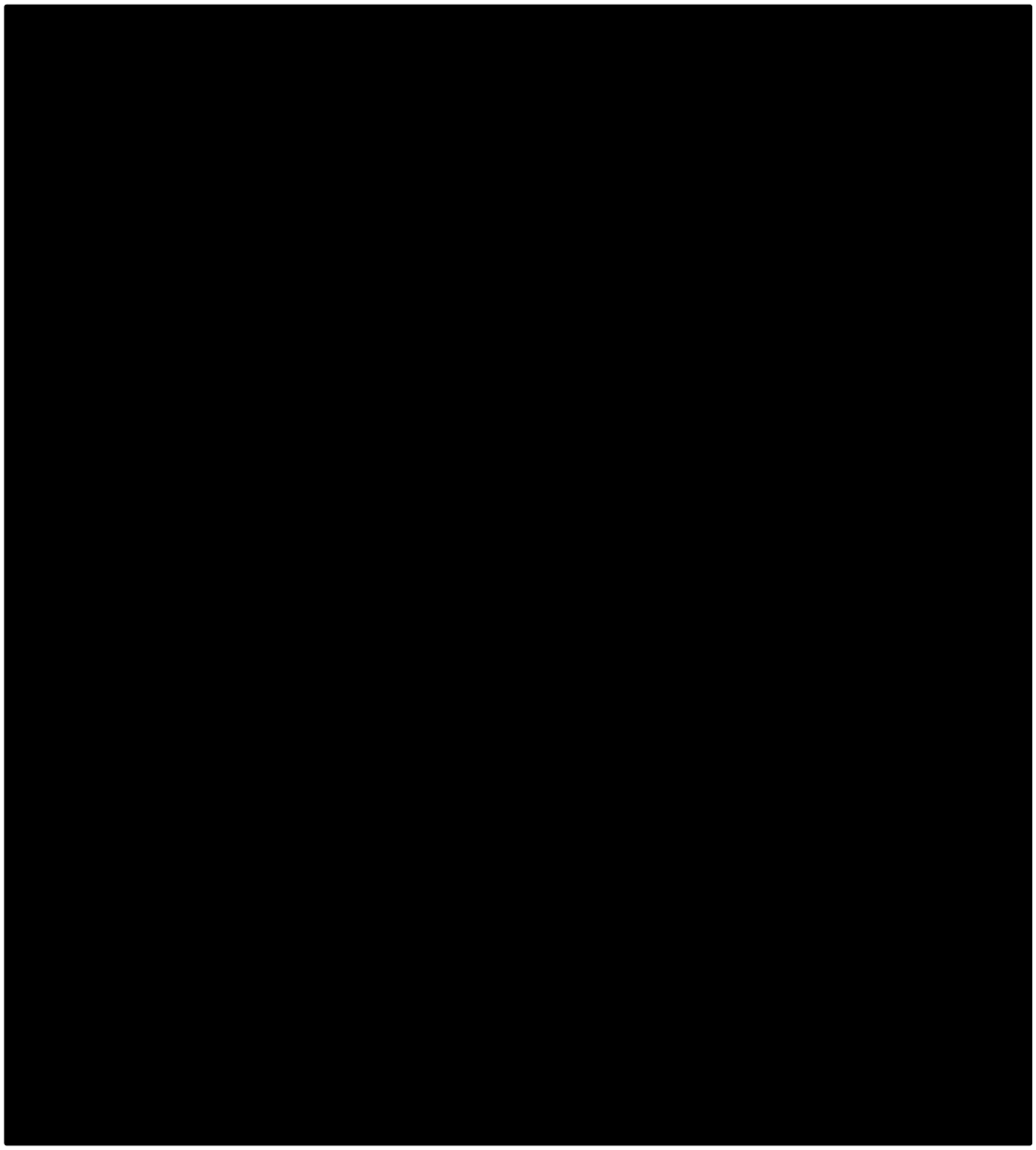
Abbreviations: BR = bendamustine and rituximab; KM = Kaplan-Meier; OS = overall survival; vs = versus
Notes: (1) The graphs above are for assessment of internal validity only without adjusting for background mortality. (2) The 6 graphs on the left are identical to the 6 graphs on the right except for the 95% CI of the parametric curve in order for clearer visual inspection

**Figure B.3.31. Visual inspection of independently fitted parametric vs KM curves for PFS – zanubrutinib (match BR)**



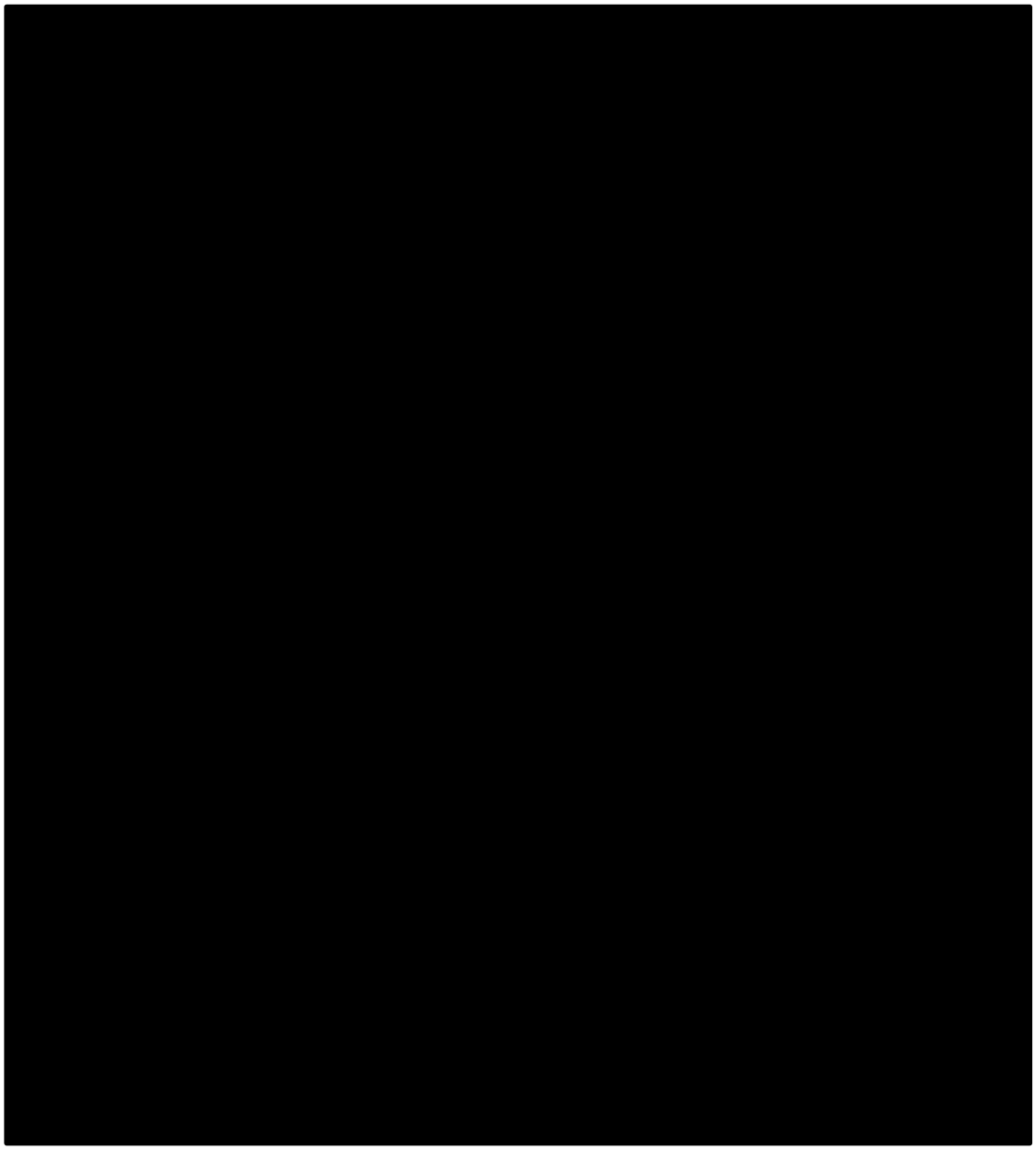
Abbreviations: BR = bendamustine and rituximab; KM = Kaplan-Meier; PFS = progression-free survival
Note: The graphs above are for assessment of internal validity only without adjusting for background mortality

**Figure B.3.32. Visual inspection of independently fitted parametric versus KM curves for PFS – BR**



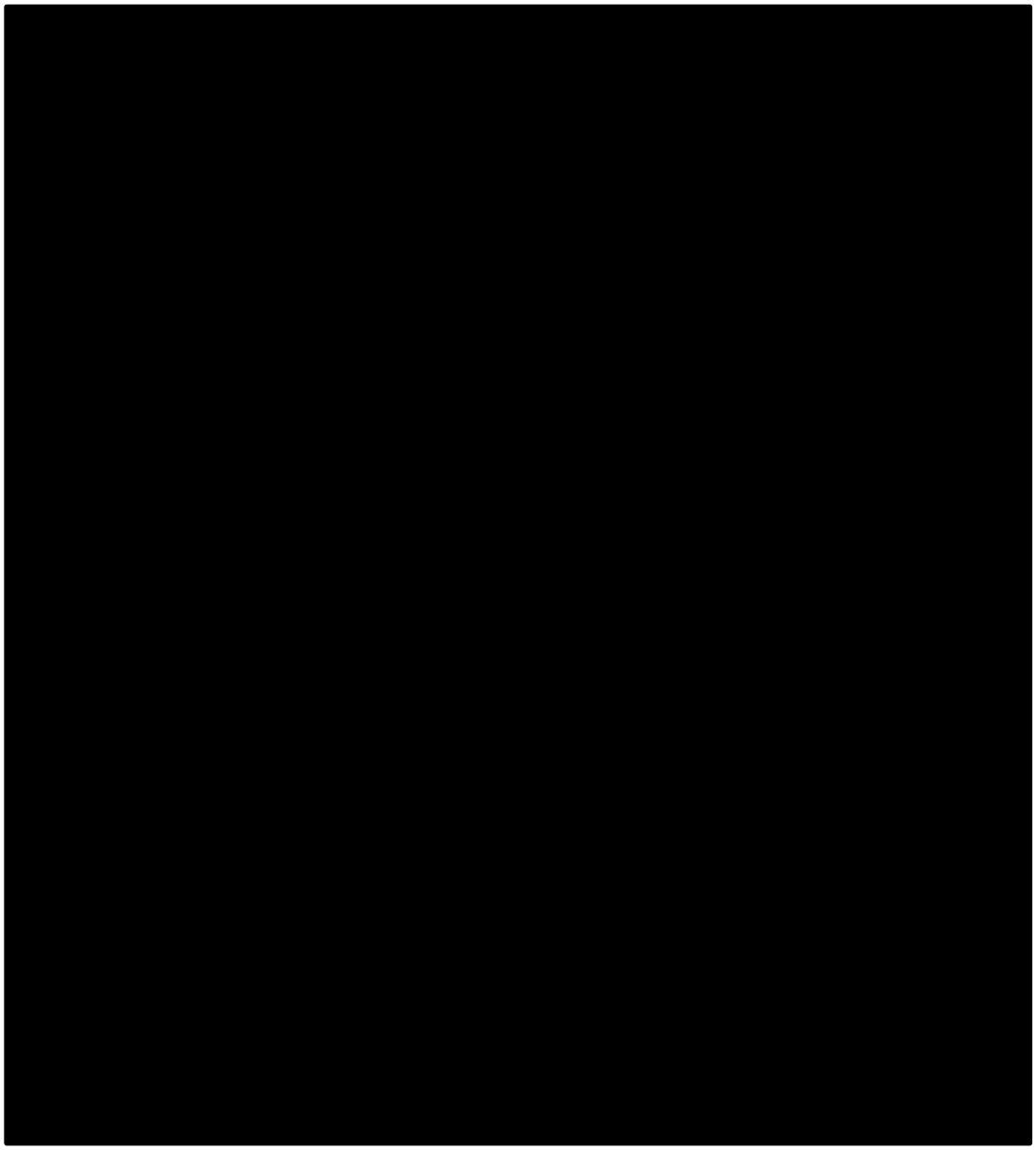
Abbreviations: BR = bendamustine and rituximab; KM = Kaplan-Meier; PFS = progression-free survival
Note: The graphs above are for assessment of internal validity only without adjusting for background mortality

**Figure B.3.33. Visual inspection of independently fitted parametric versus KM curves for OS – zanubrutinib (match BR)**



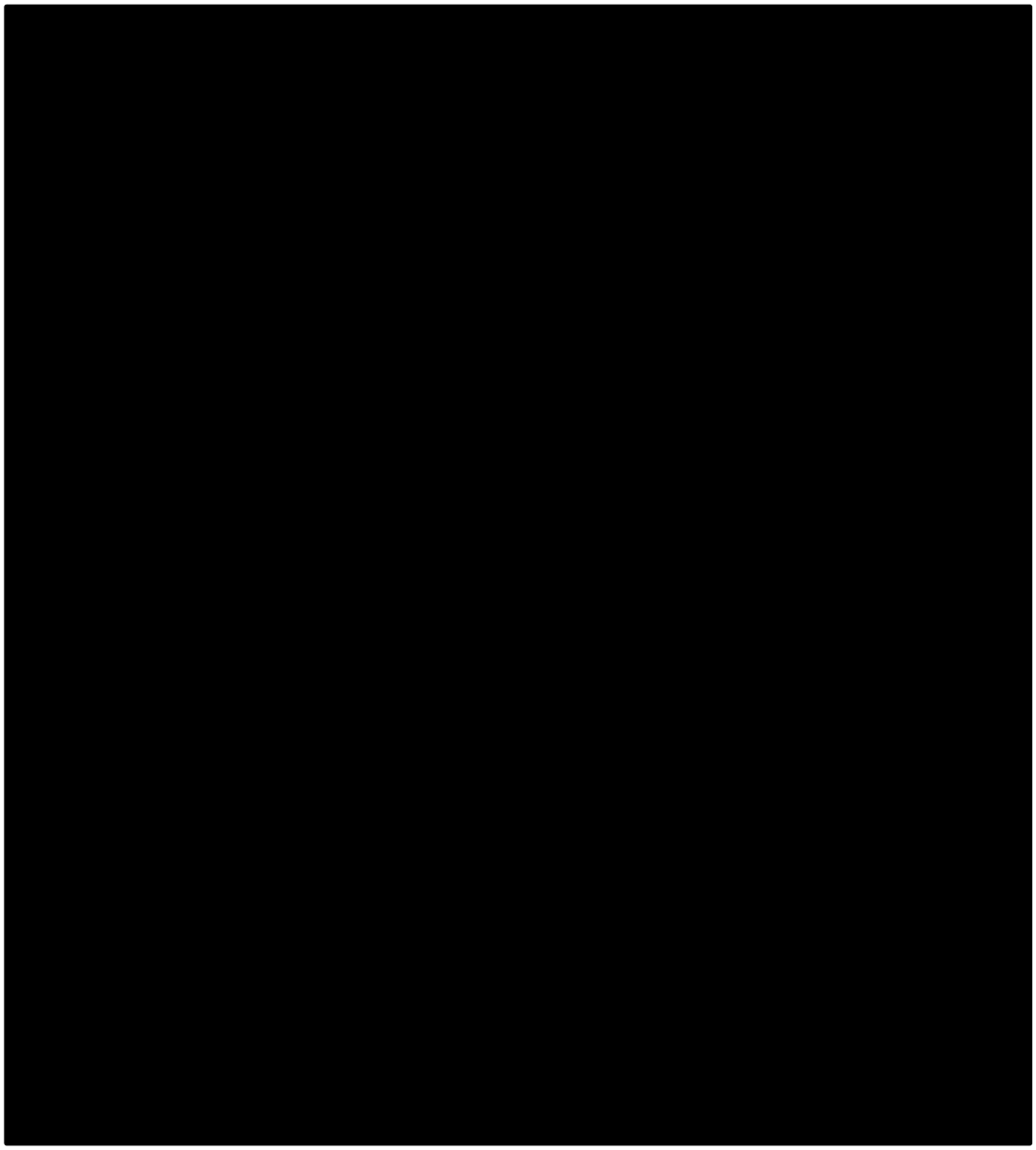
Abbreviations: BR = bendamustine and rituximab; KM = Kaplan-Meier; OS = overall survival
Note: The graphs above are for assessment of internal validity only without adjusting for background mortality

**Figure B.3.34. Visual inspection of independently fitted parametric versus KM curves for OS – BR**



Abbreviations: BR = bendamustine and rituximab; KM = Kaplan-Meier; OS = overall survival
Note: The graphs above are for assessment of internal validity only without adjusting for background mortality

**Figure B.3.35. Visual inspection of independently fitted parametric versus KM curves for TTD – zanubrutinib (match BR)**



Abbreviations: BR = bendamustine and rituximab; KM = Kaplan-Meier; TTD = time to treatment discontinuation
Note: The graphs above are for assessment of internal validity only without adjusting for background mortality

#### *B.3.3.2.3.3 Assessment of external validity of OS*

Despite that goodness-of-fit assessment supported the use of the exponential distribution for both zanubrutinib (matching BR) and BR, the fit statistics were close across parametric distributions. Therefore, and also given the immaturity of survival data, external validity was assessed.

For zanubrutinib (matching BR), as presented in Table B.3.17, all the jointly fitted parametric models generated relatively similar OS estimates with mean OS ranging between [redacted] years, which was close to but slightly lower than the OS estimates for zanubrutinib without matching adjustment (mean: [redacted] years, see Table B.3.7). In contrast, for the

independent models, other than the exponential model (mean: [redacted] years), all the distributions generated relatively shorter OS (mean: [redacted] years).

As per clinical expert opinion (see Section B.3.3.2.1.3), all the dependent models generated clinically plausible mean OS of approximately [redacted] years, among which the dependent Weibull (mean OS: [redacted] years) and dependent gamma (mean OS: [redacted] years) models were associated with monotonically increasing hazards (before and after adjusting for background mortality) that were considered to be more clinically plausible, assuming a relatively homogeneous population.

For BR, as presented in Table B.3.18, different parametric models generated a wide range of mean OS of [redacted] years. As described in Section B.3.3.2.2.3, clinical experts stated that for patients receiving chemo-immunotherapy as second- and third-line treatment, the plausible range of mean OS would be [redacted] and [redacted] years, respectively. Given that the BR population had received a median of 2 prior lines of treatments[49]), the dependent Weibull (mean OS: [redacted] years) appeared to be the most plausible. However, considering that patients in the zanubrutinib arm in ASPEN had received a median of 1 prior line of treatment, it might be more appropriate to apply the exponential, dependent Gompertz, dependent log-normal, dependent log-logistic, independent Weibull, and independent log-logistic models (with a mean OS falling between [redacted] years) for BR, among which the independent Weibull model (mean OS: [redacted] years) with monotonically increasing hazards (before and after adjusting for background mortality) was considered to be more clinically plausible for BR, assuming a relatively homogeneous population.

In summary, both the dependent Weibull and dependent gamma models were associated with clinically plausible mean survival and hazard patterns for zanubrutinib (matching BR) as well as clinically plausible hazard patterns for BR. However, both models might lead to an underestimation of the mean OS for BR. On the other hand, the independent Weibull model was associated with clinically plausible mean survival and hazard patterns for BR, but none of the independent models for zanubrutinib (matching BR) was associated with clinically plausible mean survival and hazard patterns simultaneously. Still, among all the independent models for zanubrutinib (matching BR), despite the constant hazard pattern (before adjusting for background mortality), the exponential model appeared to be the most clinically plausible with a mean OS of approximately [redacted] years. In light of the above, in the base-case analysis, an independent exponential model was applied for zanubrutinib, whereas an independent Weibull model was applied for BR. The dependent Weibull and dependent gamma models were also explored in scenario analyses.

**Table B.3.17. Landmark, median, mean and hazard patterns of OS – zanubrutinib (match BR)[a]**

| | Exponential | Weibull | Gompertz | Log-normal | Log-logistic | Gamma |
|---|---|---|---|---|---|---|
| **Jointly fitted models** | | | | | | |
| Landmark | | | | | | |
| 2 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 5 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 10 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 15 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Median (year) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Mean (year) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Hazard pattern | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically increasing | Monotonically decreasing | Increasing in the first 2 years; then decreasing | Increasing in the first 5 years; then decreasing | Monotonically increasing |
| After adjusting for background mortality | Constant for 7 years; then increasing | Monotonically increasing | Decreasing in the first 5 years; the increasing | Increasing in the first 2 years; then decreasing/ stable for 5 years; then increasing | Increasing in the first 5 years; then stable for 5 years; then increasing | Monotonically increasing |
| **Independently fitted models** | | | | | | |
| Landmark | | | | | | |
| 2 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 5 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 10 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 15 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Median (year) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Mean (year) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Hazard pattern | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically increasing | Monotonically increasing | Increasing in the first 4 years; then decreasing | Increasing in the first 6 years; then decreasing | Monotonically increasing |
| After adjusting for background mortality | Constant for 7 years; then increasing | Monotonically increasing | Monotonically increasing | Increasing in the first 3 years; then decreasing for 9 years; then increasing | Increasing in the first 6 years; then decreasing for 12 years; then increasing | Monotonically increasing |

Abbreviations: BR = bendamustine and rituximab; OS = overall survival

[a] Unless otherwise specified, the survival estimates were after adjustment by background mortality such that in any time during the model, the morality rate of modelled population would not be lower than that of general population

**Table B.3.18. Landmark, median, mean and hazard patterns of OS – BR[a]**

| | Exponential | Weibull | Gompertz | Log-normal | Log-logistic | Gamma |
|---|---|---|---|---|---|---|
| **Jointly fitted models** | | | | | | |
| Landmark | | | | | | |
| 2 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 5 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 10 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 15 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Median (year) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Mean (year) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Hazard pattern | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically increasing | Monotonically decreasing | Increasing in the first 1 year; then decreasing | Increasing in the first 2 years; then decreasing | Monotonically increasing |
| After adjusting for background mortality | Constant for 17 years; then increasing | Monotonically increasing | Decreasing in the first 13 years; the increasing | Increasing in the first 1 years; then decreasing for 10 years; then increasing | Increasing in the first 2 years; then decreasing for 11 years; then increasing | Monotonically increasing |
| **Independently fitted models** | | | | | | |
| Landmark | | | | | | |
| 2 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 5 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 10 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| 15 years | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Median (year) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Mean (year) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Hazard pattern | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically increasing | Monotonically decreasing | Increasing in the first 1 year; then decreasing | Increasing in the first 1 year; then decreasing | Monotonically increasing |
| After adjusting for background mortality | Constant for 17 years; then increasing | Monotonically increasing | Decreasing in the first 7 years; then increasing | Increasing in the first 1 year; then decreasing for 10 years; then increasing | Increasing in the first 1 year; then decreasing for 12 years; then increasing | Monotonically increasing |

Abbreviations: BR = bendamustine and rituximab; OS = overall survival

[a] Unless otherwise specified, the survival estimates were after adjustment by background mortality such that in any time during the model, the morality rate of modelled population would not be lower than that of general population

*B.3.3.2.3.4 Assessment of external validity of PFS/TTD*

The results of BIC statistics showed that the exponential distribution provided the best fit to both PFS and TTD of zanubrutinib (matching BR), which was aligned with the clinical association between disease progression and treatment discontinuation.

*B.3.3.2.3.5 Summary of model selection*

Given the above, for OS, an independent Weibull model and an independent exponential model was applied for BR and zanubrutinib (matching BR), respectively, as (1) it was associated with clinically plausible mean OS for both treatment arms, and (2) it was associated with a clinically plausible hazard pattern for BR.

For PFS and TTD of zanubrutinib (matching BR) and PFS of BR, the exponential distribution was applied in the base-case analysis, given that (1) the log-cumulative hazard plots were relatively parallel; (2) the exponential distribution was associated with the lowest BIC across all the distributions; and (3) it was aligned with the clinical association between disease progression and treatment discontinuation.

Additional scenarios of survival extrapolation were also explored (e.g., dependent Weibull and gamma models for OS; see Section B.3.8.3).

### B.3.3.3 Adverse events

AEs of Grade ≥3 that occurred in ≥5% of patients in any treatment arm were included in the model to capture the effects on costs and HRQoL. Incidence and duration of AEs for each treatment were based on the clinical studies from which the survival outcomes were obtained (Table B.3.19).

**Table B.3.19. Incidence and duration of Grade ≥3 AEs occurring in ≥5% of patients in any treatment arm**

| | AE incidence, % | | | | | | Duration, days |
|---|---|---|---|---|---|---|---|
| | **Zanubrutinib (N=101)** | **Ibrutinib (N=98)** | **Zanubrutinib adjusted to match BR ($n_{eff}$= [redacted])** | **BR (N=71)** | **Zanubrutinib adjusted to match DRC ($n_{eff}$= [redacted])** | **DRC (N=72)** | **ASPEN Safety Analysis Set (N=199)** |
| Reference | ASPEN IPD | ASPEN IPD | ASPEN IPD (match BR) | Tedeschi et al. 2015[49] | ASPEN IPD, (match DRC) | Dimopoulos et al. 2007[50] | ASPEN IPD |
| Anaemia | 4.95 | 5.10 | [redacted] | NR[a] | [redacted] | NR[a] | 17.0 |
| Hypertension | 5.94 | 11.22 | [redacted] | NR[a] | [redacted] | NR[a] | 20.9 |
| Neutropenia | 15.84 | 8.16 | [redacted] | 35.21 | [redacted] | 10.00 | 10.9 |
| Pneumonia | 0.99 | 7.14 | [redacted] | 5.63 | [redacted] | NR[a] | 21.3 |
| Thrombocytopenia | 5.94 | 3.06 | [redacted] | NR[a] | [redacted] | 0.00 | 28.8 |

Abbreviations: BR = bendamustine and rituximab; DRC = dexamethasone, rituximab and cyclophosphamide; IPD = individual patient-level data; N = number of patients evaluable; $n^{eff}$ = effective sample size; NR = not reported

[a] It was conservatively assumed that the unreported incidences were 0 for comparators

### B.3.3.4 Background mortality

UK background mortality was based on the UK National Life Tables, United Kingdom 2016–2018.[61] As specified in Section B.3.2.2, background mortality was applied in the model such that at any time during the model, the mortality rates of the modelled population would not be lower than that of the UK background mortality, adjusted by average age and sex ratio.

## B.3.4 Measurement and valuation of health effects

### B.3.4.1 Health-related quality-of-life data from clinical trials

The utility analysis was performed on the full analysis set (FAS), which comprised all patients included in ASPEN (i.e., ITT population). The EQ-5D-5L questionnaire was used to measure health utilities. Among the ITT population, patients were excluded from the utility analysis if they did not have at least one complete EQ-5D-5L measurement (utility analysis population). A total of 193 patients were included in the utility analysis.

In ASPEN, EQ-5D-5L data were collected at screening (pre-treatment), every 12 weeks during the first 48 weeks (starting Cycle 4 Day 1), then every 24 weeks (every 6 cycles) thereafter during the treatment period (Table B.3.20). The average utility value for overall ASPEN patient population before progression was estimated to be 0.791.

**Table B.3.20. Scheduled assessment of EQ-5D-5L**

| | **Pre-treatment** | **During treatment Each cycle = 28 days** | | | | | **End of treatment** |
|---|---|---|---|---|---|---|---|
| | **Screening** | **Cycle 4** | **Cycle 7** | **Cycle 10** | **Cycle 13** | **…** | **≤7 days after last dose** |
| EQ-5D-5L | x | Every 12 weeks (starting cycle 4 day 1) during the first 48 weeks (ending cycle 13 day 1), then every 24 weeks (i.e., every 6 cycles) thereafter | | | | | |

Abbreviations: EQ-5D = EuroQol 5-Dimension 5-Level

The utility analysis focused on the period before progression. While ASPEN captured some EQ-5D-5L measurements after progression, data were limited and therefore were not used to derive utility value after progression. Missing data were not imputed.

The responses obtained from the EQ-5D-5L questionnaire in ASPEN were converted into a single utility score using the UK value set. The EQ-5D-5L data were first mapped to the EuroQol 5-Dimensions 3-Level (EQ-5D-3L) using the crosswalk described by van Hout et al. (2012).[62] The EQ-5D-3L value set, proposed by Dolan (1997), was then used to derive utility values.[63] This was consistent with the position statement by NICE that the EQ-5D-5L value set for England published by Devlin et al. (2018) is not recommended.[64] Following the mapping process, the utility values ranged from 1 for the “Perfect Health” state, where all the EQ-5D dimensions are equal to 1 (11111), to -0.594 for the worst health state, where all the EQ-5D dimensions are equal to 5 (55555).

The EQ-5D-5L responses were collected repeatedly over time for the same patient. The observations tended to be correlated between time points, resulting in non-independence of the data. To account for the repeated nature of the data and explore the influence on EQ-5D-5L utility values of demographic characteristics and time from treatment, linear mixed effects models (LMM) for repeated measures were used to derive the EQ-5D-5L utility values in the pre-progression health state. LMMs utilised observations considering the correlation between repeated measurements and provided the option to include fixed and random effect terms for

time and interactions with baseline covariates. Thus, LMMs produced unbiased estimates of the impact of risk factors under the missing-at-random assumption, representing a robust method to handle missing data within reasonable limits. For this reason, LMMs are often used to analyse EQ-5D-5L data given the longitudinal and hierarchical nature of data (Level 1 = repeated measures; Level 2 = the patient).

The statistical models included EQ-5D-5L utility values as a dependent variable. To determine the relevant covariates, different regression models were implemented by including an additional independent variable in each model. The potential covariates that were investigated were:

- Treatment group and demographic characteristics:
  - Treatment ($Tx_{zanu}$) - dummy variable equal to 1 if a patient is in the zanubrutinib group
  - Age ($age_i$) – continuous variable
  - Sex ($sex_i$) – dummy variable equal to 1 if a patient is male
- Assessment time point, defined as:
  - A variable counting the days from treatment initiation ($day_t$), e.g., for Screening $day_t$ will be -35 to -1 (day) while for Cycle 4 Day 1 $day_t$ will be 84 (28 x 3 = 84) (days).
  - A variable accounting for the number of cycles (of treatment completed at the visit) e.g., for Screening $visit_t$ is equal to 0, while for Cycle 4 $visit_t$ is equal to 3.

By adding a covariate each time, three different models were fitted, where the term $U_{it}$ denoted the EQ-5D-5L utility value measured for patient *i* at time *t* and $\varepsilon_{it}$ was the residual random error for patient *i* at time *t*. A summary of the regression models was shown in Table B.3.21.

**Table B.3.21. Regression models estimated in the utility analysis**

| Model | Model specification |
|---|---|
| Model 1 | $U_{it} = \alpha + \beta_1 Tx_{Zanu} + \beta_2 age_i + \beta_3 sex_i + \varepsilon_{it}$ |
| Model 2 | $U_{it} = \alpha + \beta_1 Tx_{Zanu} + \beta_2 age_i + \beta_3 sex_i + \beta_4 day_t + \varepsilon_{it}$ |
| Model 3 | $U_{it} = \alpha + \beta_1 Tx_{Zanu} + \beta_2 age_i + \beta_3 sex_i + \beta_4 visit_t + \varepsilon_{it}$ |

All analyses were conducted using SAS. For each model, three specifications were tested including (1) random intercept, (2) random slope and (3) random intercept-slope. This specification took account of the repeated measures in the data which might introduce non-independence of EQ-5D-5L reporting. The models were fitted with identical fixed effects structures and least square mean estimates of the EQ-5D-5L utility values and the related standard errors were generated.

The regression models were subsequently assessed using the AIC and the BIC statistics. The optimal model was defined as the model which best reflected reality and generated plausible results. The optimal model was selected based on the level of significance and the magnitude of each estimated coefficients and the AIC and BIC statistics.

The random intercept specification was selected for Model 1, Model 2 and Model 3, given that its results were statistically significant. The pre-progression utility values from each model are summarised in Table B.3.22. Model 3 reports the lower AIC and BIC among all the regression

models and hence the pre-progression health state utility value from Model 3 (0.7908) is recommended for use in the cost-effectiveness model.

**Table B.3.22. Summary of pre-progression health state utility values**

| | Model 1 | | Model 2 | | Model 3 | |
|---|---|---|---|---|---|---|
| | **Zanubrutinib (N=99)** | **Ibrutinib (N=93)** | **Zanubrutinib (N=99)** | **Ibrutinib (N=93)** | **Zanubrutinib (N=99)** | **Ibrutinib (N=93)** |
| LS Mean (SE)a | 0.7917 (0.0170) | 0.7901 (0.0176) | 0.7921 (0.0068) | 0.7899 (0.0071) | 0.7919 (0.0068) | 0.7896 (0.0071) |
| AIC | -895.9 | | -892.1 | | -897.6 | |
| BIC | -886.2 | | -882.3 | | -887.9 | |
| Weighted LS Mean across treatment (SE)[b] | 0.7909 (0.0122) | | 0.7910 (0.0049) | | 0.7908 (0.0049) | |

Abbreviations: AIC = Akaike information criteria; BIC = Bayesian information criteria; LS = least square; N = number of patients evaluable; SE = standard error
[a] LS Means were adjusted at mean age of 69.07 years old and average of 68% male in the population. For Models 2 and 3, the LS Means were weighted average of the LS Means at scheduled time points, with weights as the number of observations at each time point over the total number of observations
[b] Weights were the proportions of patients in each treatment arm

### B.3.4.2 Mapping

EQ-5D-5L values were collected directly from ASPEN. Hence, no mapping was required.

### B.3.4.3 Health-related quality-of-life studies

As described in Appendix H, there were no published HRQoL studies that reported health utilities in patients with WM.

### B.3.4.4 Adverse reactions

Given the relatively short duration of AEs (as shown in Table B.3.19) and the large time interval between EQ-5D-5L observations (i.e., every 12 weeks during the first 48 week, and then every 24 weeks thereafter),[46] it is expected that QALY loss due to AEs were not captured in the health state utility. Therefore, in the base-case analysis, QALY losses due to AE Grade ≥3 were included in the model on top of the health state utilities, estimated as the sum product of AE disutilities (Table B.3.23), AE incidence and duration (Table B.3.19) for each AE. A scenario analysis was conducted without including AE disutilities (see Section B.3.8.3).

**Table B.3.23. AE disutilities**

| AE | Disutility | Source |
|---|---|---|
| Anaemia | 0.088 | NICE TA491[3] |
| Hypertension | 0.195 | Assumed to be the same as that for pneumonia, in line with the assumption adopted in NICE TA429 for ibrutinib in CLL[65] |
| Neutropenia | 0.185 | NICE TA491[3] |
| Pneumonia | 0.195 | NICE TA491[3] |
| Thrombocytopenia | 0.123 | NICE TA491[3] |

Abbreviations: AE = adverse event; CLL = chronic lymphocytic leukaemia; NICE = National Institute for Health and Care Excellence; TA = technology appraisal

### B.3.4.5 Health-related quality-of-life data used in the cost-effectiveness analysis

Total QALYs were derived by multiplying the time spent in each health state by the utility associated with the health state (Table B.3.24).

The utility for pre-progression survival health state (0.791) was estimated through a utility analysis using the EQ-5D-5L data collected in ASPEN. However, due to the very limited number of observations from patients who progressed, it was not feasible to derive a utility value for the post-progression survival health state from ASPEN trial data.

The utility for post-progression survival health state was instead calculated assuming a utility decrement of 0.100 relative to pre-progression survival health state utility (0.691). This utility decrement was based on the utility decrements for progression applied in NICE TA502 (0.10)[66] for ibrutinib in MCL and TA429 (0.098)[65] for ibrutinib in CLL, given that HRQoL data was collected until the end of primary treatment with BTK inhibitors in ASPEN. The estimated utility for post-progression survival (0.691) was then generally in line with NICE TA491 (ibrutinib in WM), where the utility value was 0.799 for the second-line progression-free health state, 0.799 for the third-line progression-free health state, 0.799 for the fourth-line progression-free health state and 0.665 for BSC health state.[3] The first three utility values were derived from the RESONATE CLL trial for ibrutinib, while the utility value for the BSC health state was based on a combination of RESONATE trial data and literature.

**Table B.3.24. Health state utilities**

| Health State | Utility | Source |
|---|---|---|
| Pre-progression survival | 0.791 | ASPEN IPD[a] |
| Post-progression survival | 0.691 | Assuming a utility decrement of 0.100 due to progression[65, 66] |

Abbreviation: IPD = individual patient-level data
[a] EQ-5D-5L collected in ASPEN was first mapped to the EQ-5D-3L using the mapping function.[62] The mapped EQ-5D-3L was then used to derive utilities using the EQ-5D-3L value sets[63]

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

### B.3.5.1 Intervention and comparators' costs and resource use

Table B.3.25 presents the drug acquisition costs applied in the model.

**Table B.3.25. Drug acquisition costs**

| Regimen | Package Size | Package Price | Reference |
|---|---|---|---|
| Zanubrutinib | 120 80mg capsules | £4,928.65 | Proposed list price |
| Ibrutinib | 28 420mg capsules | £4,292.40 | British National Formulary[67] |
| BR | | | |
| Bendamustine | 1 100mg vial | £262.02 | British National Formulary[67] |
| Rituximab[a] (MabThera) | 1 500mg vial | £873.15 | British National Formulary[67] |
| Rituximab (Truxima/Rixathon) | 1 500mg vial | £785.84 | British National Formulary[67] |
| DRC | | | |
| Dexamethasone | 10 4mg vials | £19.99 | British National Formulary[67] |
| Rituximab[a] | (as above) | | British National Formulary[67] |

| Regimen | Package Size | Package Price | Reference |
|---|---|---|---|
| Cyclophosphamide | 100 50mg tablets | £139 | British National Formulary[67] |

Abbreviations: BR = bendamustine and rituximab; DRC = dexamethasone, rituximab and cyclophosphamide; NHL = non-Hodgkin's lymphoma; WM = Waldenström's macroglobulinaemia
[a] It was assumed in the base-case analysis that MabThera accounted for 100% of use of rituximab

Table B.3.26 presents treatment administration costs applied in the model for intravenously administered drugs.

**Table B.3.26. Drug administration costs**

| Treatment | Administration route | Unit cost per administration | Reference |
|---|---|---|---|
| Zanubrutinib | Oral | 0.00 | Assumption |
| Ibrutinib | Oral | 0.00 | Assumption |
| BR | | | |
| Bendamustine | IV | 336.14 | NHS reference cost 2018-2019[68] |
| Rituximab | IV | 336.14 | NHS reference cost 2018-2019[68] |
| DRC | | | |
| Dexamethasone | IV | 336.14 | NHS reference cost 2018-2019[68] |
| Rituximab | IV | 336.14 | NHS reference cost 2018-2019[68] |
| Cyclophosphamide | Oral | 0.00 | Assumption |

Abbreviations: BR = bendamustine and rituximab; DRC = dexamethasone, rituximab and cyclophosphamide; IV = intravenous; NHS = National Health Services

### B.3.5.2 Health-state unit costs and resource use

Table B.3.27 presents the frequencies and unit costs of resource use for routine care. Given a lack of published studies reporting the resource use in patients with WM in the UK (see Appendix I), the frequency of resource use was based on those applied in NICE TA491 (ibrutinib in WM),[3] assuming that the resource use in this study would be the same as that used previously.[3]

**Table B.3.27. Frequencies and unit costs of resource use for routine care**

| | Frequency per year | | | Reference | Unit cost, £ | Reference |
|---|---|---|---|---|---|---|
| | Year 1–2 | Year 3–5 | Year 6+ | | | |
| Full blood count | 5 | 4 | 3 | NICE TA491[3] | 2.87 | NHS reference cost 2018-2019, DAPS05 Haematology[68] |
| Immunoglobulin | 5 | 4 | 3 | NICE TA491[3] | 6.72 | NHS reference cost 2018-2019, DAPS06 Immunology[68] |
| Chemistry | 5 | 4 | 3 | NICE TA491[3] | 1.14 | NHS reference cost 2018-2019, DAPS04 Clinical biochemistry[68] |
| Haematologist | 5 | 4 | 3 | NICE TA491[3] | 135.59 | NHS reference cost 2018-2019, WF01A Clinical haematology, consultant-led, non-admitted face to face follow-up[68] |
| Plasma viscosity | 5 | 4 | 3 | NICE TA491[3] | 6.75 | NHS reference cost 2018-2019, DAPS06 Immunology[68] |

| | Frequency per year | | | Reference | Unit cost, £ | Reference |
|---|---|---|---|---|---|---|
| | Year 1–2 | Year 3–5 | Year 6+ | | | |
| Paraprotein | 5 | 4 | 3 | NICE TA491[3] | 1.13 | NHS reference cost 2018-2019, DAPS04 Clinical biochemistry[68] |

Abbreviations: DAPS = Directly Accessed Pathology Services; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; TA = technology appraisal; WM = Waldenström's macroglobulinaemia

### B.3.5.3 Adverse reaction unit costs and resource use

Total management costs for AEs Grade ≥3 were applied as a one-off cost during the first model cycle, estimated as the sum product of the AE incidence (Table B.3.19) and the associated unit costs (Table B.3.28) of each AE.

**Table B.3.28. AE costs**

| AE type | Unit cost, £ | Source |
|---|---|---|
| Infections (mainly sepsis)[a] | 1,481.76 | NHS reference cost 2018-2019[68] |
| AEs other than infections[b] | 179.94 | NHS reference cost 2018-2019[68] |

Abbreviation: AE = adverse event; NHS = National Health Service

[a] The cost was estimated based on the weighted average of costs for Infections or other complications of procedures, without interventions, with CC Score 0-<4 (codes: WH07F – WH07G in NHS reference cost 2018-2019)

[b] The cost was estimated based on the weighted average of costs for Non-Admitted Face to Face Attendance - Clinical Haematology (codes: WF01A, WF01B, WF01C, WF01D, WF02A, WF02B, WF02C, WF02D in NHS reference cost 2018-2019)

### B.3.5.4 Miscellaneous unit costs and resource use

#### *B.3.5.4.1 Subsequent treatment*

Subsequent treatments were included in the model, where it was assumed that patients would receive subsequent treatments upon disease progression based on PFS. Although subsequent anticancer treatment was collected in ASPEN for BTK inhibitors, the data were immature. In addition, time to next anticancer treatment was not available for BR and DRC. Therefore, PFS was applied as a proxy for time to next anticancer treatment for these therapies. Data from literature and previous HTA submissions were used to inform the subsequent treatment use and distribution in the model, as presented in Table B.3.29. The proportion of patients receiving subsequent treatment upon progression was obtained from NICE TA491 (ibrutinib in WM) and was assumed to be the same across treatment arms.[3] Distribution of subsequent treatments was based on the first UK WM registry report from the Rory Morrison Registry.[1]

**Table B.3.29. Subsequent treatment use and distribution**

| Treatment regimen at model entry | Subsequent treatment use, % | Subsequent treatment distribution, % | | |
|---|---|---|---|---|
| | | Ibrutinib | BR | DRC |
| Zanubrutinib (with or without matching adjustment) | 86[a] | 0.0 [b] | 60.4 [b] | 39.6 [b] |
| Ibrutinib | 86[a] | 0.0 [b] | 60.4 [b] | 39.6 [b] |
| BR | 86[a] | 72.0 [b] | 0.0 | 28.0 [c] |
| DRC | 86[a] | 75.0 [c] | 25.0 [d] | 0.0 [c] |

Abbreviations: BR = bendamustine and rituximab; DRC = dexamethasone, rituximab and cyclophosphamide
[a] The estimate of 86% was based on the proportion of patients receiving third-line treatment among patients progressing from second-line treatment for WM based on UK clinical experts' opinions, reported in UK NICE TA491 (ibrutinib in WM)[3]
[b] The uptake of BR (32%; n=14/43) and DRC (21%; n=9/43) in patients with treatment-naïve WM[1] were adjusted such that the sum equals 100%
[c] The uptake of ibrutinib (18%) and DRC (7%) in patients with relapsed/refractory WM[1] were adjusted such that the sum equals 100%
[d] The uptake of ibrutinib (18%) and BR (6%) in patients with relapsed/refractory WM were adjusted such that the sum equals 100%

For patients receiving subsequent BTK inhibitors, it was assumed that patients would be treated with ibrutinib until death, based on the time patients spent in the post-progression survival health state. For patients treated with subsequent chemo-immunotherapy (i.e., BR or DRC), two lines of subsequent treatments (i.e., two treatment courses, each with 6 cycles) were assumed. This assumption was based on the number of treatment lines modelled in previous ibrutinib WM models in the UK and Italy (i.e., second- to fourth-line, and best supportive care afterwards).[3, 69]

The costs of drug and drug administration of subsequent ibrutinib, BR and DRC were the same as those as primary treatment (see Table B.3.25 and Table B.3.26).

#### *B.3.5.4.2 Terminal care*

A one-time terminal care cost (£7,978.35) was applied upon death, estimated based on sex-specific terminal care costs reported in published literature.[70]

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

### B.3.6.1 Summary of base-case analysis inputs

**Table B.3.30. Summary of variables applied in the economic analysis**

| Variable | Deterministic value (base-case analysis) | Distribution in DSA and PSA | Lower bound for DSA | Upper bound for DSA | Reference to section in submission |
|---|---|---|---|---|---|
| **General model settings** | | | | | |
| Time horizon, years | 30 | Fixed | | | B.3.2.2 |
| Discounting per year – costs | 3.5% | Fixed | | | |
| Discounting per year – clinical outcomes | 3.5% | Fixed | | | |
| **Baseline patient characteristics** | | | | | |
| Female proportion | 33.33% | Beta | 26.97% | 40.01% | B.3.3.1 |
| Mean age, year | 69.53 | Normal | 68.22 | 70.84 | |
| Body surface area, $m^2$ | 1.86 | Normal | 1.83 | 1.89 | |

| Variable | Deterministic value (base-case analysis) | Distribution in DSA and PSA | Lower bound for DSA | Upper bound for DSA | Reference to section in submission |
|---|---|---|---|---|---|
| **Survival parameters** | | | | | |
| PFS, OS and TTD – zanubrutinib | Parametric model | Multivariate normal | | | B.3.3.2 |
| PFS, OS and TTD – ibrutinib | Parametric model | Multivariate normal | | | |
| PFS – BR | Parametric model | Multivariate normal | | | |
| PFS and OS – DRC | Parametric model | Multivariate normal | | | |
| **AEs** | | | | | |
| AE incidence – zanubrutinib | | | | | B.3.3.3 |
| Anaemia | 4.95% | Beta | 1.63% | 9.96% | |
| Hypertension | 5.94% | Beta | 2.22% | 11.31% | |
| Neutropenia | 15.84% | Beta | 9.84% | 22.95% | |
| Pneumonia | 0.99% | Beta | 0.02% | 3.65% | |
| Thrombocytopenia | 5.94% | Beta | 2.22% | 11.31% | |
| AE incidence – ibrutinib | | | | | |
| Anaemia | 5.10% | Beta | 1.68% | 10.24% | |
| Hypertension | 11.22% | Beta | 5.77% | 18.18% | |
| Neutropenia | 8.16% | Beta | 3.60% | 14.35% | |
| Pneumonia | 7.14% | Beta | 2.93% | 13.01% | |
| Thrombocytopenia | 3.06% | Beta | 0.63% | 7.27% | |
| AE incidence – zanubrutinib (match BR) | | | | | |
| Anaemia | [redacted] | Beta | [redacted] | [redacted] | |
| Hypertension | [redacted] | Beta | [redacted] | [redacted] | |
| Neutropenia | [redacted] | Beta | [redacted] | [redacted] | |
| Pneumonia | [redacted] | Beta | [redacted] | [redacted] | |
| Thrombocytopenia | [redacted] | Beta | [redacted] | [redacted] | |
| AE incidence – BR | | | | | |
| Anaemia | 0.00% | Beta | 0.00% | 0.00% | |
| Hypertension | 0.00% | Beta | 0.00% | 0.00% | |
| Neutropenia | 35.21% | Beta | 24.54% | 46.68% | |
| Pneumonia | 5.63% | Beta | 1.56% | 12.06% | |
| Thrombocytopenia | 0.00% | Beta | 0.00% | 0.00% | |
| AE incidence – zanubrutinib (match DRC) | | | | | |
| Anaemia | [redacted] | Beta | [redacted] | [redacted] | |
| Hypertension | [redacted] | Beta | [redacted] | [redacted] | |
| Neutropenia | [redacted] | Beta | [redacted] | [redacted] | |
| Pneumonia | [redacted] | Beta | [redacted] | [redacted] | |
| Thrombocytopenia | [redacted] | Beta | [redacted] | [redacted] | |
| AE incidence – DRC | | | | | |
| Anaemia | 0.00% | Beta | 0.00% | 0.00% | |
| Hypertension | 0.00% | Beta | 0.00% | 0.00% | |
| Neutropenia | 10.00% | Beta | 4.21% | 17.91% | |
| Pneumonia | 0.00% | Beta | 0.00% | 0.00% | |
| Thrombocytopenia | 0.00% | Beta | 0.00% | 0.00% | |
| AE duration, days | | | | | |
| Anaemia | 17.00 | Gamma | 10.5 | 25.0 | |
| Hypertension | 20.90 | Gamma | 11.4 | 33.2 | |
| Neutropenia | 10.90 | Gamma | 8.0 | 14.3 | |
| Pneumonia | 21.30 | Gamma | 9.5 | 37.8 | |
| Thrombocytopenia | 28.80 | Gamma | 10.0 | 57.4 | |

| Variable | Deterministic value (base-case analysis) | Distribution in DSA and PSA | Lower bound for DSA | Upper bound for DSA | Reference to section in submission |
|---|---|---|---|---|---|
| AE costs, £ | | | | | |
| Infections (including pneumonia and sepsis) | 1,481.76 | Gamma | 958.91 | 2116.55 | B.3.5.3 |
| AEs other than infections | 179.94 | Gamma | 116.45 | 257.03 | |
| **Mortality** | | | | | |
| Background mortality | Age- and sex-specific estimates | Fixed | | | B.3.3.4 |
| **Treatment costs, £** | | | | | |
| Drug acquisition costs | | | | | |
| Zanubrutinib, per 120 80mg capsules | 4,928.65 | Fixed | | | B.3.5.1 |
| Ibrutinib, per 28 420mg tablets | 4,292.40 | Fixed | | | |
| Bendamustine, per 100mg | 262.02 | Fixed | | | |
| Rituximab, per 500mg | 873.15 | Fixed | | | |
| Dexamethasone, per 10 4mg vials | 19.99 | Fixed | | | |
| Cyclophosphamide, per 100 50mg tablets | 139 | Fixed | | | |
| Dose intensity | | | | | |
| Zanubrutinib | 97.64% | Beta | 97.64% | 97.64% | B.3.2.3**Error! Reference source not found.** |
| Ibrutinib | 98.18% | Beta | 98.18% | 98.18% | |
| BR | 100.00% | Beta | 100.00% | 100.00% | |
| DRC | 100.00% | Beta | 100.00% | 100.00% | |
| Administration cost | | | | | |
| IV administration | 336.14 | Gamma | 217.53 | 480.14 | B.3.5.1 |
| **Healthcare resource use** | | | | | |
| Healthcare resource use per year | | | | | |
| Year 1-2: Full blood count, immunoglobulin, chemistry, haematologist, plasma viscosity, paraprotein | 5 | Gamma | 3.2 | 7.1 | B.3.5.2 |
| Year 3-5: Full blood count, immunoglobulin, Chemistry, Haematologist, Plasma viscosity, Paraprotein | 4 | Gamma | 2.6 | 5.7 | |
| Year 6+: Full blood count, Immunoglobulin, Chemistry, Haematologist, Plasma viscosity, Paraprotein | 3 | Gamma | 1.9 | 4.3 | |

| Variable | Deterministic value (base-case analysis) | Distribution in DSA and PSA | Lower bound for DSA | Upper bound for DSA | Reference to section in submission |
|---|---|---|---|---|---|
| Healthcare resource use cost, £ | | | | | |
| Full blood count | 2.87 | Gamma | 1.86 | 4.11 | B.3.5.2 |
| Immunoglobulin | 6.72 | Gamma | 4.35 | 9.60 | |
| Chemistry | 1.14 | Gamma | 0.73 | 1.62 | |
| Haematologist | 135.59 | Gamma | 87.75 | 193.68 | |
| Plasma viscosity | 6.75 | Gamma | 4.37 | 9.65 | |
| Paraprotein | 1.13 | Gamma | 0.73 | 1.61 | |
| **Subsequent treatment** | | | | | |
| Subsequent treatment use | 86% | Beta | 38% | 100% | B.3.5.4 |
| Subsequent treatment distribution following zanubrutinib or ibrutinib | | | | | |
| Ibrutinib | 0.0% | Fixed | | | |
| BR | 60.4% | Fixed | | | |
| DRC | 39.6% | Fixed | | | |
| Subsequent treatment distribution following BR | | | | | |
| Ibrutinib | 72.0% | Fixed | | | |
| BR | 0.0% | Fixed | | | |
| DRC | 28.0% | Fixed | | | |
| Subsequent treatment distribution following DRC | | | | | |
| Ibrutinib | 75.0% | Fixed | | | |
| BR | 25.0% | Fixed | | | |
| DRC | 0.0% | Fixed | | | |
| HR of time to initiating subsequent BTK inhibitors relative to PFS curve | 1.0 | Fixed | | | |
| HR of time to ending subsequent BTK inhibitors relative to OS curve | 1.0 | Fixed | | | |
| **Terminal care costs, £** | | | | | |
| Terminal care cost per mortality event | 7,978.35 | Gamma | 5163.17 | 11396.31 | B.3.5.4 |
| **Utilities** | | | | | |
| Health state utilities | | | | | |
| Pre-progression survival | 0.791 | Beta | 0.781 | 0.801 | B.3.4.5 |
| Post-progression survival | 0.691 | Beta | 0.681 | 0.701 | |
| AE disutility | | | | | |
| Anaemia | 0.088 | Beta | 0.057 | 0.125 | B.3.4.4 |
| Hypertension | 0.195 | Beta | 0.125 | 0.277 | |
| Neutropenia | 0.185 | Beta | 0.118 | 0.263 | |
| Pneumonia | 0.195 | Beta | 0.125 | 0.277 | |
| Thrombocytopenia | 0.123 | Beta | 0.079 | 0.175 | |

Abbreviations: AE = adverse event; BR = bendamustine and rituximab; DRC = dexamethasone, rituximab and cyclophosphamide; DSA = deterministic sensitivity analysis; HR = hazard ratio; ITT = intention-to-treat; OS = overall survival; PFS = progression-free survival; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year; TTD = time to treatment discontinuation

### B.3.6.2 Assumptions

Table B.3.31 presents the key assumptions of the base-case analysis.

**Table B.3.31. Key model assumptions**

| Category | Assumptions | Justification |
|---|---|---|
| Time horizon | 30 years | This is in line with the mean baseline age of patients enrolled in ASPEN (i.e., 69.5 years)[46] and the lifetime horizon assumption (i.e., 30 years) adopted in the previous model supporting the ibrutinib WM NICE appraisal (TA491).[3] |
| Comparison of zanubrutinib versus ibrutinib | The hazards were proportional over time between the zanubrutinib and the ibrutinib arm for each of the outcome separately, including OS, PFS and TTD. | The proportional hazard assumption was assessed through log-cumulative hazard plots which showed relatively straight and parallel curves. |
| Comparison of zanubrutinib versus DRC | It was assumed that the clinical outcomes were the same between relapsed/refractory patients treated with DRC and treatment-naïve patients treated with DRC. | There was a paucity of available clinical evidence (see Appendix D).<br>This was a conservative assumption, considering that the patients treated with zanubrutinib included a mix of relapsed/refractory patients and treatment-naïve patients unsuitable for chemo-immunotherapy. |
| Comparison of zanubrutinib versus BR | It was assumed that the clinical outcomes were similar between relapsed/refractory patients treated with BR and treatment-naïve patients treated with BR. | Although a MAIC was conducted to match the relapsed/refractory subpopulation (N=83) in the zanubrutinib arm to the BR population (see Section B.2.9), the associated MAIC results were not used to inform the cost-effectiveness analysis. This was because it would have negligible impact on model results but introduce more uncertainties.<br>More specifically, the KM curves of both PFS and OS overlapped between the ITT and relapsed/refractory populations of the zanubrutinib arm, likely because the relapsed/refractory population accounts for the majority of the ITT population. Therefore, it was expected that the deterministic results would be highly similar.<br><br>Given that the survival data for zanubrutinib are immature and hence associated with considerable uncertainty, restricting the analyses to the relapsed/refractory population would further reduce the effective sample size and statistical power and would then lead to even more uncertainty in the cost-effectiveness analysis that relied on the MAIC results for survival extrapolations. |
| Matching adjustment | All the unobserved prognostic factors or effect modifiers were balanced between the zanubrutinib population and the comparator populations. | The MAIC could only match zanubrutinib patient to match comparator populations on baseline characteristics reported in the comparator trial publications. It is not feasible to account for any variables that were not reported or observed, in the absence of IPD for the comparator trials. |
| Utility post progression | It was assumed that the utility decrement value due to disease progression in WM population would be the same as the utility decrement (0.10) as applied in previous UK NICE TA502 (0.10)[66] for ibrutinib in MCL and TA429 (0.098)[65] for ibrutinib in CLL. | There were no published HRQoL studies that reported health utilities in patients with WM. The utility for pre-progression survival health state (0.791) was estimated through a utility analysis using the EQ-5D-5L data collected in ASPEN. However, due to the very limited number of observations from patients who progressed, it was not feasible to derive a utility value for the post-progression survival health state from |

| | | |
|---|---|---|
| | | ASPEN trial data. Therefore, this analysis relied on prior NICE appraisals for ibrutinib in non-WM lymphoma indications to inform the utility decrement due to disease progression. |
| Subsequent treatment | The PFS curves were applied as proxy for time to next anticancer treatment. | Although subsequent anticancer treatment was collected in ASPEN for BTK inhibitors, the data was immature. In addition, time to next anticancer treatment was not available for BR and DRC. |
| | For patients receiving BTK inhibitors as subsequent treatment, the OS curves were applied as proxy for time to ending subsequent treatment with BTK inhibitors. That is, for patients receiving subsequent BTK inhibitors, it was assumed that patients would be on ibrutinib until death, based on the time patients spent in the post-progression survival health state. | There was a lack of data to inform when patients would discontinue subsequent BTK inhibitors, and therefore, an assumption had to be made. |
| | For patients treated with subsequent chemo-immunotherapy (i.e., BR or DRC), 2 lines of subsequent treatments (i.e., 2 treatment courses, each with 6 cycles) were assumed. | This assumption was based on<br>(1) the number of previous lines of treatment in ASPEN (with a median of one) and<br>(2) the number of treatment lines modelled in previous ibrutinib WM models in the UK and Italy (i.e., second- to fourth-line, and best supportive care afterwards).[3, 69] |

Abbreviations: BR = bendamustine and rituximab; DRC = dexamethasone, rituximab and cyclophosphamide; EQ-5D-5L = EuroQol 5-Dimensions 5-Level; IPD = individual patient-level data; ITT, intention-to-treat, NICE = National Institute for Health and Care Excellence; OS = overall survival; PFS = progression-free survival; QALY = quality-adjusted life year; TTD = time to treatment discontinuation; UK = United Kingdom; WM = Waldenström's macroglobulinaemia

## B.3.7 Base-case results

### B.3.7.1 Base-case incremental cost-effectiveness analysis results

Table B.3.32 presents the base-case analysis results for the 3 pairwise comparisons of zanubrutinib versus ibrutinib, zanubrutinib versus BR, and zanubrutinib versus DRC, separately, at the proposed list price. Zanubrutinib was associated with greater QALYs and higher costs when compared with ibrutinib, BR and DRC separately, leading to incremental cost-effectiveness ratios (ICERs) of £[redacted], and [redacted] per QALY, respectively. Disaggregated results are presented in Appendix J.

**Table B.3.32. Base-case pairwise results**

| Technology | Total costs,£ | Total LYG | Total QALYs | Incremental costs, £ | Incremental LYG | Incremental QALYs | ICER, £/QALY |
|---|---|---|---|---|---|---|---|
| Pairwise comparison 1 | | | | | | | |
| Zanubrutinib | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | - |
| Ibrutinib | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Pairwise comparison 2 | | | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | - |
| BR | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Pairwise comparison 3 | | | | | | | |

| Technology | Total costs,£ | Total LYG | Total QALYs | Incremental costs, £ | Incremental LYG | Incremental QALYs | ICER, £/QALY |
|---|---|---|---|---|---|---|---|
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | - |
| DRC | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |

Abbreviations: BR = bendamustine and rituximab; DRC = dexamethasone, rituximab and cyclophosphamide; ICER = incremental cost-effectiveness ratio; LY = life years; QALY = quality-adjusted life years

## B.3.8 Sensitivity analyses

### B.3.8.1 Probabilistic sensitivity analysis

The PSA sampled from the distribution of each model parameter for a total of 1,000 simulations, with summary results and comparison with base-case results presented in Table B.3.33. The results for total costs and total QALYs from the probabilistic analysis were similar to those of the deterministic base-case analysis, indicating the model is structurally stable.

**Table B.3.33. Mean results of PSA (1,000 runs) and comparison with base-case results**

| | Total costs, £ | | Total QALYs | | ICER, £/QALY | |
|---|---|---|---|---|---|---|
| | Base case | PSA | Base case | PSA | Base case | PSA |
| Zanubrutinib | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Ibrutinib | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |

Abbreviations: BR = bendamustine and rituximab; CI = confidence interval; DRC = dexamethasone, rituximab and cyclophosphamide; ICER = incremental cost-effectiveness ratio; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year

The results of each probabilistic model run are presented on the cost-effectiveness plane for each pairwise comparison (Figure B.3.36 to Figure B.3.41). The spread of the points horizontally illustrates the uncertainty in results of total QALYs and the spread of the points vertically demonstrates the uncertainty in the cost results.

**Figure B.3.36. Scatter plot of PSA results on total costs and QALYs – zanubrutinib vs ibrutinib**

**Figure B.3.37. Scatter plot of PSA results on incremental costs and QALYs – zanubrutinib vs ibrutinib**

**Figure B.3.38. Scatter plot of PSA results on total costs and QALYs – zanubrutinib (match BR) vs BR**

**Figure B.3.39. Scatter plot of PSA results on incremental costs and QALYs – zanubrutinib (match BR) vs BR**

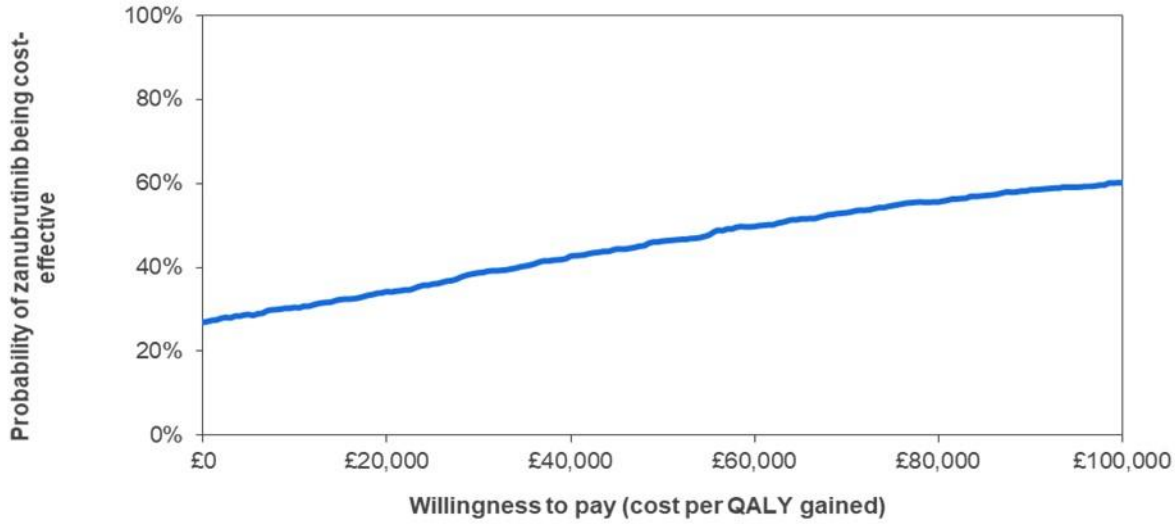
**Figure B.3.40. Scatter plot of PSA results on total costs and QALYs – zanubrutinib (match DRC) vs DRC**

**Figure B.3.41. Scatter plot of PSA results on incremental costs and QALYs – zanubrutinib (match DRC) vs DRC**

The uncertainty associated with each treatment in terms of probability of zanubrutinib being cost effective is presented over the range of willingness-to-pay (WTP) values in the form of a cost-effectiveness acceptability curve (Figure B.3.42 to Figure B.3.44).

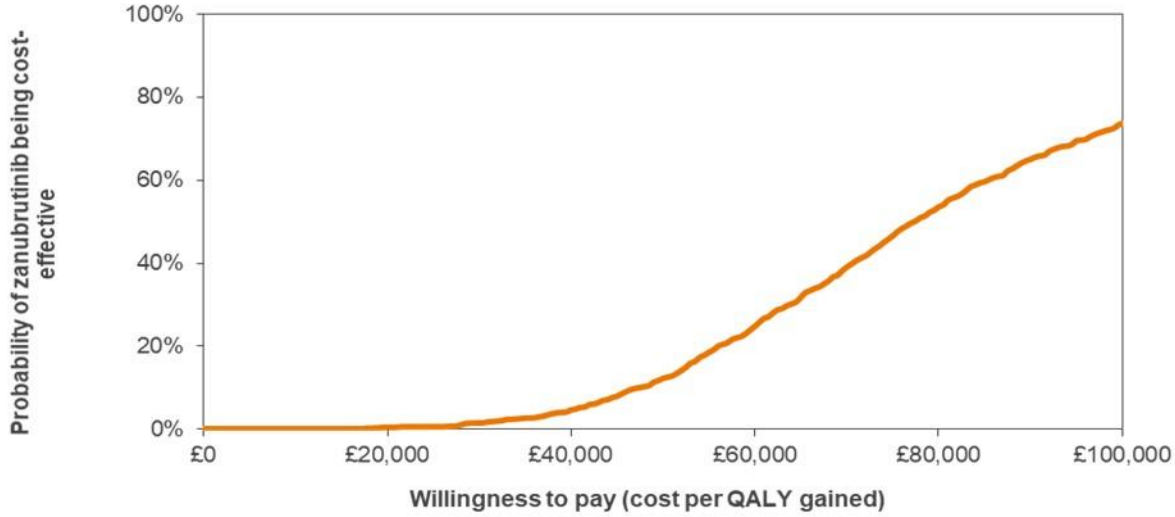
At a £30,000/QALY WTP threshold, the probabilities of zanubrutinib being cost-effective were estimated to be 39%, 1% and 0%, compared to ibrutinib, BR and DRC, respectively, whereas at a £50,000/QALY WTP threshold, the probabilities of zanubrutinib being cost-effective were estimated to be 46%, 12% and 1%, compared to ibrutinib, BR and DRC, respectively.

**Figure B.3.42. Cost-effectiveness acceptability curve – zanubrutinib vs ibrutinib**



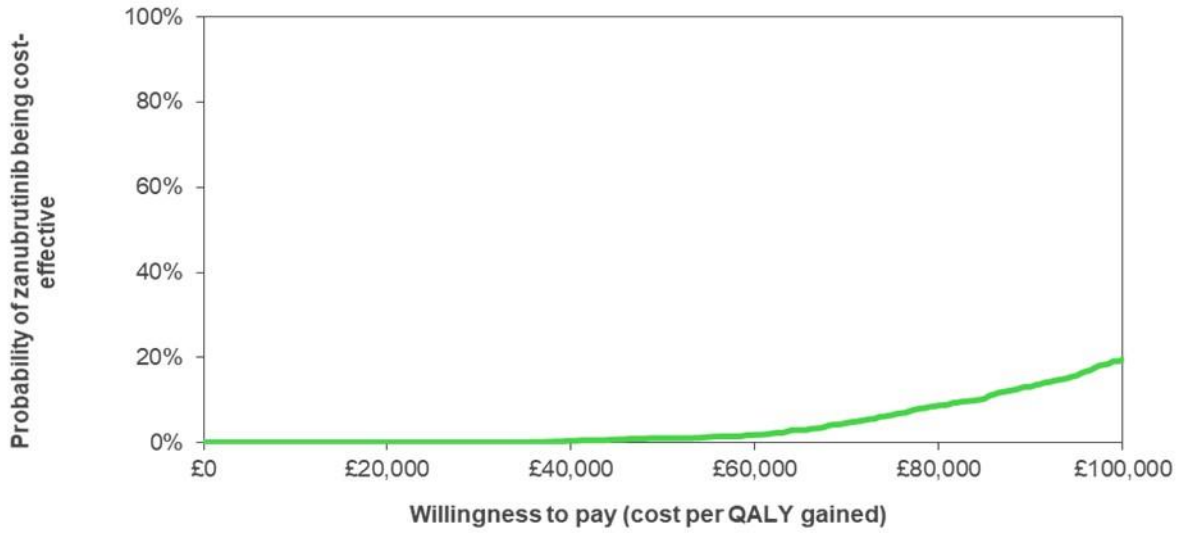

Abbreviation: QALY = quality-adjusted life year; vs = versus

**Figure B.3.43. Cost-effectiveness acceptability curve – zanubrutinib (match BR) vs BR**




Abbreviations: BR = bendamustine and rituximab; QALY = quality-adjusted life year; vs = versus

**Figure B.3.44. Cost-effectiveness acceptability curve – zanubrutinib (match DRC) vs DRC**




Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; QALY = quality-adjusted life year; vs = versus

### B.3.8.2 Deterministic sensitivity analysis

The results of the DSA are presented using tornado plots that show how parameter uncertainty would impact key model results. The top-10 parameters that created the widest range in each model result are displayed in Figure B.3.45 to Figure B.3.53.

As shown in Figure B.3.47, Figure B.3.50 and Figure B.3.53, the top drivers of ICERs included the proportion of patients treated with subsequent treatment and average baseline age (note: the uncertainties of survival parameters were not examined in the DSA but in the PSA through Cholesky decomposition).

**Figure B.3.45. Tornado plot of parameter impact on incremental costs – zanubrutinib vs ibrutinib**

[redacted]

**Figure B.3.46. Tornado plot of parameter impact on incremental QALYs – zanubrutinib vs ibrutinib**

**Figure B.3.47. Tornado plot of parameter impact on incremental costs per QALY gained – zanubrutinib vs ibrutinib**

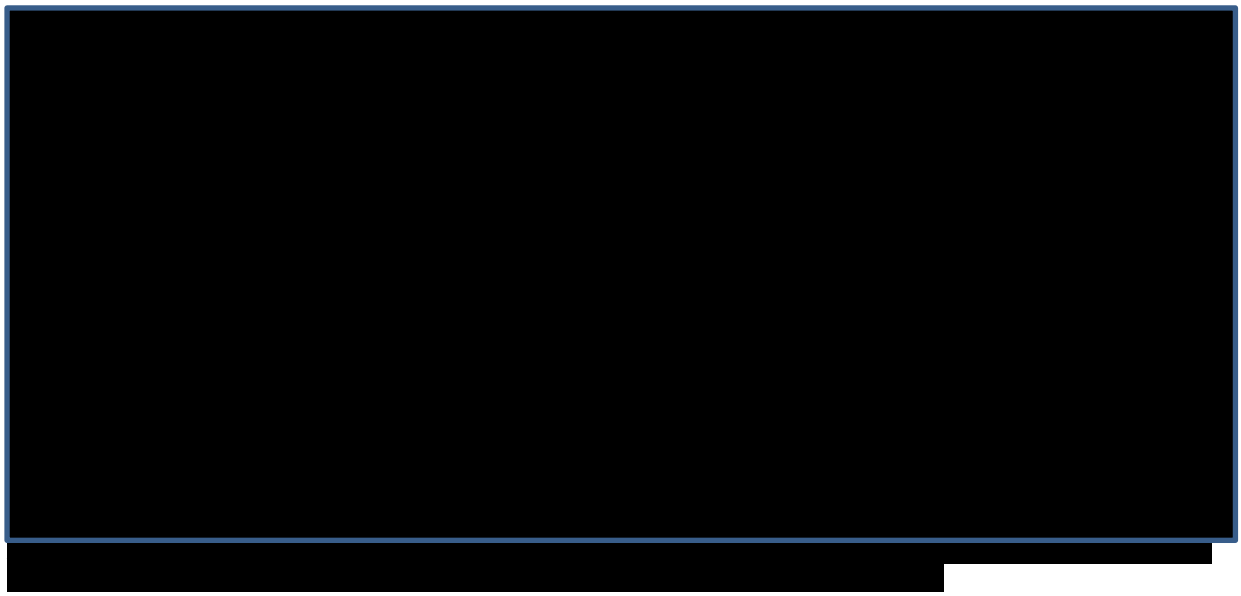


**Figure B.3.48. Tornado plot of parameter impact on incremental costs – zanubrutinib (match BR) vs BR**

**Figure B.3.49. Tornado plot of parameter impact on incremental QALYs – zanubrutinib (match BR) vs BR**

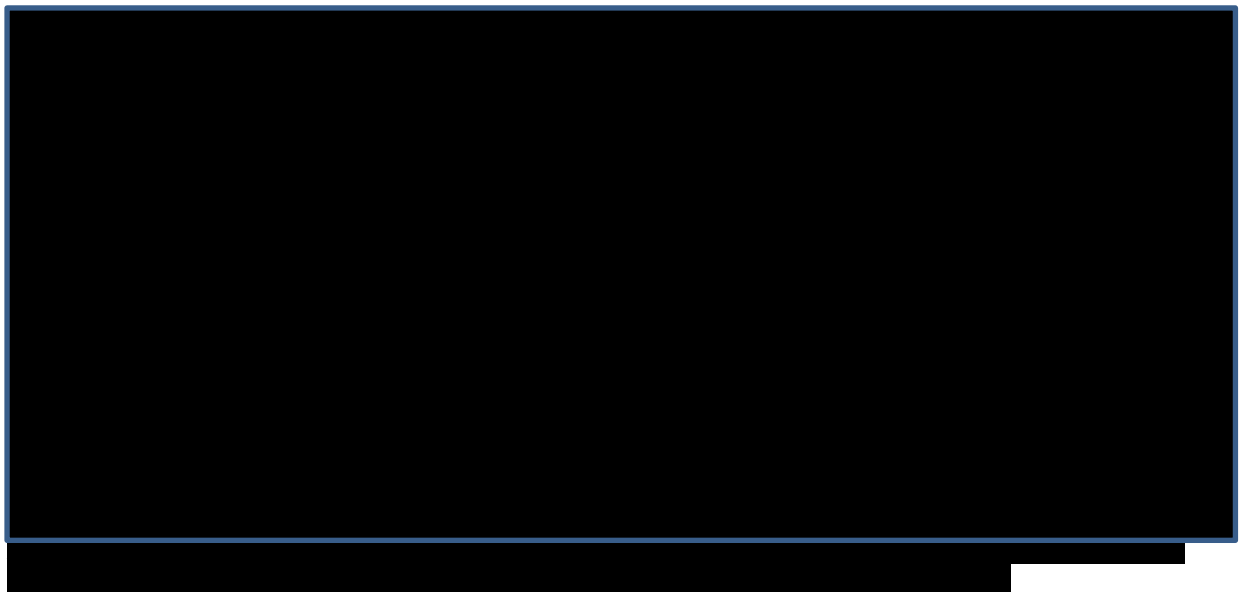


**Figure B.3.50. Tornado plot of parameter impact on incremental costs per QALY gained – zanubrutinib (match BR) vs BR**

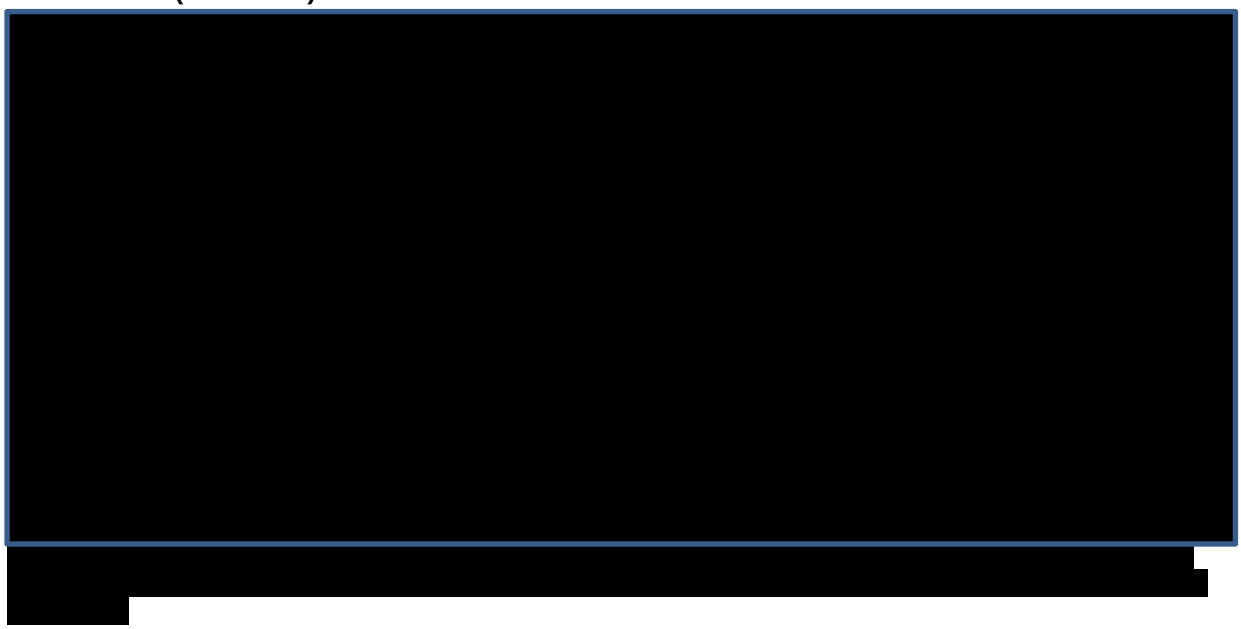


**Figure B.3.51. Tornado plot of parameter impact on incremental costs – zanubrutinib (match DRC) vs DRC**

**Figure B.3.52. Tornado plot of parameter impact on incremental QALYs – zanubrutinib (match DRC) vs DRC**

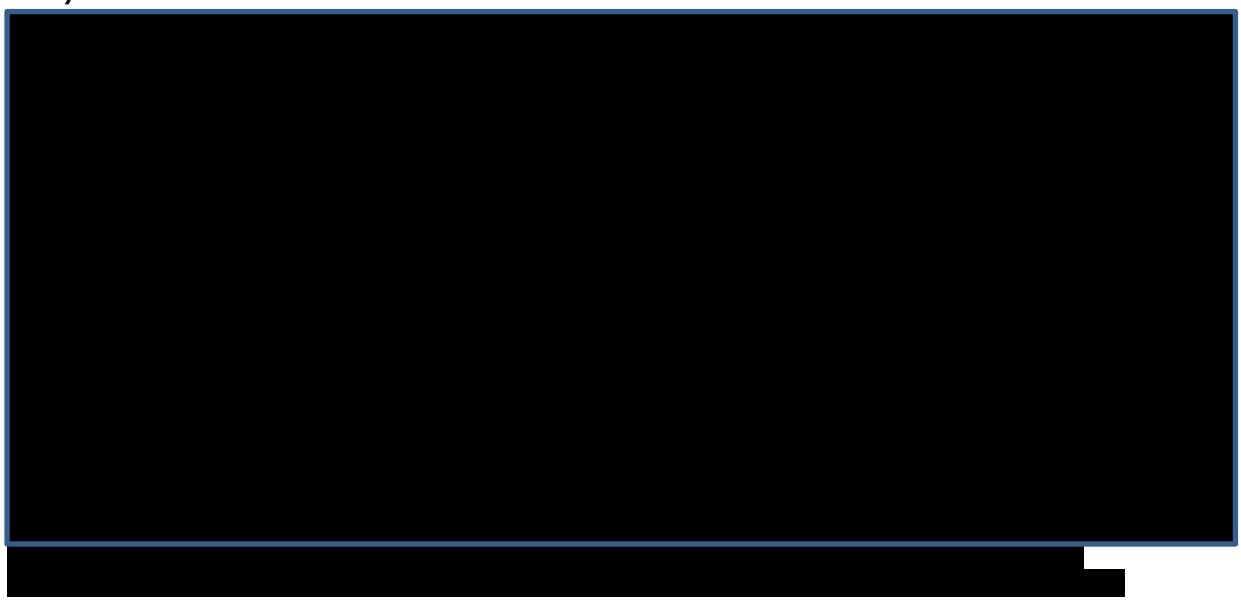


**Figure B.3.53. Tornado plot of parameter impact on incremental costs per QALY gained – zanubrutinib (match DRC) vs DRC**

### B.3.8.3 Scenario analysis

Table B.3.34 to Table B.3.36 present the results of scenario analyses conducted to assess the impact of uncertainty of model assumptions and inputs on the model results, especially those not varied in the DSA or PSA.

In the comparison between zanubrutinib and ibrutinib, the ICER fell between £50,000 and £70,000 per QALY gained in most of the tested scenarios. The ICER fluctuated most from the base-case setting when the hazard ratio was applied to ibrutinib after 30 months in the modeling approach for OS/PFS/TTD of zanubrutinib. In addition, adopting a 10-year time horizon had a relatively high impact on the ICER.

In the comparison between zanubrutinib (matching BR) and BR, the ICER fell between £60,000 and £90,000 per QALY gained in most of the tested scenarios. The ICER fluctuated most from the base-case setting when the time horizon was set to 10 years. In addition, subsequent treatment inclusion and settings, as well as the discounting rate being set at 1.5% for cost and 0% for QALY had a relatively high impact on the ICER.

In the comparison between zanubrutinib (matching DRC) and DRC, the ICER fell between £100,000 and £140,000 per QALY gained in most of the tested scenarios. The ICER fluctuated most from the base-case setting when the time horizon was set to 10 years. Additionally, subsequent treatment inclusion and settings had a relatively high impact on the ICER.

**Table B.3.34. Scenario analyses results – zanubrutinib vs ibrutinib**

| Parameter | Base-case value or setting | Alternative value or setting | Incremental costs, £ | Incremental QALYs | ICER, £/QALY |
|---|---|---|---|---|---|
| Base case | | | [redacted] | [redacted] | [redacted] |
| Discounting rate per year for costs and QALYs | 3.5% for both costs and QALYs | Costs and QALYs: 0% | [redacted] | [redacted] | [redacted] |
| | | Costs and QALYs: 1.5% | [redacted] | [redacted] | [redacted] |
| | | Costs: 0%; QALYs 1.5% | [redacted] | [redacted] | [redacted] |
| | | Costs: 1.5%; QALYs 0% | [redacted] | [redacted] | [redacted] |
| Time horizon | 30 years | 10 years | [redacted] | [redacted] | [redacted] |
| | | 20 years | [redacted] | [redacted] | [redacted] |
| Modeling approach for OS/PFS/TTD of zanubrutinib | ASPEN data-based parametric models for OS/PFS/TTD | ASPEN data-based parametric models for OS/PFS/TTD in the first 30 months; after 30 months, applying HR=1.00 to ibrutinib OS/PFS/TTD | [redacted] | [redacted] | [redacted] |
| Modeling approach for OS of zanubrutinib | ASPEN data-based parametric model for OS | ASPEN data-based parametric model for OS in the first 30 months; after 30 months, applying time-varying HRs of PFS-OS of DRC to zanubrutinib PFS | [redacted] | [redacted] | [redacted] |
| Health state utility for post-progression survival | 0.691 | 0.650 | [redacted] | [redacted] | [redacted] |
| | | 0.600 | [redacted] | [redacted] | [redacted] |
| Inclusion of disutility of adverse events | Yes | No | [redacted] | [redacted] | [redacted] |
| Proportion of patients receiving subsequent treatment upon progression on primary treatment | 86% | 100% | [redacted] | [redacted] | [redacted] |
| | | 70% | [redacted] | [redacted] | [redacted] |
| Inclusion of subsequent treatment cost | Yes | No | [redacted] | [redacted] | [redacted] |

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; HR = hazard ratio; ICER = incremental cost-effectiveness ratio; OS = overall survival; PFS = progression-free survival; QALY = quality-adjusted life year; TTD = time to treatment discontinuation; vs = versus

**Table B.3.35. Scenario analyses results – zanubrutinib (match BR) vs BR**

| Parameter | Base-case value or setting | Alternative value or setting | Incremental costs, £ | Incremental QALYs | ICER, £/QALY |
|---|---|---|---|---|---|
| Base case | | | [redacted] | [redacted] | [redacted] |
| Discounting rate per year for costs and QALYs | 3.5% for both costs and QALYs | Costs and QALYs: 0% | [redacted] | [redacted] | [redacted] |
| | | Costs and QALYs: 1.5% | [redacted] | [redacted] | [redacted] |
| | | Costs: 0%; QALYs: 1.5% | [redacted] | [redacted] | [redacted] |
| | | Costs: 1.5%; QALYs: 0% | [redacted] | [redacted] | [redacted] |
| Time horizon | 30 years | 10 years | [redacted] | [redacted] | [redacted] |
| | | 20 years | [redacted] | [redacted] | [redacted] |
| Baseline patient characteristics | Baseline characteristics of ASPEN ITT population (N=201) | Baseline characteristics of ASPEN ITT population, zanubrutinib arm (match BR; $n_{eff}$=[redacted]) | [redacted] | [redacted] | [redacted] |
| Parametric model type for OS of zanubrutinib (match BR) vs BR | Independent exponential model for zanubrutinib (match BR); independent Weibull model for BR | Dependent Weibull model | [redacted] | [redacted] | [redacted] |
| | | Dependent gamma model | [redacted] | [redacted] | [redacted] |
| Health state utility for post-progression survival | 0.691 | 0.650 | [redacted] | [redacted] | [redacted] |
| | | 0.600 | [redacted] | [redacted] | [redacted] |
| Inclusion of disutility of adverse events | Yes | No | [redacted] | [redacted] | [redacted] |
| Proportion of patients receiving subsequent treatment upon progression on primary treatment | 86% | 100% | [redacted] | [redacted] | [redacted] |
| | | 70% | [redacted] | [redacted] | [redacted] |
| Number of subsequent lines of treatment for patients receiving chemo-immunotherapy as subsequent treatment | 2 | 3 | [redacted] | [redacted] | [redacted] |
| | | 4 | [redacted] | [redacted] | [redacted] |
| Inclusion of subsequent treatment cost | Yes | No | [redacted] | [redacted] | [redacted] |

| Parameter | Base-case value or setting | Alternative value or setting | Incremental costs, £ | Incremental QALYs | ICER, £/QALY |
|---|---|---|---|---|---|
| Approach of including subsequent treatment | Cost only by relying on post-progression survival duration for subsequent BTK inhibitors and fixed number of treatment course for subsequent chemo-immunotherapy | Both costs and effects by relying on lump-sum incremental costs and incremental QALYs from previous technology appraisal for ibrutinib in WM | [redacted] | [redacted] | [redacted] |

Abbreviations: BR = bendamustine and rituximab; DRC = dexamethasone, rituximab and cyclophosphamide; ICER = incremental cost-effectiveness ratio; ITT = intention-to-treat, N = number of patients evaluable; $n^{eff}$ = effective sample size; OS = overall survival; PFS = progression-free survival; QALY = quality-adjusted life year; vs = versus; WM = Waldenström's macroglobulinaemia

**Table B.3.36. Scenario analyses results – zanubrutinib (match DRC) vs DRC**

| Parameter | Base-case value or setting | Alternative value or setting | Incremental costs, £ | Incremental QALYs | ICER, £/QALY |
|---|---|---|---|---|---|
| Base case | | | [redacted] | [redacted] | [redacted] |
| Discounting rate per year for costs and QALYs | 3.5% for both costs and QALYs | Costs and QALYs: 0% | [redacted] | [redacted] | [redacted] |
| | | Costs and QALYs: 1.5% | [redacted] | [redacted] | [redacted] |
| | | Costs: 0%; QALYs 1.5% | [redacted] | [redacted] | [redacted] |
| | | Costs: 1.5%; QALYs 0% | [redacted] | [redacted] | [redacted] |
| Time horizon | 30 years | 10 years | [redacted] | [redacted] | [redacted] |
| | | 20 years | [redacted] | [redacted] | [redacted] |
| Baseline patient characteristics | Baseline characteristics of ASPEN ITT population (N=201) | Baseline characteristics of ASPEN ITT population, zanubrutinib arm (match DRC; $n_{eff}$=[redacted]) | [redacted] | [redacted] | [redacted] |
| Parametric model type for OS of zanubrutinib (match DRC) and DRC | Dependent model with gamma distribution | Dependent Gompertz model | [redacted] | [redacted] | [redacted] |
| | | Dependent Weibull model | [redacted] | [redacted] | [redacted] |
| Health state utility for post-progression survival | 0.691 | 0.650 | [redacted] | [redacted] | [redacted] |
| | | 0.600 | [redacted] | [redacted] | [redacted] |
| Inclusion of disutility of adverse events | Yes | No | [redacted] | [redacted] | [redacted] |
| Proportion of patients receiving subsequent treatment upon progression on primary treatment | 86% | 100% | [redacted] | [redacted] | [redacted] |
| | | 70% | [redacted] | [redacted] | [redacted] |

| Parameter | Base-case value or setting | Alternative value or setting | Incremental costs, £ | Incremental QALYs | ICER, £/QALY |
|---|---|---|---|---|---|
| Number of subsequent lines of treatment for patients receiving chemo-immunotherapy as subsequent treatment | 2 | 3 | [redacted] | [redacted] | [redacted] |
| | | 4 | [redacted] | [redacted] | [redacted] |
| Inclusion of subsequent treatment cost | Yes | No | [redacted] | [redacted] | [redacted] |
| Approach of including subsequent treatment | Cost only by relying on post-progression survival duration for subsequent BTK inhibitors and fixed number of treatment course for subsequent chemo-immunotherapy | Both costs and effects by relying on lump-sum incremental costs and incremental QALYs from previous technology appraisal for ibrutinib in WM | [redacted] | [redacted] | [redacted] |

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; ICER = incremental cost-effectiveness ratio; ITT = intention-to-treat; HR = hazard ratio; N = number of patients evaluable; $n^{eff}$ = effective sample size; OS = overall survival; PFS = progression-free survival; QALY = quality-adjusted life year; TTD = time to treatment discontinuation; vs = versus; WM = Waldenström's macroglobulinaemia

### B.3.8.4 Summary of sensitivity analyses results

The mean probabilistic results are aligned with the deterministic results for all three treatment comparisons, indicating the model is structurally stable.

For the comparison with ibrutinib, the probabilities of zanubrutinib being cost-effective was 39% and 46% at WTP thresholds of £30,000 and £50,000 per QALY gained, respectively. The ICER fell between £50,000 and £70,000 per QALY gained in most of the tested scenario analyses. Scenario analyses and DSA indicated that the ICER was most sensitive to the modeling approach for OS/PFS/TTD of zanubrutinib, time horizon and the proportion of patients receiving subsequent treatments.

For the comparison with BR, the probabilities of zanubrutinib being cost-effective were 1% and 12% at WTP thresholds of £30,000 and £50,000 per QALY gained, respectively. The ICER was £60,000 and £90,000 per QALY gained in most of the tested scenario analyses. The scenario analyses and the DSA showed that the ICER was most sensitive to the time horizon, discounting rate, and the inclusion and proportion of patients receiving subsequent treatments.

For the comparison with DRC, the probabilities of zanubrutinib being cost-effective were 0% and 1% at WTP thresholds of £30,000 and £50,000 per QALY gained, respectively. The ICERs fell between £100,000 and £140,000 in most of the tested scenario analyses. Scenario analyses and the DSA showed that the ICER was most sensitive to the time horizon and the inclusion and proportion of patients receiving subsequent treatments.

## B.3.9 Subgroup analysis

No subgroup was modelled for this economic evaluation.

## B.3.10 Validation

### B.3.10.1 Validation of cost-effectiveness analysis

#### *B.3.10.1.1 Internal validation*

A range of items have been tested to manage quality control of the model (Table B.3.37), all of which yielded positive results.

**Table B.3.37. Quality control items (selected)**

| Category | Item |
|---|---|
| Logical tests | Set all utility values equal to 1 and set all disutilities to zero. QALYs should equal LYs. |
| | Set all utility/disutility values to zero. There should be zero QALYs accrued for all included treatments. |
| | Set all mortality rates (including background mortality) to 1. All patients should be dead in cycle 1, but still produce (some) expected costs and QALYs (due to half cycle, and one-off costs/disutilities) |
| | Set mortality rate (including background mortality) at age X to 1. All patients should survive until age X, and still produce expected costs and QALYs |
| | If included, set all AE probabilities to zero. Make sure that no AEs occur, and that AE-related costs and disutilities are also estimated to be zero. |
| | Set unit costs for all included treatments to zero. Estimated treatment costs should be zero. |
| | Halve and double treatment unit costs for each treatment. Estimated undiscounted treatment costs should also halve and double in response. |
| | For other included cost categories: halve, double, and set to zero. Ensure that all undiscounted model results respond as expected. |
| | If included as an input: increase and decrease treatment durations. Do treatment costs increase and decrease appropriately in response? |
| | Explore higher and lower time horizons. LYs, QALYs and total costs should increase/decrease with longer/shorter time horizons. Set time horizon to zero – all costs and outcomes should be zero/undefined |
| | Set the discount rate of benefits to 100%. Total QALYs should dramatically decrease. Repeat for cost discount rate. |
| | Set the discount rate of benefits to 0%. Total discounted QALYs should increase and match undiscounted results exactly. Repeat for cost discount rate. |
| Technical implementation | Check that half cycle correction has been appropriately applied. |
| | Check that background mortality is correctly applied (for the correct age, adjusted for cycle length, reactive to changes in age and gender distribution). Pay special attention to the last model cycle, and any assumptions made for ages that fall outside of the life table (e.g., 100+) |
| | Has discounting been appropriately applied using the correct formula? ((1+p)^(-t)). And is it implemented separately for costs and benefits? (check cell references) |

| Category | Item |
|---|---|
| | Check that the sum of all health state membership in the model sums to 1 for all cycles. Check that this isn't simply the result of one state's membership being equal to 1 minus all others (this can hide errors) |
| | Check that probabilities and rates have been handled correctly (i.e., rates are converted to probabilities before being used as transitions). |
| | Check that the starting distribution of health states is correct, and consistent across the included treatments. Check that it makes sense given the decision problem (e.g., patients who have the disease, vs patients who have been diagnosed with disease) |
| | Ensure that the cumulative probability to die in any given cycle is equal to or greater than that of the age-matched general population mortality. Use the background mortality tables to check |
| | Confirm that the relationship with the cycle length and time horizon is correct. |
| | Following on from the above, ensure that treatment durations are correct by confirming that the model is incurring treatment costs for the correct number of cycles. Ensure any stopping rules are appropriately timed. |
| | Confirm that disutilities are correctly subtracted from QALYs, by ensuring that duration is taken into account in the calculations. |

Abbreviations: AE = adverse event; LY = life year; QALY = quality-adjusted life year

#### *B.3.10.1.2 External validation*

##### *B.3.10.1.2.1 Inclusion of comparators and clinical trials*

The comparators and associated clinical trials identified from the SLR were finalised based on a combination of:

- Recommendation by WM treatment guidelines per ESMO[12], International Workshop on Waldenström's Macroglobulinemia[71], and Australia Medical and Scientific Advisory Group[72]
- Real-world treatment patterns according the UK Rory Morrison Registry, which showed that ibrutinib, BR and DRC are the most commonly used regimens in patients with relapsed or refractory diseases[1]
- A medical advisory board meeting in the EU.[73]

##### *B.3.10.1.2.2 Validation of long-term survival extrapolation*

As described in Section B.3.3.2, PFS, OS, and TTD, the selection of parametric models for survival was based on both internal validity, assessed by AIC and BIC fit statistics and visual inspection, and external validity, assessed by published estimates and clinical experts' opinions on clinical plausibility of the extrapolated survival.[74]

## B.3.11 Interpretation and conclusions of economic evidence

This model was developed to assess the cost effectiveness of zanubrutinib for the treatment of adult patients with WM who have received at least one prior line of therapy, or as first-line treatment for patients unsuitable for chemo-immunotherapy, compared with ibrutinib, BR, and DRC, separately.

The comparison with ibrutinib was based on the head-to-head comparison from ASPEN. Results of the base-case analysis showed that zanubrutinib was associated with an additional discounted [redacted] additional discounted QALYs, and a decrease in discounted costs

of [redacted] compared with ibrutinib over a 30-year lifetime horizon. As such, treatment with zanubrutinib rather than ibrutinib cost an additional [redacted] gained.

The comparison with BR was based on the results of an MAIC. Results of the base-case analysis indicated that zanubrutinib was associated with an additional discounted [redacted] additional discounted QALYs, and increased discounted costs of [redacted] over a 30-year lifetime horizon. As such, treatment with zanubrutinib rather than BR cost an additional [redacted] gained.

The comparison with DRC was also based on an MAIC. This comparison showed that zanubrutinib was associated with an additional discounted [redacted] additional discounted QALYs, and increased discounted costs of [redacted] over a 30-year lifetime horizon. As such, treatment with zanubrutinib rather than DRC was estimated to cost an additional [redacted] gained. Despite the substantially higher ICER compared with the comparisons with ibrutinib and with BR, it is important to note that the data for DRC were obtained from treatment-naïve patients suitable for chemo-immunotherapy (see Appendix D for the discussions surrounding the availability of clinical evidence for DRC), which would likely bias the results against zanubrutinib. As such, the ICER for the comparison with DRC could potentially be overestimated.

Extensive sensitivity and scenario analyses were conducted to test the robustness of the base-case results. The results of most of the sensitivity and scenario analyses were aligned with the results of the base-case analysis for all three pairwise comparisons. The DSA indicated that the model outcomes were most sensitive to the proportion of patients receiving subsequent treatments for all three comparisons. For the comparisons with BR and DRC specifically, the ICERs were also sensitive to the time horizon.

The major limitation of the analysis lies with the immaturity of the survival data for BTK inhibitors (i.e., zanubrutinib, ibrutinib), which causes uncertainties of the results of long-term survival extrapolation. To mitigate this limitation, clinical experts were consulted as to the clinical plausibility of the extrapolated survival. [redacted]

## B.4. References


1. WMUK. First UK Waldenström's Macroglobulinaemia Registry Report. 2018. Available at: https://www.wmuk.org.uk/sites/default/files/2020-01/The%20Rory%20Morrison%20WM%20Registry%20Report%20%281%29%20-%20March%202018.pdf (accessed: August 2020).
2. Tam CS, et al. A Randomized Phase 3 Trial of Zanubrutinib Versus Ibrutinib in Symptomatic Waldenstrom Macroglobulinemia:The Aspen Study. Blood. 2020.
3. © National Institute for Health and Care Excellence. Single Technology Appraisal Ibrutinib for treating Waldenström's macroglobulinaemia [ID884]. 2016. Available at: https://www.nice.org.uk/guidance/ta491/documents/committee-papers (accessed: August 2020).
4. BeiGene. Zanubrutinib Summary of Product Characteristics. 2020.
5. Hendriks RW. Drug discovery: New Btk inhibitor holds promise. Nat Chem Biol. 2011;7(1):4-5.
6. Tam CS, et al. Three-year follow-up of treatment-naïve and previously treated patients with Waldenström macroglobulinemia (WM) receiving single-agent zanubrutinib. Journal of Clinical Oncology. 2020;38(15_suppl):8051-8051.
7. Tam CS, et al. A head-to-head Phase III study comparing zanubrutinib versus ibrutinib in patients with Waldenstrom macroglobulinemia. Future Oncol. 2018;14(22):2229-2237.
8. Kapoor P, et al. Diagnosis and Management of Waldenstrom Macroglobulinemia: Mayo Stratification of Macroglobulinemia and Risk-Adapted Therapy (mSMART) Guidelines 2016. JAMA Oncol. 2017;3(9):1257-1265.
9. Mazzucchelli M, et al. Waldenstrom's Macroglobulinemia: An Update. Mediterr J Hematol Infect Dis. 2018;10(1):e2018004.
10. Dimopoulos MA, Kastritis E. How I treat Waldenström macroglobulinemia. Blood. 2019;134(23):2022-2035.
11. Swerdlow SH, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. Blood. 2016;127(20):2375-2390.
12. Kastritis E, et al. Waldenstrom's macroglobulinaemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2019;30(5):860-862.
13. Paludo J, Ansell SM. Waldenstrom macroglobulinemia: biology, genetics, and therapy. Blood Lymphat Cancer. 2016;6:49-58.
14. Wang W, Lin P. Lymphoplasmacytic lymphoma and Waldenstrom macroglobulinaemia: clinicopathological features and differential diagnosis. Pathology. 2020;52(1):6-14.
15. Varettoni M, et al. Pattern of somatic mutations in patients with Waldenstrom macroglobulinemia or IgM monoclonal gammopathy of undetermined significance. Haematologica. 2017;102(12):2077-2085.
16. Treon SP, et al. MYD88 Mutations and Response to Ibrutinib in Waldenstrom's Macroglobulinemia. N Engl J Med. 2015;373(6):584-586.
17. Treon SP, et al. Genomic Landscape of Waldenstrom Macroglobulinemia and Its Impact on Treatment Strategies. J Clin Oncol. 2020;38(11):1198-1208.
18. Office for National Statistics. Population estimates for the UK, England and Wales, Scotland and Northern Ireland: mid-2019. 2020. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2019estimates (accessed: October 2020).
19. Owen RG, et al. Guidelines on the diagnosis and management of Waldenstrom macroglobulinaemia. Br J Haematol. 2014;165(3):316-333.
20. Office for National Statistics. Cancer Statistics Registrations, England, 2017 dataset. 2019. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancerregistrationstatisticscancerregistrationstatisticsengland (accessed: August 2020).
21. International Waldenstrom's Macroglobulinemia Foundation. About Waldenstrom's Macroglobilinemia Fact Sheet. 2018.
22. Owen RG, et al. Waldenstrom macroglobulinemia. Development of diagnostic criteria and identification of prognostic factors. Am J Clin Pathol. 2001;116(3):420-428.
23. Maqbool MG, et al. A practical guide to laboratory investigations at diagnosis and follow up in Waldenstrom macroglobulinaemia: recommendations from the Medical and Scientific Advisory Group, Myeloma Australia, the Pathology Sub-committee of the Lymphoma and Related Diseases Registry and the Australasian Association of Clinical Biochemists Monoclonal Gammopathy Working Group. Pathology. 2020;52(2):167-178.

24. Tomkins O, et al. An Analysis from the WM UK Rory Morrison Registry - How Does Waldenström's Macroglobulinemia Affect Younger Patients? Blood. 2019;134(Supplement_1):1543-1543.
25. Morel P, et al. International prognostic scoring system for Waldenstrom macroglobulinemia. Blood. 2009;113(18):4163-4170.
26. Castillo JJ, et al. Ibrutinib in Waldenstrom macroglobulinemia: latest evidence and clinical experience. Ther Adv Hematol. 2016;7(4):179-186.
27. American Cancer Society. If You Have Waldenstrom Macroglobulinemia. Available at: https://www.cancer.org/cancer/waldenstrom-macroglobulinemia/if-you-have-waldenstrom-macroglobulinemia.html (accessed: August 2020).
28. Frustaci AM, et al. Health-Related Quality of Life in Waldenstrom Macroglobulinemia (WM) and IgM Monoclonal Gammopathy of Undeterminated Significance (IgM-MGUS). Blood. 2018;132(Supplement 1):3593-3593.
29. Dren NM, et al. A Cross-Sectional Study Examining the Effects of Patient Information Level on Healthcare Experience in 2 Patient Populations: Extranodal Natural Killer T-Cell Lymphoma (ENKTL) and Waldenstrom Macroglobulinemia (WM). Blood. 2019;134(Supplement_1):3422-3422.
30. Lymphoma Coalition. Waldenstrom Macroglobulinemia; Subtype review. Available at: https://www.iwmf.com/sites/default/files/docs/publications/SubtypeReviewofWM2016-LymphomaCoalition.pdf (accessed: August 2020).
31. Karcher RA. Global burden of Waldenstrom macroglobulinemia: A systematic literature review and evidence gap analysis. ISPOR; 18-22 May 2019; New Orleans, USA.
32. Treon SP, et al. Paradoxical increases in serum IgM and viscosity levels following rituximab in Waldenstrom's macroglobulinemia. Ann Oncol. 2004;15(10):1481-1483.
33. Vincent C, et al. Adverse events in British hospitals: preliminary retrospective record review. BMJ. 2001;322(7285):517-519.
34. NHS Digital. Hospital Admitted Patient Care Activity, 2018-19: Diagnosis. 2019. Available at: https://digital.nhs.uk/data-and-information/publications/statistical/hospital-admitted-patient-care-activity/2018-19 (accessed: October 2020).
35. © National Institute for Health and Care Excellence. Ibrutinib for treating Waldenstrom's macroglobulinaemia [TA491]. November 2017. Available at: www.nice.org.uk/guidance/ta491 (accessed: October 2020).
36. Grunenberg A, Buske C. Treatment options for relapsed Waldenström's Macroglobulinemia. HemaSphere. 2019;3(S2):65-67.
37. Tang B, editor Current Treatment Patterns and Associated Outcomes in Waldenström Macroglobulinemia and Related Hematologic Malignancies: A Systematic Literature Review. American Society of Clinical Oncology; 2019 31 May-4 June 2019; Chicago, Illinois.
38. Leleu X, et al. Update on therapeutic options in Waldenstrom macroglobulinemia. Eur J Haematol. 2009;82(1):1-12.
39. D'Souza A, et al. Waldenstrom macroglobulinaemia: the key questions. Br J Haematol. 2013;162(3):295-303.
40. Oza A, Rajkumar SV. Waldenstrom macroglobulinemia: prognosis and management. Blood Cancer J. 2015;5:e394.
41. Gustine JN, et al. Ibrutinib discontinuation in Waldenstrom macroglobulinemia: Etiologies, outcomes, and IgM rebound. Am J Hematol. 2018;93(4):511-517.
42. Paydas S. Management of adverse effects/toxicity of ibrutinib. Crit Rev Oncol Hematol. 2019;136:56-63.
43. Garcia-Sanz R, et al. Updated results of the ASPEN trial from a cohort of patients with MYD88 wild-type (MYD88WT) Waldenström macroglobulinemia (WM). Journal of Clinical Oncology. 2020;38(15_suppl):e20056-e20056.
44. Trotman J, et al. Zanubrutinib for the treatment of patients with Waldenström macroglobulinemia: three years of follow-up. Blood. 2020.
45. Tam CS, et al. ASPEN: Results of a phase III randomized trial of zanubrutinib versus ibrutinib for patients with Waldenström macroglobulinemia (WM). Journal of Clinical Oncology. 2020;38(15_suppl):8007-8007.
46. BeiGene. Clinical Study Report: ASPEN BGB-3111-302. June 2020.
47. Dimopoulos M. Updated results of the ASPEN trial from a cohort of patients with MYD88 wild-type (MYD88WT) Waldenström macroglobulinemia (WM). American Society of Clinical Oncology; 29-31 May 2020; Virtual2020.
48. BeiGene. Clinical Study Report: BGB-3111-AU-003 2020.

49. Tedeschi A, et al. Bendamustine and rituximab combination is safe and effective as salvage regimen in Waldenstrom macroglobulinemia. Leuk Lymphoma. 2015;56(9):2637-2642.
50. Dimopoulos MA, et al. Primary treatment of Waldenström macroglobulinemia with dexamethasone, rituximab, and cyclophosphamide. J Clin Oncol. 2007;25(22):3344-3349.
51. Kastritis E, et al. Dexamethasone, rituximab, and cyclophosphamide as primary treatment of Waldenström macroglobulinemia: final analysis of a phase 2 study. Blood. 2015;126(11):1392-1394.
52. Phillippo D, et al. NICE DSU support document 18: methods for population-adjusted indirect comparisons in submissions to NICE. Sheffield, UK: Decision Support Unit; 2016.
53. Treon SP, et al. Ibrutinib in previously treated Waldenström's macroglobulinemia. N Engl J Med. 2015;372(15):1430-1440.
54. Signorovitch JE, et al. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. Value Health. 2012;15(6):940-947.
55. BeiGene. Summary of Clinical Safety. May 2020.
56. © National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. 2013. Available at: https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781 (accessed: September 15th 2020).
57. © National Institute for Health and Care Excellence. Ibrutinib for treating Waldenstrom's macroglobulinaemia. Final Appraisal Determination. 2016.
58. Janssen-Cilag. Ibrutinib Summary of Product Characteristics. 2021. Available at: https://www.ema.europa.eu/en/documents/product-information/imbruvica-epar-product-information_en.pdf (accessed: February 2021).
59. Latimer NR. NICE Decision Support Unit Technical Support Documents. Survival Analysis For Economic Evaluations Alongside Clinical Trials - Extrapolation with Patient-Level Data. London: National Institute for Health and Care Excellence (NICE)



60. Dimopoulos MA, et al. Phase 3 trial of ibrutinib plus rituximab in Waldenström's macroglobulinemia. N Engl J Med. 2018;378(25):2399-2410.
61. Office for National Statistics. National Life Tables, United Kingdom 2016-2018. 2020. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables (accessed: September 2020).
62. van Hout B, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value Health. 2012;15(5):708-715.
63. Dolan P, Kahneman D. Interpretations Of Utility And Their Implications For The Valuation Of Health*. The Economic Journal. 2008;118(525):215-234.
64. Devlin NJ, et al. Valuing health-related quality of life: An EQ-5D-5L value set for England. Health Econ. 2018;27(1):7-22.
65. © National Institute for Health and Care Excellence. Ibrutinib for treating chronic lymphocytic leukaemia [ID749]. Committee Papers 2016.
66. © National Institute for Health and Care Excellence. Ibrutinib for treating relapsed or refractory mantle cell lymphoma [ID753]. Committee Papers. 2016.
67. © National Institute for Health and Care Excellence. British National Formulary. 2020. Available at: https://bnf.nice.org.uk/ (accessed: August 6th 2020).
68. UK Department of Health. NHS reference cost 2018-2019. 2020. Available at: https://www.england.nhs.uk/national-cost-collection/#ncc1819 (accessed: October 16th 2020).
69. Aiello A, et al. Cost-effectiveness analysis of ibrutinib in patients with Waldenstrom macroglobulinemia in Italy. Journal of market access & health policy. 2017;5(1):1393308.
70. Round J, et al. Estimating the cost of caring for people with cancer at the end of life: A modelling study. (1477-030X (Electronic)).
71. Leblond V, et al. Treatment recommendations from the Eighth International Workshop on Waldenstrom's Macroglobulinemia. Blood. 2016;128(10):1321-1328.
72. Talaulikar D, et al. Treatment of patients with Waldenström Macroglobulinaemia: clinical practice guidelines from the Myeloma Foundation of Australia Medical and Scientific Advisory Group Intern Med J. 2017;47(1):35-49.
73. BeiGene. EU Medical Advisory Board 10 June 2020 (Virtual). Zanubrutinib and BTK Inhibitors for Treatment of Waldenström's Macroglobulinemia (WM). 2020.
74. BeiGene. Data on file. Survival Extrapolation Validation Meeting Minutes. 2020.

## B.5. Appendices

**Appendix C.** **Summary of product characteristics (SmPC) and European public assessment report (EPAR)**

**Appendix D.** **Identification, selection and synthesis of clinical evidence**

**Appendix E.** **Subgroup analysis**

**Appendix F.** **Adverse reactions**

**Appendix G.** **Published cost-effectiveness studies**

**Appendix H.** **Health-related quality-of-life studies**

**Appendix I.** **Cost and healthcare resource identification, measurement and valuation**

**Appendix J.** **Clinical outcomes and disaggregated results from the model**

**Appendix K.** **Checklist of confidential information**

**Appendix L.** **Clinical effectiveness – supplementary information**

# NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

## Single technology appraisal

## Zanubrutinib for treating Waldenström's macroglobulinaemia [ID1427]

## Clarification questions

**May 2021**

| File name | Version | Contains confidential information | Date |
|---|---|---|---|
| ID1427 Zanubrutinib clarification questions v3.1 20Jan2022 ACIC | 3.1 | Yes | 20 January 2022 |

**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

## *Literature Searches*

A1. The clinical and cost-effectiveness searches were run up to September 2020 and are now 6 months out of date. Please check whether any relevant literature has been published since the last search date and clarify what impact this might have on the clinical and cost-effectiveness systematic reviews.

As the literature was searched on 24 September 2020 and the submission dossier was submitted on the 12 March 2021, the searches were conducted within the 6-month timeframe acceptable to NICE. As such, no additional searches were carried out. However, based on this feedback, a separate targeted search based on the current search syntax was carried out to explore whether any new data were published within this timeframe. Two unique clinical publications that may have been included were identified. However, both of these publications would not have been relevant to the network-meta analysis:

- Castillo JJ, Gustine JN, Meid K, et al. Response and Survival Outcomes to Ibrutinib Monotherapy for Patients With Waldenström Macroglobulinemia on and off Clinical Trials. Hemasphere. 2020;4(3):e363
- Dimopoulos M, Sanz RG, Lee HP, et al. Zanubrutinib for the treatment of MYD88 wild-type Waldenström macroglobulinemia: a substudy of the phase 3 ASPEN trial. Blood Adv. 2020; 4(23):6009-6018

Regarding ibrutinib (data published in Castillo), data from the ASPEN study were used in the indirect treatment comparison. As this data originates from a randomised head-to-head trial investigating ibrutinib versus zanubrutinib, this is considered more robust and appropriate as it directly compares both interventions.

Regarding the sub study of the ASPEN trial, this sub study aimed to specifically investigate the safety and efficacy of zanubrutinib in Waldenström’s

macroglobulinaemia (WM) patients with MYD88$^{WT}$. The results of ASPEN are already presented in Section B.2.6.1.2 of the company submission (CS).

**A2.** Please explain the rationale for limiting the clinical and cost-effectiveness Embase searches to English language only. Please describe what steps were taken to mitigate for potential language bias.

The rationale for limiting the searches to English literature only was based on guidance provided by NICE; Chapter 5.4 of *Developing NICE guidelines: the manual* states that with regards to limits and filters, searches should be limited to studies reported in English.[1] Hence, a filter was used to include English literature only. In addition, according to the Cochrane handbook for Systematic Reviews of Interventions "Evidence indicates that excluding non-English studies does not change the conclusions of most systematic reviews (Morrison et al 2012, Jiao et al 2013, Hartling et al 2017), although exceptions have been observed for complementary and alternative medicine (Moher et al 2003, Pham et al 2005, Wu et al 2013). There is, however, also research related to language bias that supports the inclusion of non-English studies in systematic reviews (Egger et al 1997)."[2] A potential language bias may exist when the given intervention is more commonly provided in countries and regions that speak languages other than English, for instance in the case of Chinese herbal medicine.[2] However, given that trial publications for zanubrutinib and its comparators were identified from English literature, and no other publications within other languages were to be expected (e.g. not a therapeutic that is specifically given in other non-English speaking regions), searching English literature was appropriate.

## *Literature Searches - Clinical effectiveness*

**A3.** Section D.1.1.1 of the appendices to the company submission reports inclusion of the Database of Abstracts of Reviews of Effectiveness (DARE) within the Cochrane Library search, however Table B.5.4 does not present search results for the DARE database. As DARE was removed from the Cochrane Library in

September 2018, please clarify how DARE was searched and provide a full search strategy, reporting the hits per search line.

DARE was mistakenly added to the text. As DARE has not been part of the Cochrane library since September 2018 it was not searched.

## *Literature Searches - Cost effectiveness*

**A4.** Section G.1.1.2 of the appendices to the company submission reports EconLit was searched, however the search strategy is missing from Appendix G. Please provide a full search strategy, reporting the hits per search line.

The ProQuest database was used to search Embase, Medline and EconLit simultaneously. Hence, the search lines display the number of hits originating from all three sources.

**A5.** The Embase search reported in Table B.5.29 of the appendices to the company submission has duplicated search lines. Please clarify why these search terms are repeated within the strategy:

A. Lines S6, S10-14 are repeated later in the strategy as lines S39-43.

B. Please clarify why lines S14 and S6 is repeated later in the strategy as lines S78 and S79 respectively.

| **S6** | **MESH.EXACT("Economics")** |
|---|---|
| S10 | EMB.EXACT("Economic aspect") |
| S11 | EMB.EXACT("Socioeconomics") |
| S12 | MESH.EXACT("Economics, pharmaceutical") |
| S13 | EMB.EXACT("Health economics") |
| S14 | MESH.EXACT("Costs and cost analysis") |

| **S38** | **MESH.EXACT("Economics")** |
|---|---|
| S39 | EMB.EXACT("Economic aspect") |
| S40 | EMB.EXACT("Socioeconomics") |
| S41 | MESH.EXACT("Economics, pharmaceutical") |
| S42 | EMB.EXACT("Health economics") |
| S43 | MESH.EXACT("Costs and cost analysis") |

| | |
|---|---|
| **S78** | **MESH.EXACT.EXPLODE("Costs and cost analysis")** |
| S79 | EMB.EXACT("Economics") |

C. Please clarify why line S29 is repeated later in the strategy as line S77.

| | |
|---|---|
| **S29** | **TI,IF(economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed)** |

| | |
|---|---|
| **S77** | **TI,IF(Economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed)** |

The search syntax for the economic searches consists of combining search terms for WM with 1) cost-effectiveness (S4–S37), 2) healthcare resource use and costs (S38–S87) or 3) utilities (S88–S119). Each of the above highlighted duplicate parts of the search strategy originate from either the S6–S14 & S29; cost-effectiveness, or S38–S43 & S77; healthcare costs and resource use) and aims to capture publications pertaining to these topics. Given that both topics relate, similar search syntax has been used to capture these publications. This indeed means that there is overlap between the search syntax between the two topics. However, duplicates are automatically removed by ProQuest, as the software has been designed to filter out the duplicates by default in the final number of hits.

**A6.** The inclusion criteria in Table B.5.30 of the appendices to the company submission states that studies conducted in all countries would be considered for inclusion. As the cost-effectiveness search of Embase (Table B.5.29) was limited to English language, please clarify how the language limit may have restricted recall of international publications.

The rationale for limiting the searches to English literature only was based on guidance provided by NICE.[1] Given that this search intended to support the UK submission, a conscious decision was made to only include English literature.

## *Decision Problem*

**A7.** The NICE scope lists autologous stem cell transplantation (SCT), in people for whom autologous SCT is suitable as a relevant comparator for people who have had at least 1 prior therapy. Please clarify why this comparator was not included in the company submission.

The eligibility criteria for the search were based on the draft scope set by NICE, published in July 2020. As such, the current interventions listed reflect the draft scope. Only after the search was carried out, a final scope was published by NICE (27 October 2020) including SCT as one of the comparators (less than 5 months prior to submission). As such searches were carried out according to the draft scope and not the final scope. In addition, data from the UK WM Rory Morrison registry showed that 3% of all WM patients were considered for SCT.[3] Hence, SCT was not considered a relevant comparator.

## *Zanubrutinib Studies*

**A8.** The proportion of >75 year olds was higher in the zanubrutinib group compared with the ibrutinib group (company submission, Table B.2.7, Page 33). The disease affects older people more, but this also means that people over 75 years of age may have more room for improvement. Please explain how the different proportion of people over 75 years of age in each arm of ASPEN may bias the results. Please explain how this was adjusted for in the analyses? If so, please provide adjusted results. If no adjustment was made please explain why.

WM is a disease of older people with a median onset of 70 years. Patients who are older or who have comorbidities typically have a worse prognosis than those who are young and/or fit. [redacted].[4]

[redacted]

**Table 1 Subgroup analysis of the IRC-assessed VGPR/CR rate in Cohort 1 (ITT Analysis Set)**

| Subgroup | Zanubrutinib n/N (%) | Ibrutinib n/N (%) | Risk Difference [a] (95% CI), % |
|---|---|---|---|
| All patients | [redacted] | [redacted] | [redacted] |
| Age group | | | |
| ≤75 years | [redacted] | [redacted] | [redacted] |
| >75 years | [redacted] | [redacted] | [redacted] |

Data cut-off 31 August 2019

[a] Unstratified rate difference and 95% CI.

Abbreviations: CI = confidence interval; CR = complete response; IRC = independent review committee; ITT = intention-to-treat; VGPR = very good partial response

Source: BeiGene 2020[4]

[redacted]

**Table 2. CMH test for IRC-assessed VGPR/CR rate in Cohort 1 (ITT Analysis Set)**

| | Adjusted for randomisation stratification factors only, without age | Adjusted for randomisation stratification factors and age (>65 vs ≤65 years) | Adjusted for randomisation stratification factors and age (>75 vs ≤75 years) |
|---|---|---|---|
| Risk difference (zanubrutinib-ibrutinib), % (95% CI)[a] | [redacted] | [redacted] | [redacted] |
| p-value[b] | [redacted] | [redacted] | [redacted] |

Data cut-off 31 August 2019
[a] Mantel-Haenszel common risk difference stratified by the randomisation stratification factors, i.e. prior anti-cancer therapy (0, 1–3, >3) and *CXCR4* (mutated vs. WT/UNK), and age.
[b] Cochran-Mantel-Haenszel stratified by the randomisation stratification factors, i.e., prior anti-cancer therapy (0, 1–3, >3) and *CXCR4* (Mutated vs. WT/UNK), and age.
Abbreviations: CI = confidence interval; CMH = Cochran-Mantel-Haenszel; CR = complete response; IRC = independent review committee; ITT = intention-to-treat; UNK = unknown; VGPR = very good partial response; WT = wild type

**A9.** Please explain why patients with cardiovascular disease and those taking warfarin were excluded from the ASPEN trial. Please provide an estimate of the proportion of UK patients with cardiovascular disease or taking warfarin within the target population (the population described in the scope).

Such exclusion criteria are common in clinical trials in order to prevent patients with severe underlying comorbidities being exposed to potential side effects. Atrial fibrillation is a well-known adverse event (AE) associated with ibrutinib treatment; in the pivotal multicentre trial of ibrutinib monotherapy in previously treated patients with WM, incidence of atrial fibrillation was 12.7%.[5] Occurrence of atrial fibrillation is associated with an increased risk of death from any cause and an increased risk of cardiovascular diseases including all-cause cardiovascular mortality, major cardiovascular events, ischaemic heart disease, sudden cardiac death, heart failure and peripheral arterial disease.[6]

A special warning regarding cardiac risk factors is included in the draft Summary of Product Characteristics (SmPC) for zanubrutinib: “Cases of atrial fibrillation and atrial flutter have been reported particularly in patients with cardiac risk factors, hypertension, acute infections, and a previous history of atrial fibrillation.[7]” As underscored in the SmPC, “patients with severe cardiovascular disease were excluded from [ibrutinib] clinical studies”.[7] Consequently, BeiGene did not want to

expose patients with underlying comorbidities to known or unknown side-effects, patients with cardiovascular disease were excluded from the ASPEN trial.

Ibrutinib use is also associated with increased rate of bleeding events in patients treated for any indication and for patients treated for WM.[8] These include minor and major bleeding events, some of these events being fatal.[8] Therefore, patients with concomitant treatment with anticoagulants are excluded from ibrutinib clinical trials. In addition, the SmPC for ibrutinib states that “warfarin or other vitamin K antagonists should not be administered concomitantly with [ibrutinib]”.[9] Consequently, patients with concomitant treatment with warfarin were excluded from ASPEN trial. However, concomitant heparin treatment was allowed.[4]

**A10.** The cut-off date for the ASPEN study was 31/08/2019 (see B.2.3.5.2 company submission). Are there any further follow-up data available? Or are there further analyses planned?

Yes, the ASPEN study is an ongoing study. A follow-up analysis of safety and efficacy has been conducted with the cut-off date of 31 August 2020 and is documented in an efficacy addendum. This efficacy addendum was submitted to the European Medicines Agency (EMA) as part of the “Day 120 responses” on 18 February 2021 during the ongoing centralised procedure. The efficacy addendum includes efficacy data as judged by the investigators, with an amendment planned to include efficacy data as judged by the independent review committee (IRC) with the same cut-off date (submission to EMA currently planned in late May 2021 as part of the ongoing centralised procedure). The addendum submitted to EMA is included with the response to these questions, and the amendment will be provided to NICE when available.

**A11.** In Tables B.2.10-B.2.16, could you please clarify whether the medians are presented with interquartile ranges or 95% confidence intervals.

Medians in Tables B.2.10–B.2.16 are presented with 95% confidence intervals.

**A12. Priority Question: Please extend figures B.2.4 and B.2.5 to 30 months, as in figures B.3.24 onwards, including the number of patients at risk. Please also provide the number of patients at risk in figures B.3.24 to B.3.28, as in figures B.2.4 and B.2.5.**

Progression-free survival (PFS) and overall survival (OS) curves extended to [redacted] are presented in Figure 1 and Figure 2,Abbreviations: CS = company submission; CSR = clinical study report; PFS = progression-free survival

Figure 2 respectively.

**Figure 1. PFS extended to [redacted] in Cohort 1 (Figure B.2.4 in CS; Figure 8 in the CSR)**

Abbreviations: CS = company submission; CSR = clinical study report; PFS = progression-free survival

**Figure 2. Figure B.2.5 OS extended to [redacted] in Cohort 1 (Figure B.2.5 in CS; Figure 11 in the CSR)**

Abbreviations: CS = company submission; CSR = clinical study report; PFS = progression-free survival

**A13.** In section B.2.6.1.1.6, please provide effect estimates as well as P values.

Reduction of IgM level from baseline over time was compared between the two treatment arms. The analysis was done using a likelihood based linear mixed model for repeated measures (MMRM) and a non-parametric comparison of the area under curve (AUC) of the IgM over time by Mantel-Haenszel test.

In the MMRM approach, treatment effect was estimated by the difference in the slope of the IgM reduction (in log scale) over time, i.e., the interaction between time and treatment arm. A covariance structure of compound symmetry was assumed in the analysis. The slope of the log IgM change from baseline over time was [redacted] and [redacted] for zanubrutinib and ibrutinib, respectively, in the ITT analysis set with a 2-sided p-value of [redacted] for the time and treatment interaction (Table 3).

**Table 3. Analysis of IgM (log) change from baseline in Cohort 1 (ITT Analysis Set)**

| Parameter | Zanubrutinib (N=102) | Ibrutinib (N=99) |
|---|---|---|
| Slope over time (months) | [redacted] | [redacted] |
| p value (2-sided)[a] | [redacted] | |
| AUC | | |
| Top 25 percentile | [redacted] | [redacted] |
| Median | [redacted] | [redacted] |
| p-value (2 sided)[b] | [redacted] | |

Data cut-off 31 August 2019
[a] From repeated measure mixed effect model with time as continuous variable and treatment arm and time as fixed effects. P-value is for the interaction between treatment arm and time effects.
[b] From Mantel-Haenszel test.
Abbreviations: AUC = area under curve; IgM = immunoglobulin M; ITT = intention-to-treat
Source: BeiGene 2020[4]

In addition to the MMRM analysis, which assumes a linear trend in the IgM change (in log scale) over time, a non-parametric AUC analysis was performed to compare zanubrutinib with ibrutinib. The AUC distributions are shown in Figure 3.

**Figure 3. AUC distribution in Cohort 1 (ITT Analysis Set)**

Abbreviations: AUC = area under the curve; ITT= intention-to-treat

The median AUC was [redacted] for zanubrutinib and [redacted] for ibrutinib in the ITT analysis set. The treatment arm difference is larger among the patients with higher AUC. The top 25 percentile AUC was [redacted] and [redacted] for zanubrutinib and ibrutinib, respectively.

[redacted]

**A14.** In section B.2.6.1.1.7, please provide both effect estimates and P values for the comparison of outcomes between zanubrutinib and ibrutinib.

A linear MMRM was performed on the QLQ-C30 global health score. The model included the repeated measurement (including baseline) of QLQ-C30 global health status/QoL scale up to Cycle 25 Day 1 as dependent variable and treatment arm, randomisation stratification factors, i.e. prior anti-cancer therapy (0, 1–3, >3) and *CXCR4* (mutated vs WT/UNK), intercept, and slope of time as fixed effects. The random subject effects included subject random intercept and subject random slope of time. There was no statistically significant difference between the two treatment arms for QLQ-C30 global health status/QoL scale scores with a treatment difference of -0.69 and a p-value of 0.751 (Table 4). No other treatment effects or p-values were estimated for the patient-reported outcomes.

**Table 4. Analysis of EORTC QLQ-C30 Questionnaire Global Health Status/QoL in Cohort 1 (ITT Analysis Set)**

| | **Ibrutinib (N = 99)** | **Zanubrutinib (N = 102)** |
|---|---|---|
| n | 99 | 101 |
| LSmean (SE) a | 69.07 (2.369) | 68.38 (2.299) |
| Treatment difference a | | -0.69 |
| 95% CI | | (-4.95, 3.57) |
| p-value | | 0.751 |

Data cut-off 31 August 2019
Abbreviations: ITT = intention-to-treat; LS = least squared; SE = standard error; QoL = quality of life

n = number of patients with any baseline or postbaseline result on or before cycle 25 day 1.
[a] Based on a linear MMRM. The model includes the repeated measurement (including baseline) of QLQ-C30 global health status/QoL scale up to Cycle 25 Day 1 as dependent variable and treatment arm, randomisation stratification factors (from IRT), intercept and slope of time as fixed effects. The random subject effects includes subject random intercept and subject random slope of time. The randomisation stratification factors include the *CXCR4* mutation status (WHIM vs WT/Unknown) and prior line of therapy (1–3 vs. >3 for relapsed/refractory population and 0 vs 1–3 vs >3 for overall population). Ibrutinib is the reference group.

## A15. Priority Question: Please provide data for the number, reasons for withdrawing and timing of patients who withdrew from ASPEN, including whether the patients had a response to treatment, and if so, the magnitude of their response.

As of 31 August 2019, a total of 6 patients had withdrawn consent from study treatment (5 patients in Cohort 1, and 1 in Cohort 2). Table 5 summarises the timing, descriptive narrative of events upon consent withdrawal decision by patient (did not resume dosing as instructed) and the best overall response to treatment for these patients.

**Table 5. Patients withdrawing from ASPEN**

| Subject | Cohort | RR/TN | Treatment | Consent Withdrawal Study Day | Best overall response by investigator | Best overall response by IR | Narrative |
|---|---|---|---|---|---|---|---|
| 3109001 | Cohort 1 | RR | Zanubrutinib | 48 | MR | MR | Patient stopped taking study drug, preceded by a temporary dose hold for SAE Grade 3 atrioventricular block second degree. Patient had medical history of atrioventricular block and cerebrovascular accident. |
| 3382001 | Cohort 1 | TN | Zanubrutinib | 368 | MR | PR | "Patient decision due to medical condition and not due to specific AE" was referenced. Patient had also been in IgM flare for about 6 months with two dose holds due to AE (Grade 4 gastric haemorrhage and Grade 2 eye haemorrhage). |
| 3531002 | Cohort 2 | RR | Zanubrutinib | 559 | MR | PR | Patient decided to stop taking study drug, preceded by a temporary dose hold for SAE Grade 2 respiratory infection. Patient had baseline condition of chronic obstructive pulmonary disease. |
| 3539001 | Cohort 1 | RR | Zanubrutinib | 471 | VGPR | VGPR | Patient decided to stop taking study drug, preceded by a temporary dose hold for SAE Grade 4 atrioventricular block complete and concurrent Grade 2 infections. Patient had ongoing baseline condition of bundle branch block left; and had maintained VGPR response for about 1 year on study treatment. |
| 3539002 | Cohort 1 | RR | Zanubrutinib | 120 | SD | SD | Patient decision, preceded by patient-initiated dose holds with reference to AEs and baseline Grade 3 hypertension. Ongoing AEs at the time included Grade 2 depression, oedema and cardiac failure. Patient also had ongoing baseline condition of Grade 2 atrial fibrillation. |
| 3806002 | Cohort 1 | TN | Zanubrutinib | 56 | PR | MR | Patient decided to stop taking study drug, preceded by recurrent dose holds primarily due to Grade 2 contusion and Grade 1 dyspnoea, and experienced withdrawal symptoms including worsening arthralgia. |

Abbreviations: AE = adverse event; IR = independent review; MR = minor response; PR = partial response; RR = relapsed/refractory; SAE = serious adverse event; SD = stable disease; TN = treatment naïve; VGPR = very good partial response

**A16.** Table B.5.8 (appendices to company submission) has 134 female patients in ASPEN – is this an error?

This was an error in the CS. Table B.5.8 should read 134 **male** patients.

**A17.** It is stated in the company submission (page 26) that "Patients with *MYD88*$^{WT}$ disease or with undetermined *MYD88* mutation status were enrolled in Cohort 2 (Arm C)" of ASPEN. Please confirm that Cohort 1 did not include any patients with *MYD88*$^{WT}$ disease or with undetermined *MYD88* mutation status. Please also clarify whether assessment of *MYD88* mutation status is standard UK practice for the population defined in the scope (Adults with Waldenström's macroglobulinaemia who have had at least 1 prior therapy, or whose disease is untreated, for whom chemo-immunotherapy is unsuitable).

Cohort 1 did not include any patients with *MYD88*$^{WT}$ or with undetermined *MYD88* status. A UK WM clinical expert confirmed that testing for *MYD88* mutation is the standard of care at most of the 24 WM centres in the UK, which have treated 90% of the UK WM patient population since 2016.

**A18.** As stated in the company submission (page 24), 33 UK patients were included in the ASPEN study. Please specify how many UK patients were included in each arm of Cohort 1.

A total of 33 UK patients were randomised in the ASPEN study (30 randomised to Cohort 1 and 3 to Cohort 2), and 32 patients were treated. In Cohort 1, 13 patients were treated with ibrutinib, and 16 of 17 randomised were treated with zanubrutinib.

## *Systematic Review*

**A19.** The ERG noticed that some studies that are listed in Table B.5.6 are not listed in Table B.5.7. (appendices to company submission) Could the company please provide detailed justification as to why some of those studies included in the review

and presented in table B.5.6 were not included in the indirect treatment comparison and presented in Table B.5.7 (While a full list of excluded studies is included in an Excel file, the reasons for exclusion are not).

A. The reasons for excluding studies from the matching adjusted indirect comparison (MAIC) (Table B.5.8) were also not provided. Please provide a justification for excluding studies from the MAIC?

The list of studies that were included in Table B.5.6 (identified from the clinical SLR) are summarised in Table 6, along with the rationales why some of the studies were not included in Table B.5.7 (that was specific for the discussion of the studies for potential inclusion in indirect treatment comparisons).

**Table 6. Studies identified in clinical SLR**

| Study | Citation | Inclusion in Table B.5.7 and rationale for the exclusion |
|---|---|---|
| Abeykoon 2020 | Abeykoon, J.P. et al., Ibrutinib monotherapy outside of clinical trial setting in Waldenström macroglobulinaemia: practice patterns, toxicities and outcomes. British journal of haematology. 2020. 188:394-403 | Not included.<br>The comparison between zanubrutinib and ibrutinib relied on the phase 3 randomised ASPEN trial, the only study, to our knowledge and the findings of the SLR, that directly compares zanubrutinib with ibrutinib, which was considered to be more robust than relying on indirect treatment comparisons. Therefore, ibrutinib studies identified from the clinical SLR were not assessed in Table B.5.7 regarding whether these studies were appropriate for inclusion in indirect treatment comparisons. |
| Buske 2016 | Buske, C. et al., Single-agent Ibrutinib is efficacious and well tolerated in Rituximab-refractory patients with Waldenstrom's Macroglobulinemia (WM): initial results from an international, multicenter, open-label phase 3 substudy (iNNOVATETM). Oncology research and treatment. Conference: jahrestagung der deutschen, osterreichischen und schweizerischen gesellschaften fur hamatologie und medizinische onkologie 2016. Germany. Conference start: 20161014. Conference end: 20161018. 2016. 39:119 | |
| Buske 2009 | Buske, C. et al., The addition of rituximab to front-line therapy with CHOP (R-CHOP) results in a higher response rate and longer time to treatment failure in patients with lymphoplasmacytic lymphoma: Results of a randomised trial of the German Low-Grade Lymphoma Study Group (GLSG). Leukemia. 2009. 23:153-161 | Not included.<br>R-CHOP was included as a comparator of interest in the draft scope (which the clinical SLR was based on) but not included in the final scope. Therefore, R-CHOP was not included for assessment of potential inclusion in indirect comparisons. |
| Byrd 1999 | Byrd, J.C. et al., Rituximab therapy in Waldenstrom's macroglobulinemia: Preliminary evidence of clinical activity. Annals of Oncology. 1999. 10:1525-1527 | Included |
| Castillo 2018 | Castillo, J. et al., Response and survival for primary therapy combination regimens and maintenance rituximab in Waldenström | Included. The company reached out to the first author, Dr. Castillo. He confirmed that “in our paper, all the patients were in |

| Study | Citation | Inclusion in Table B.5.7 and rationale for the exclusion |
|---|---|---|
| | macroglobulinaemia. British Journal of Haematology. 2018. 181:77-85 | the frontline setting and all patients were fit enough to be good candidates for chemo-immunotherapy. I think it would not be fair to compare a mixed pool of patients (treatment naïve and relapsed/refractory) treated with zanubrutinib versus a purely treatment-naïve group." |
| Dimopoulos 2020 | Dimopoulos, M.A. et al., Aspen: results of a phase 3 randomised trial of zanubrutinib versus ibrutinib for patients with waldenstrom macroglobulinemia (WM). Hemasphere. 2020. 4:71- | Not included.<br>This is a publication for the ASPEN trial, the phase 3 trial for zanubrutinib in WM, which is not relevant to Table B.5.7 which aimed for assessment of the comparator studies for potential inclusion in indirect treatment comparisons. |
| Dimopoulos 2016a | Dimopoulos, M.A. et al., Efficacy and safety of single-agent ibrutinib in rituximab-refractory patients with Waldenstrom's macroglobulinemia (WM): initial results from an international, multicenter, open-label phase 3 substudy (iNNOVATETM). British journal of haematology. Conference: 36th world congress of the international society of hematology. United kingdom. 2016. 173:82 | Not included.<br>The comparison between zanubrutinib and ibrutinib relied on the phase 3 randomised ASPEN trial, the only study, to our knowledge and the findings of the SLR, that directly compares zanubrutinib to ibrutinib, which was considered to be more robust than relying on indirect treatment comparisons. Therefore, ibrutinib studies identified from the clinical SLR were not assessed in Table B.5.7 regarding whether these studies were appropriate for inclusion in indirect treatment comparisons. |
| Dimopoulos 2016b | Dimopoulos, M.A. et al., Ibrutinib for patients with rituximab-refractory Waldenstrom's macroglobulinaemia (iNNOVATE): an open-label substudy of an international, multicentre, phase 3 trial. Lancet oncology. 2016. | |
| Dimopoulos 2016c | Dimopoulos, M.A. et al., Single-agent ibrutinib in rituximab-refractory patients with waldenstrom's macroglobulinemia (WM): updated results from a multicenter, open-label phase 3 substudy (innovatetm). Haematologica. Conference: 21st congress of the european hematology association. Denmark. 2016. 101:256-257 | |
| Dimopoulos 2002a | Dimopoulos, M.A. et al., Treatment of Waldenström's macroglobulinemia with rituximab. Journal of Clinical Oncology. 2002. 20:2327-2333 | Included |
| Dimopoulos 2002b | Dimopoulos, M.A. et al., Extended rituximab therapy for previously untreated patients with Waldenström's macroglubulinemia. Clinical Lymphoma. 2002. 3:163-166 | Included |
| Dimopoulos 2007 | Dimopoulos, M.A. et al., Primary treatment of Waldenström macroglobulinemia with dexamethasone, rituximab, and cyclophosphamide. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2007. 25:3344-9 | Included |
| Gertz 2009 | Gertz, M.A. et al., Clinical value of minor responses after 4 doses of rituximab in Waldenström macroglobulinaemia: A follow-up of the Eastern Cooperative Oncology Group | Included |

| Study | Citation | Inclusion in Table B.5.7 and rationale for the exclusion |
|---|---|---|
| | E3A98 trial. British Journal of Haematology. 2009. 147:677-680 | |
| Gertz 2004 | Gertz, M.A. et al., Multicenter phase 2 trial of rituximab for Waldenström macroglobulinemia (WM): An Eastern Cooperative Oncology Group Study (E3A98). Leukemia and Lymphoma. 2004. 45:2047-2055 | Included |
| Kastritis 2015 | Kastritis, E. et al., Dexamethasone, rituximab, and cyclophosphamide as primary treatment of Waldenström macroglobulinemia: final analysis of a phase 2 study. Blood. 2015. 126:1392-1394 | Included |
| Kyle 2000 | Kyle, R.A. et al., Waldenström's macroglobulinaemia: a prospective study comparing daily with intermittent oral chlorambucil. British journal of haematology. 2000. 108:737-742 | Included |
| Ngan 2003 | Ngan, S. et al., Waldenstrom's macroglobulinemia: a retrospective analysis of 40 patients from 1972 to 2001. Seminars in oncology. 2003. 30:236-8 | Included |
| Paludo 2018 | Paludo, J, et al., Bendamustine and rituximab (BR) versus dexamethasone, rituximab, and cyclophosphamide (DRC) in patients with Waldenström macroglobulinemia. Annals of Hematology. 2018. 97:1417-1425 | Included |
| Paludo 2016a | Paludo, J. et al., Bendamustine and rituximab versus dexamethasone, rituximab and cyclophosphamide in patients with waldenstrom macroglobulinemia (WM). Blood. 2016. 128 | Not included.<br>Only full-text articles were considered for inclusion in indirect treatment comparisons. Hence, Paludo 2016a, a conference abstract was not included in Table B.5.7. |
| Paludo 2016b | Paludo, J. et al., Dexamethasone, rituximab and cyclophosphamide (DRC) as salvage therapy for Waldenstrom macroglobulinemia. Blood. Conference: 58th annual meeting of the american society of hematology, ASH 2016. United states. Conference start: 20161203. 2016. 128: | Not included.<br>Only full-text articles were considered for inclusion in indirect treatment comparisons. Hence, Paludo 2016b, a conference abstract was not included in Table B.5.7. |
| Paludo 2017 | Paludo, J. et al., Dexamethasone, rituximab and cyclophosphamide for relapsed and/or refractory and treatment-naïve patients with Waldenstrom macroglobulinemia. British Journal of Haematology. 2017. 179:98-105 | Included |
| Souchet 2016 | Souchet, L. et al., Efficacy and long-term toxicity of the rituximab-fludarabine-cyclophosphamide combination therapy in Waldenstrom's macroglobulinemia. American Journal of Hematology. 2016. 91:782-786 | Included |
| Tam 2020 | Tam, C.S. et al., A randomised phase 3 trial of Zanubrutinib versus Ibrutinib in Symptomatic Waldenström Macroglobulinemia:The ASPEN study. Blood. 2020. | Not included.<br>This is a publication for the ASPEN trial, the phase 3 trial for zanubrutinib in WM, which is not relevant to Table B.5.7 which aimed to assess the comparator studies for potential inclusion in indirect treatment comparisons. |

| Study | Citation | Inclusion in Table B.5.7 and rationale for the exclusion |
|---|---|---|
| Tam 2005 | Tam, C.S. et al., Fludarabine combation therapy is highly effective in first-line and salvage treatment of patients with Waldenström's macroglobulinemia. Clinical Lymphoma and Myeloma. 2005. 6:136-139 | Included |
| Tedeschi 2015 | Tedeschi, A. et al., Bendamustine and rituximab combination is safe and effective as salvage regimen in Waldenström macroglobulinemia. Leukemia and Lymphoma. 2015. 56:2637-2642 | Included |
| Tedeschi 2012 | Tedeschi, A. et al., Fludarabine plus cyclophosphamide and rituximab in Waldenstrom macroglobulinemia: an effective but myelosuppressive regimen to be offered to patients with advanced disease. Cancer. 2012. 118:434-43 | Included |
| Tedeschi 2013 | Tedeschi, A. et al., Fludarabine, cyclophosphamide, and rituximab in salvage therapy of waldenström's macroglobulinemia. Clinical Lymphoma, Myeloma and Leukemia. 2013. 13:231-234 | Included |
| Treon 2005 | Treon, S.P. et al., CHOP plus rituximab therapy in Waldenström's macroglobulinemia. Clinical Lymphoma. 2005. 5:273-277 | Not included.<br>R-CHOP was included as a comparator of interest in the draft scope (which the clinical SLR was based on) but not included in the final scope. Therefore, R-CHOP was not included for assessment of potential inclusion in indirect comparisons. |
| Treon 2011 | Treon, S.P. et al., Bendamustine therapy in patients with relapsed or refractory Waldenström's macroglobulinemia. Clinical Lymphoma, Myeloma and Leukemia. 2011. 11:133-135 | Included |
| Treon 2001 | Treon, S.P. et al., CD20-directed antibody-mediated immunotherapy induces responses and facilitates hematologic recovery in patients with Waldenstrom's Macroglobulinemia. Journal of Immunotherapy. 2001. 24:272-279 | Included |
| Treon 2015 | Treon, S.P. et al., Ibrutinib in previously treated Waldenström's macroglobulinemia. New England Journal of Medicine. 2015. 372:1430-1440 | Not included.<br>The comparison between zanubrutinib and ibrutinib relied on the phase 3 randomised ASPEN trial, the only study, to our knowledge and the findings of the SLR, that directly compares zanubrutinib with ibrutinib, which was considered to be more robust than relying on indirect treatment comparisons. Therefore, ibrutinib studies identified from the clinical SLR were not assessed in Table B.5.7 regarding whether these studies were appropriate for inclusion in indirect treatment comparisons. |
| Treon 2020 | Treon, S.P. et al., Long-Term Follow-Up of Ibrutinib Monotherapy in Symptomatic, Previously Treated Patients With Waldenström Macroglobulinemia. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2020. | |
| Treon 2009 | Treon, S.P. et al., Long-term outcomes to fludarabine and rituximab in Waldenström macroglobulinemia. Blood. 2009. 113:3673-3678 | Included |

| Study | Citation | Inclusion in Table B.5.7 and rationale for the exclusion |
|---|---|---|
| Trotman 2020 | Trotman, J. et al., Zanubrutinib for the treatment of patients with Waldenström macroglobulinemia: three years of follow-up. Blood. 2020. | Not included.<br>This is a publication for the ASPEN trial, the phase 3 trial for zanubrutinib in WM, which is not relevant to Table B.5.7 which aimed to assess the comparator studies for potential inclusion in indirect treatment comparisons. |

Abbreviations: R-CHOP = rituximab, cyclophosphamide, doxorubicin, vincristine and prednisone; SLR = systematic literature review; WM = Waldenström's macroglobulinaemia

## *Indirect Comparisons*

**A20.** In section D.1.4 (appendices to company submission), the company states: “A quality assessment of studies included in the NMA is provided in Table B.5.13” – should “NMA” be “MAIC”?

Correct; this was a typographical error in the CS. “NMA” should be “MAIC”.

**A21.** Please provide a comparison of the baseline characteristics of patients in the three studies included in the MAIC, and justify that the populations are similar enough to be combined in the MAIC.

A comparison of baseline patient characteristics are provided below based on the available patient characteristics, including (1) the exact population characteristics presented in Table B.5.10–Table B.5.12 in the appendices to the CS, and (2) chi-square tests for statistical comparisons of the patient characteristics (all of which were either binary or categorical) that were not presented in the initial company submission.

*Zanubrutinib versus DRC*

Baseline characteristics for patients treated with zanubrutinib and DRC before matching are shown in Table 7.

**Table 7. Baseline characteristics (zanubrutinib vs DRC)**

| Baseline characteristics | Before matching | | |
|---|---|---|---|
| | Zanubrutinib, % (n=102) | DRC, % (n=72) | p-value |
| Age ≤65 years | 40.6 | 37.5 | 0.753 |
| Age 65–≤69 years | 6.9 | 12.5 | 0.288 |
| Age >69 years | 52.5 | 50.0 | 0.760 |
| Platelet count <100 $\times 10^9$/L | 11.9 | 4.2 | 0.101 |
| Haemoglobin <100 g/L | 46.5 | 56.9 | 0.217 |
| Presence of extramedullary disease: lymphadenopathy (by investigator) | 59.4 | 38.9 | 0.009 |
| Presence of extramedullary disease: splenomegaly (by investigator) | 15.8 | 31.9 | 0.016 |

Abbreviations: DRC = dexamethasone-rituximab-cyclophosphamide

Compared with the population treated with zanubrutinib, the population treated with DRC were treatment naïve and suitable for chemo-immunotherapies (N=72), in contrast to the zanubrutinib arm in the ASPEN trial (n=102) which included a mix of relapsed/refractory (n=83) and treatment-naïve patients considered unsuitable for chemo-immunotherapy (n=19). As discussed in the CS, such differences in patient populations could not be adjusted via MAIC, which might have led to an underestimation of the relative survival benefit of zanubrutinib compared with DRC, assuming that PFS and OS outcomes are more favourable in the treatment-naïve population (suitable for chemo-immunotherapy) compared with both the treatment-naïve population (unsuitable for chemo-immunotherapy) and the relapsed/refractory population. However, given the limited clinical evidence for DRC in WM, it is still a valuable and conservative option to include this study in the MAIC.

In addition to the difference in prior line of treatment and suitability for chemo-immunotherapy, a statistically significantly lower proportion of patients with extramedullary disease with lymphadenopathy (p=0.009) and significantly higher proportion of patients with extramedullary disease with splenomegaly (p=0.016) were observed for the DRC population. The differences in the rate of cytopenias (higher rate of haemoglobin <100 g/L and platelet count <110 X $10^9$/L) noted in the zanubrutinib treatment arm may also reflect worse disease. The other characteristics were comparable between the populations, based on the results of the chi-square tests. After matching adjustment, these patient characteristics were balanced between the populations (as shown in the Table 7). Matching adjustment led to a fair

effective sample size of 53 for the zanubrutinib arm, indicating modest differences in the populations.

*Zanubrutinib versus BR*

Baseline characteristics for patients treated with zanubrutinib and DRC before matching are shown in Table 8.

**Table 8. Baseline characteristics (zanubrutinib vs BR)**

| **Baseline Characteristics** | **Before Matching** | | |
|---|---|---|---|
| **Zanubrutinib in the ITT population of ASPEN** | **Zanubrutinib, % (n = 102)** | **BR, % (n = 71)** | **p-value** |
| Age ≤72 years | 58.4 | 50.0 | 0.352 |
| 0–2 prior lines of therapy | 79.2 | 50.0 | 0.0001 |
| IgM ≤38.15 g/L | 63.4 | 50.0 | 0.117 |
| IPSSWM score, intermediate risk | 37.6 | 30.4 | 0.388 |
| IPSSWM score, high risk | 45.5 | 48.2 | 0.867 |
| Presence of extramedullary disease: either splenomegaly or adenopathy (by investigator) | 61.4 | 42.3 | 0.020 |
| **Zanubrutinib in the relapsed/refractory subpopulation of ASPEN** | **Zanubrutinib, % (n = 83)** | **BR, % (n = 71)** | **p-value** |
| Age ≤72 years | 61.4 | 50.0 | 0.196 |
| 0–2 prior lines of therapy | 74.7 | 50.0 | 0.003 |
| IgM ≤38.15 g/L | 65.1 | 50.0 | 0.101 |
| IPSSWM score, intermediate risk | 36.1 | 30.4 | 0.584 |
| IPSSWM score, high risk | 44.6 | 48.2 | 0.730 |
| Presence of extramedullary disease: either splenomegaly or adenopathy (by investigator) | 63.9 | 42.3 | 0.009 |

Abbreviation: BR = bendamustine-rituximab; IgM = immunoglobulin M; IPSSWM = International Prognostic Scoring System for Waldenström's Macroglobulinaemia; ITT = intention-to-treat

Compared with the population treated with zanubrutinib, the population treated with BR included a statistically significantly higher proportion of heavily treated patients (p=0.0001 when adjusting the overall population of zanubrutinib to match the BR population) and a significantly higher proportion of patients with extramedullary disease with either splenomegaly or adenopathy (p=0.02 when adjusting the overall population of zanubrutinib to match the BR population). The other characteristics were comparable between the populations, based on the results of the chi-square tests. After matching adjustment, these patient characteristics were balanced between the populations, which led to a fair effective sample size of 50 for the zanubrutinib arm. The results of adjusting the relapsed/refractory subpopulation of

zanubrutinib are similar to the results when adjusting the overall population of zanubrutinib.

In addition to the observed patient characteristics discussed above between zanubrutinib and DRC and between zanubrutinib and BR, there were additional factors to consider that could increase the uncertainties as to the assessment of similarities between populations, including (1) the known differences in study characteristics (e.g., geographic location, year of study; see responses to questions A22 and A25), and (2) unobserved patient characteristics.

**A22. Priority question: In the MAIC, the individual participant data (IPD) of the ASPEN study (zanubrutinib arm only) are matched, so that the baseline characteristics of patients from the ASPEN study match the summary data of baseline characteristics in each of the two other single-arm studies. Please justify, with evidence, that the populations in each of the two other single-arm studies are representative of the population described in the scope (Adults with Waldenström's macroglobulinaemia who have had at least 1 prior therapy, or whose disease is untreated, for whom chemo-immunotherapy is unsuitable) in UK clinical practice.**

A. **If there is any indication that the populations in the single arm studies do not reflect the population in the scope then please discuss the potential implications for prognosis of the people included in the single arm studies and the treatment effect of zanubrutinib versus any of the comparators.**

The definition of the population with WM of the studies included in the MAIC are broadly in line with the population of interest per the NICE scope. In addition, both studies were EU-based and therefore more likely to be representative of the WM population in the UK (compared to the non-EU-based studies identified from the clinical SLR with relatively smaller sample size), which was also one of the criteria for inclusion of studies in the MAIC in the initial company submission. Despite the

general alignment in patient population and geographic location, there are several differences to be acknowledged for each study, as summarised in Table 9.

**Table 9. Patient populations for DRC and BR**

| | Population of interest per the NICE scope | Population per the DRC study[10,11] | Population per the BR study[12] |
|---|---|---|---|
| Population of WM | Adults with WM who have had at least 1 prior therapy, or whose disease is untreated, for whom chemo-immunotherapy is unsuitable | Adults with previous untreated WM, who were suitable for treatment with chemo-immunotherapy such as DRC, the treatment of interest of this interventional study | Adults with relapsed/refractory WM |
| Geographic location of interest / study | UK | Greece | Italy |
| Year of participant enrolment | N/A | 2002–2006 | NR |
| Median follow-up of study | N/A | 8 years | 19 months |
| Patient characteristics | | | |
| Age, year | | | |
| Mean (SD) | N/A | NR | NR |
| Median | N/A | 72 | 69 |
| Range | N/A | 49–88 | 33–89 |
| >65, n (%) | N/A | NR | 63% |
| Female proportion | N/A | 25 (35.2%) | 45 (62.5%) |
| IgM, g/L | | | |
| Mean (SD) | N/A | NR | NR |
| Median | N/A | 38.15 | NR |
| Range | N/A | 2.4–96.2 | NR |
| Platelet count, $10^9$/L | | | |
| Mean (SD) | N/A | NR | NR |
| Median | N/A | NR | NR |
| Range | N/A | NR | NR |
| ≤100, n (%) | N/A | NR | 3 (4.2%) |
| Haemoglobin, g/L | | | |
| Mean (SD) | N/A | NR | NR |
| Median | N/A | NR | NR |
| Range | N/A | NR | NR |
| <100, n (%) | N/A | NR | 41 (56.9%) |
| Prior line of treatment | | | |
| Median | N/A | 2 | N/A |
| Range | N/A | 1-5 | N/A |
| 0, n (%) | N/A | NR | N/A |
| 1–3, n (%) | N/A | NR | N/A |
| >3, n (%) | N/A | NR | N/A |
| Prior treatment regimen, n (%) | | | |
| Nucleoside analogue-containing therapies | N/A | 21 (29.6%) | N/A |
| Bortezomib-containing therapies | N/A | 7 (9.9%) | N/A |
| Cyclophosphamide-containing therapies | N/A | 64 (90.1%) | N/A |

| | Population of interest per the NICE scope | Population per the DRC study[10,11] | Population per the BR study[12] |
|---|---|---|---|
| Rituximab alone or in combination therapy | N/A | 55 (77.5%) | N/A |
| Extramedullary disease, n (%) | | | |
| Adenopathy and/or splenomegaly | N/A | 30 (42.3%) | N/A |
| Lymphadenopathy | N/A | NR | 28 (38.9%) |
| Splenomegaly | N/A | NR | 23 (31.9%) |
| IPSSWM score, n (%) | | | |
| Low risk | N/A | 12 (21.4%[a]) | NR |
| Intermediate risk | N/A | 17 (30.4%[a]) | NR |
| High risk | N/A | 27 (48.2%[a]) | NR |

[a] Based on 56 patients.
Abbreviations: BR = bendamustine-rituximab; DRC = dexamethasone-rituximab-cyclophosphamide; IgM = immunoglobulin M; IPSSWM = International Prognostic Scoring System for Waldenström's Macroglobulinaemia; N/A = not applicable; NR = not reported; SD = standard deviation; WM = Waldenström's macroglobulinaemia

For the DRC study, the enrolled patients with WM represented a relatively healthier population who were previously untreated (upon study enrolment) and were suitable for chemo-immunotherapy, compared with the population of interest in the NICE scope, which could potentially bias the comparison against zanubrutinib. On the other hand, the study was conducted in Greece and was initiated more than 10 years ago, which makes it potentially questionable to which extent the study population was representative of the WM population in the UK in more recent years. However, one advantage associated with the limitation of the year of study is that the relatively longer follow-up duration allowed for assessment of long-term outcomes in patients with WM.

For the BR study, although the WM population included in the Italy-based BR study were more comparable to the ASPEN trial and more aligned with the population of interest per the NICE scope (relative to the level of alignment between the DRC study and the NICE scope), uncertainties remain to which extent the study population was representative of recent clinical practice in the UK.

In summary, uncertainties regarding representativeness exist, but the studies included in the MAIC were more likely to be representative of the UK-based WM population compared with the other studies identified from the clinical SLR, as discussed in detail in the Appendix D.1.2 in the CS.

Further assessments of the representativeness were conducted by comparing the reported patient characteristics of the participants enrolled in the two single-arm studies to the participants from the UK WMUK Rory Morrison Registry up to 2018 (a registry with a total of 579 WM patients registered from 19 hospitals across the UK).[3] As shown in Table 10, the populations were relatively comparable, based on the reported estimates of female proportion, age, and IPSSWM (note that the IPSSWM score accounts for 5 key prognostic factors, including age, haemoglobin, platelet, β-2 microglobulin and monoclonal IgM).

**Table 10. Comparison of BR and DRC patient populations with the WMUK Rory Morrison Registry population**

| Baseline patient characteristics | Population per the first UK WM registry report from the WMUK Rory Morrison Registry[3] | Population per the DRC study[10,11] | Population per the BR study[12] |
|---|---|---|---|
| Female proportion | 38% | 38% | 35% |
| Age, year | Median 60–69 | 69 | 72 |
| IPSSWM | 0: 19%<br>1: 15%<br>2: 31%<br>3: 13%<br>4: 17%<br>5: 4%<br>6: 0%<br>(among 253 patients with IPSSWM data) | Not reported | Low (0): 22%<br>Intermediate (1–2): 30%<br>High (>2): 48%<br>(among 56 patients with IPSSWM data) |

Abbreviations: BR = bendamustine-rituximab; DRC = dexamethasone-rituximab-cyclophosphamide; IPSSWM = International Prognostic Scoring System for Waldenström's Macroglobulinaemia; WM = Waldenström's macroglobulinaemia; UK = United Kingdom

**A23. Priority Question: Please explain how PFS was used as an outcome for DRC (dexamethasone, rituximab and cyclophosphamide) when in Table B.5.7 (company submission, Appendix D), PFS is listed as "not estimable" in the Dimopoulos[reference #50] and Kastritis[reference #51] studies, and PFS was not mentioned in the second paragraph of B.2.9.1 for DRC?**

a. **Alternatively, if data are available, please update Table B.5.7 and section B.2.9.1**

The Kaplan-Meier (KM) curve of PFS for DRC reported in Dimopoulos et al. 2007 with shorter follow-up was estimable, due to the availability of the number of patients

at risk at baseline and the markers for censoring.[10] However, the KM curve of PFS reported in Kastritis et al. 2015, despite its longer follow-up, was lacking both the markers for censoring and the number of patients at risk at each time interval.[11] However, it is still estimable based on an extra assumption that there was no censoring until the end of the follow-up. The KM curve of PFS reported in Kastritis et al. 2015 with longer follow-up was used to inform the MAIC and then the cost-effectiveness analysis. The assumption made as to censoring has no impact on the MAIC results and minimal impact on the long-term extrapolation of PFS for DRC used to inform the cost-effectiveness analysis. The impact was also mitigated by the availability of long-term mature survival data for DRC.

**A24. Priority question: Why does table B.5.8 (company submission, Appendix D) have the Overall survival (OS) Kaplan Meier curve listed as "NR" (not reported), when there is an overall survival curve presented in Tedeschi 2015[reference #49], and there appears to be data for overall survival for bendamustine rituximab (BR)?**

This was a typographical error. "NR" should be "Reported".

**A25. Priority question: Please update Table B.2.19 key assumptions of the MAIC (company submission) to provide validation for the following assumptions:**

a. **The outcomes for all trials in the MAIC were measured similarly enough to avoid bias (The Evidence Review Group (ERG) note that this appears to be not true for DRC, which counted deaths from any cause in PFS, unlike BR, where deaths from non-WM causes were censored).**

b. **That all relevant outcomes, including OS and PFS and any other variables that cannot be balanced and matched, are unaffected by year and location**

**of study (which are not matching variables and not balanced between zanubrutinib and its comparators).**

c. **That matching on the dichotomous or categorical form of continuous variables, e.g. IgM concentration (≤40 versus >40 g/L), is sufficient to remove bias from the matched variables.**

For A25b and A25c, an updated Table B.2.19 is shown below in Table 11 with updates in *italics*.

**Table 11. Key assumptions of the MAIC (UPDATED Table B.2.19)**

| Category | Assumption | Justification |
|---|---|---|
| Population (zanubrutinib matched to DRC vs DRC) | The zanubrutinib arm in ASPEN (N=102), including a mix of relapsed/refractory population (N=83) and treatment-naïve population unsuitable for chemo-immunotherapy (N=19), was adjusted to match the DRC population which included only treatment-naïve population suitable for chemo-immunotherapy (N=72), assuming that the discrepancies in patient populations had limited impact on the MAIC results. | This assumption was necessitated by the limited availability of clinical evidence (see **Error! Reference source not found.**).<br>It should be acknowledged that such differences in patient populations might have led to an underestimation of the relative survival benefit of zanubrutinib compared with DRC, assuming that PFS and OS outcomes are more favourable in the treatment-naïve population (suitable for chemo-immunotherapy) compared with both treatment-naïve population (unsuitable for chemo-immunotherapy) and relapsed/refractory population. |
| Matching variables | It was assumed that any unobserved key prognostic factors were well balanced between the zanubrutinib arm and comparator arms, such that the MAIC results were robust with limited biases. | It is rarely possible to completely adjust for all unobserved or unreported baseline patent characteristics, which is a general limitation of a MAIC. Despite that, the outcome comparison was conducted before and after matching adjustment, which consistently showed survival benefit of zanubrutinib compared with the comparators. |
| *Matching variables* | *The categorisation of the continuous variables (e.g., IgM concentration ≤40 versus >40 g/L) does not bias the MAIC result.* | *The categorisation was necessitated given the availability of summary data of baseline patient characteristics from the comparator trial publications where only categorical data was available.*<br><br>*In addition to the comparator trial publications, a supplemental review of studies identified from the clinical SLR was conducted, which showed that the same categorisation was commonly adopted in clinical studies (e.g., Treon et al. 2015 NEJM per the phase 2 trial for ibrutinib; Castillo et al. 2018 Brit J Hematol per a study of BR) and therefore deemed clinically relevant.* |

| Category | Assumption | Justification |
|---|---|---|
| *Matching variables* | *It was assumed that the differences in study characteristics, such as geographic location (Europe, Australia or New Zealand and North America for the ASPEN trial; Greece for the DRC study; Italy for the BR study) and year of study, would not bias the MAIC result.* | *It was infeasible to adjust for these factors in the MAIC.* |

Abbreviations: BR = bendamustine-rituximab; DRC = dexamethasone, rituximab and cyclophosphamide; IgM = immunoglobulin M; MAIC = matching-adjusted indirect comparison; N = number of patients evaluable; OS = overall survival; PFS = progression-free survival; SLR = systematic literature review

For A25a, because it was stated in Tedeschi et al. 2015 (i.e. the BR study), "*OS was calculated from the start of therapy to death from any cause*.", no update was made to the table of assumptions of MAIC. The company also double checked Dimopoulos et al. 2007/Kastritis et al. 2015 (i.e. the DRC study) and found no description of the censoring but judging by the wording of "overall survival", it was assumed that the definition of "event" was "death from any-cause", as typically defined in clinical trials.

**A26. Priority Question: Please repeat all MAIC analyses using any available study with sufficient data for any outcome for any comparator, and not just only the two best studies, however judged.**

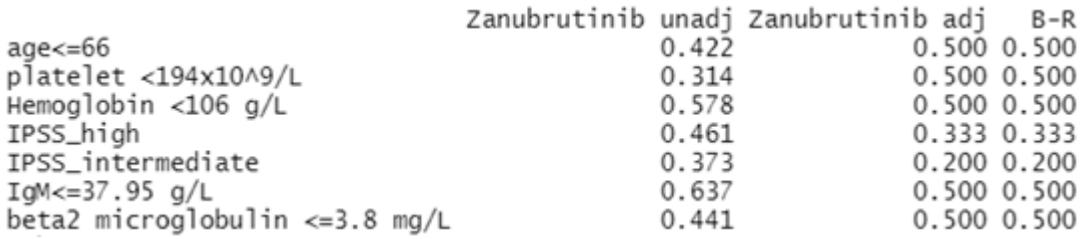
Based on Table B.5.7 in Appendix D of the CS, further feasibility assessments were conducted to determine which of the studies were feasible for inclusion in the MAIC.

Results of the additional MAIC analyses are presented below, with one set of MAIC per row, depending on the data availabilities for the comparator population (e.g. availabilities of survival outcomes in the form of KM curves, availabilities of subgroup-specific outcomes). Limitations of these analyses should be considered when interpreting these analysis results, including but not limited to geographic location, sample size, alignment in population definition, and feasibility for further inclusion in the cost-effectiveness analyses), as described below for each set of MAIC.

1. Adjusting the overall zanubrutinib arm (N=102) to the BR arm (N=60) of Paludo 2018 (including 17 treatment-naïve patients suitable for chemo-

immunotherapies and 43 relapsed/refractory patients treated with BR from a single centre in the US)

**Figure 4. Baseline characteristics (zanubrutinib [overall] vs BR) – Paludo 2018[13]**



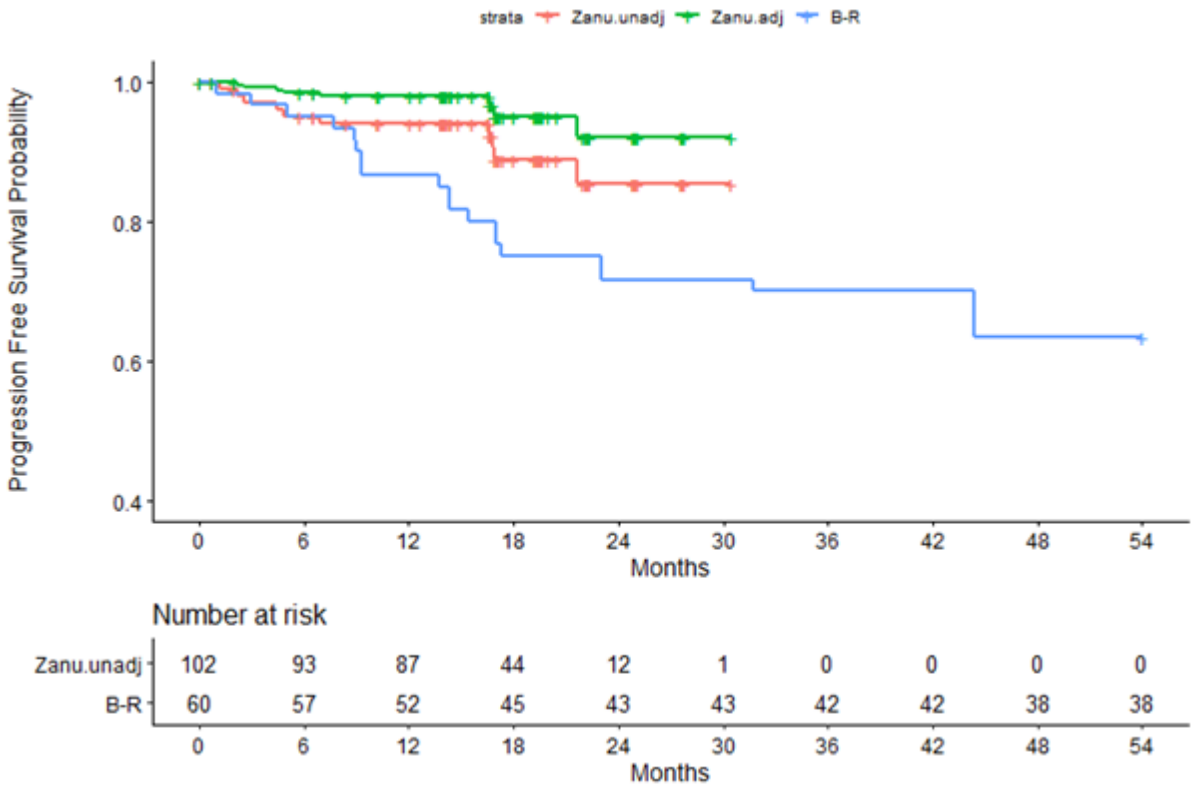

Abbreviations: BR = bendamustine-rituximab; IgM = immunoglobulin M; IPSS = International Prognostic Scoring System

**Table 12. PFS (zanubrutinib [overall] vs BR) – Paludo 2018[13]**

| | Before matching adjustment | | After matching adjustment | |
|---|---|---|---|---|
| | **P-value, log-rank test** | **HR of zanubrutinib vs comparator (95% CI)** | **P-value, log-rank test** | **HR of zanubrutinib vs comparator (95% CI)** |
| PFS | 0.0497 | 0.458 (0.207 1.016) | 0.0050 | 0.200 (0.076, 0.524) |
| OS | Not reported | Not reported | Not reported | Not reported |

Note: Due to a lack of publicly reported number of patients at risk at baseline and the markers for censoring for the comparator arm from the associated study publication, it was no censoring until the end of the follow-up (see response to clarification question A23).
Abbreviations: BR = bendamustine-rituximab; CI = confidence interval; HR = hazard ratio; OS = overall survival; PFS = progression-free survival

**Figure 5. PFS (zanubrutinib [overall] vs BR) – Paludo 2018[13]**




Abbreviations: BR = bendamustine-rituximab; PFS = progression-free survival

2. Adjusting the overall zanubrutinib arm, Relapsed/Refractory Analysis Set (n=83) to the BR arm, relapsed/refractory subgroup (n=43) of Paludo 2018 (including 17 treatment-naïve patients *suitable* for chemo-immunotherapies and 43 R/R patients treated with BR from a single centre in the US)

**Figure 6. Baseline characteristics (zanubrutinib [Relapsed/Refractory Analysis Set] vs BR) – Paludo 2018[13]**

| | Zanubrutinib unadj | Zanubrutinib adj | B-R |
|---|---|---|---|
| age<=66 | 0.458 | 0.500 | 0.500 |
| platelet <194x10^9/L | 0.313 | 0.500 | 0.500 |
| Hemoglobin <106 g/L | 0.530 | 0.500 | 0.500 |
| IPSS_high | 0.446 | 0.333 | 0.333 |
| IPSS_intermediate | 0.361 | 0.200 | 0.200 |
| IgM<=37.95 g/L | 0.651 | 0.500 | 0.500 |
| beta2 microglobulin <=3.8 mg/L | 0.446 | 0.500 | 0.500 |

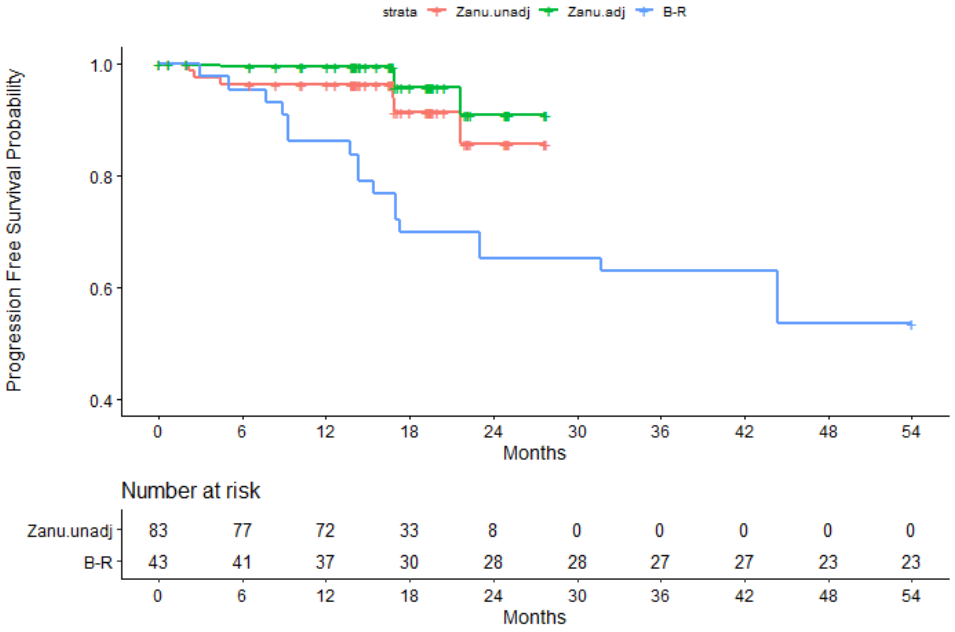
Of note, the patient baseline characteristics are only available for the overall population, whereas the PFS KM is available for the relapsed/refractory subgroup. In order for exploratory analyses, the patient baseline characteristics of the overall population were used for this matching adjustment, assuming that the patient profiles were the same between the relapsed/refractory subgroup (n=43) and the overall population (including 17 treatment-naïve patients suitable for chemo-immunotherapies and 43 relapsed/refractory patients).
Abbreviations: BR = bendamustine-rituximab; IgM = immunoglobulin M; IPSS = International Prognostic Scoring System; KM = Kaplan-Meier; PFS = progression-free survival

**Table 13. PFS (zanubrutinib [Relapsed/Refractory Analysis Set] vs BR) – Paludo 2018[13]**

| | Before matching adjustment | | After matching adjustment | |
|---|---|---|---|---|
| | P-value, log-rank test | HR of zanubrutinib vs comparator (95% CI) | P-value, log-rank test | HR of zanubrutinib vs comparator (95% CI) |
| PFS | 0.009 | 0.294 (0.112, 0.773) | 0.004 | 0.103 (0.028, 0.381) |
| OS | Not reported | Not reported | Not reported | Not reported |

Note: Due to a lack of publicly reported number of patients at risk at baseline and the markers for censoring for the comparator arm from the associated study publication, it was no censoring until the end of the follow-up (see response to clarification question A23).
Abbreviations: BR = bendamustine-rituximab; CI = confidence interval; HR = hazard ratio; OS = overall survival; PFS = progression-free survival

**Figure 7. PFS (zanubrutinib [Relapsed/Refractory Analysis Set] vs BR) – Paludo 2018[13]**




Abbreviations: BR = bendamustine-rituximab; PFS = progression-free survival

3. Adjusting the overall zanubrutinib arm (N=102) to the BR arm (N=57) of Castillo 2018 (including 57 treatment-naïve patients *suitable* for chemo-immunotherapies from a single centre in the US)

Of note, in Table B.5.7 in the Appendix D of the CS, “treatment line not reported” was specified. However, based on the description of the patient population below, it is likely that these were treatment-naïve patients *suitable* for chemo-immunotherapies:

“*We searched our database for WM patients who received primary therapy with Benda-R, BDR or CDR between January 2005 and December 2016. All patients met diagnostic criteria for WM and criteria for treatment initiation based on the*

*recommendations made by the 2nd International Workshop for WM (IWWM) (Kyle et al, 2003; Owen et al, 2003).”*[14]

The company contacted the author, Dr. Castillo, who confirmed that “in our paper, all the patients were in the frontline setting and all patients were fit enough to be good candidates for chemo-immunotherapy. I think it would not be fair to compare a mixed pool of patients (treatment naïve and relapsed/refractory) treated with zanubrutinib versus a purely treatment-naïve group.”

**Figure 8. Baseline characteristics (zanubrutinib vs BR) – Castillo 2018**[14]

| | Zanubrutinib unadj | Zanubrutinib adj | B-R |
|---|---|---|---|
| age<=65 | 0.402 | 0.368 | 0.368 |
| platelet <100x10^9/L | 0.118 | 0.158 | 0.158 |
| Hemoglobin <115 g/L | 0.706 | 0.404 | 0.404 |
| IPSS_high | 0.461 | 0.351 | 0.351 |
| IPSS_intermediate | 0.373 | 0.298 | 0.298 |
| IgM<=40 g/L | 0.657 | 0.632 | 0.632 |
| beta2 microglobulin <=3 mg/L | 0.265 | 0.316 | 0.316 |

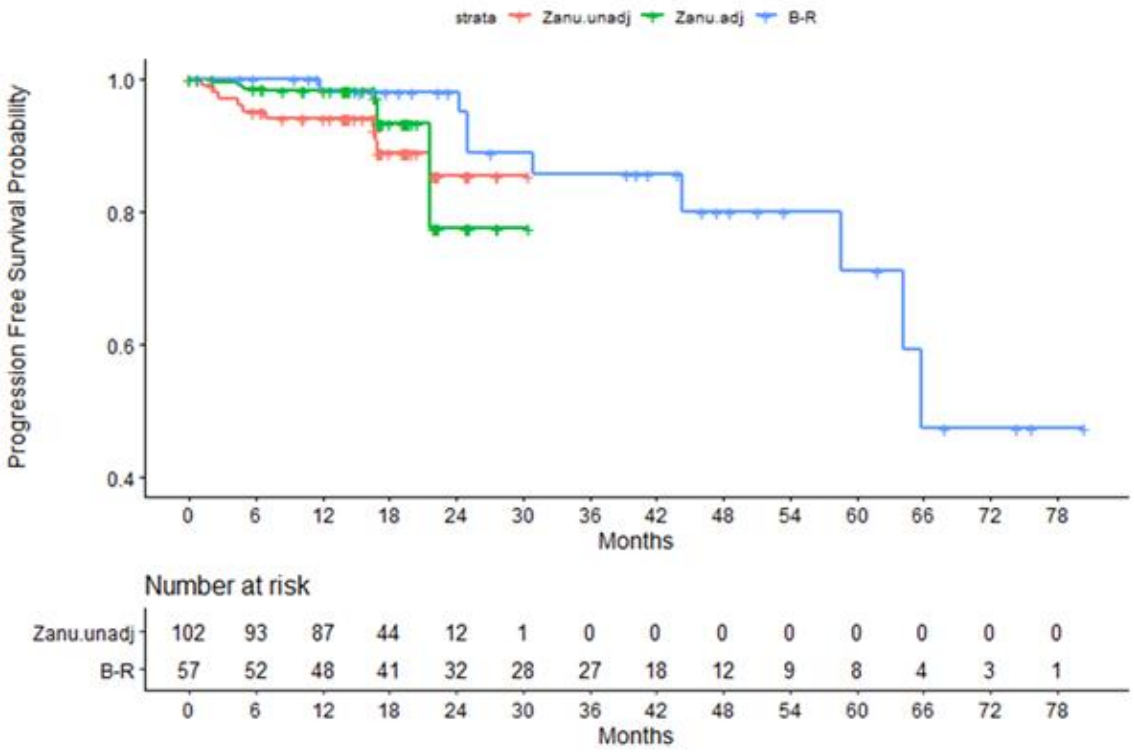
Abbreviations: BR = bendamustine-rituximab; IgM = immunoglobulin M; IPSS = International Prognostic Scoring System

**Table 14. PFS and OS (zanubrutinib vs BR) – Castillo 2018**[14]

| | Before matching adjustment | | After matching adjustment | |
|---|---|---|---|---|
| | **P-value, log-rank test** | **HR of zanubrutinib vs comparator (95% CI)** | **P-value, log-rank test** | **HR of zanubrutinib vs comparator (95% CI)** |
| PFS | 0.109 | 2.680 (0.774, 9.277) | 0.202 | 3.690 (0.797, 17.079) |
| OS | 0.232 | 2.657 (0.508, 13.884) | 0.208 | 4.429 (0.678, 28.906) |

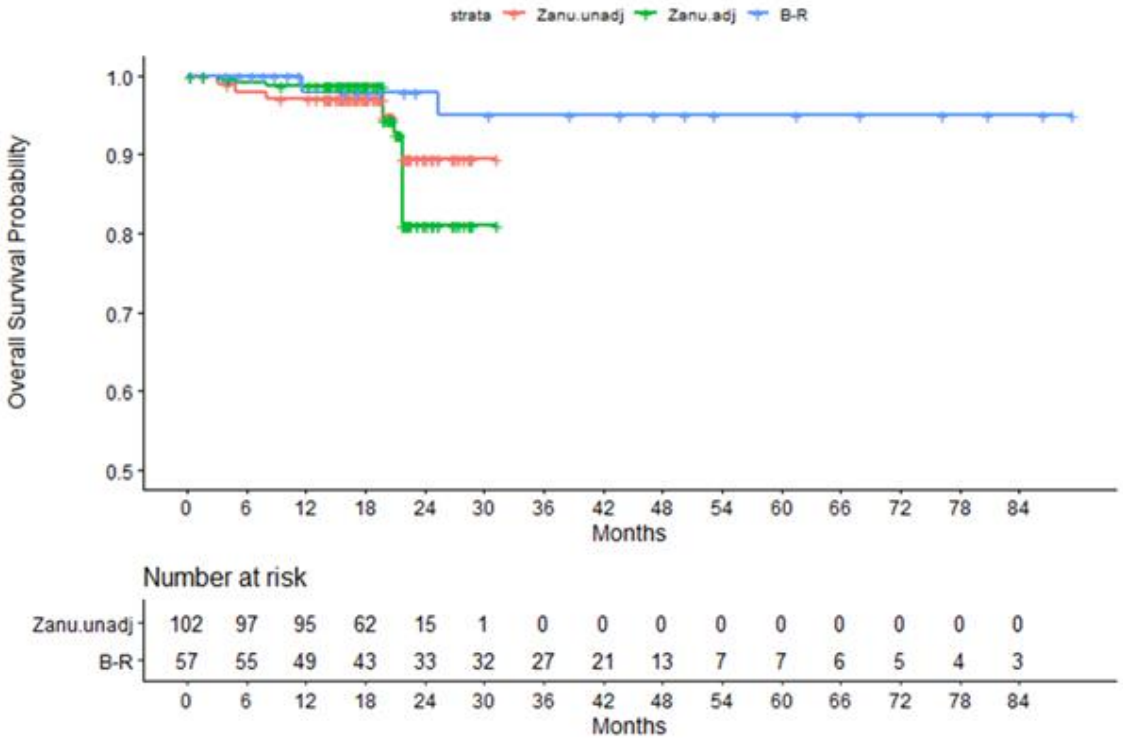
Abbreviations: BR = bendamustine-rituximab; CI = confidence interval; HR = hazard ratio; OS = overall survival; PFS = progression-free survival

**Figure 9. PFS (zanubrutinib vs BR) – Castillo 2018[14]**




Abbreviations: BR = bendamustine-rituximab; PFS = progression-free survival

**Figure 10. OS (zanubrutinib vs BR) – Castillo 2018[14]**




Abbreviations: BR = bendamustine-rituximab; OS = overall survival

4. Adjusting the overall zanubrutinib arm (N=102) to the DRC arm (N=100) of Paludo 2017/2018 (including 50 treatment-naïve patients *suitable* for chemo-immunotherapies and 50 R/R patients from either a single centre or multiple centres in the US)

**Figure 11. Baseline characteristics (zanubrutinib [overall] vs DRC) – Paludo 2017/2018[13,15]**

| | Zanubrutinib unadj | Zanubrutinib adj | B-R |
|---|---|---|---|
| age<=68 | 0.451 | 0.500 | 0.500 |
| platelet <196x10^9/L | 0.324 | 0.500 | 0.500 |
| Hemoglobin <105 g/L | 0.569 | 0.500 | 0.500 |
| IPSS_high | 0.461 | 0.550 | 0.550 |
| IPSS_intermediate | 0.373 | 0.263 | 0.262 |
| IgM<=40.85 g/L | 0.657 | 0.500 | 0.500 |
| beta2 microglobulin <=3.4 mg/L | 0.353 | 0.500 | 0.500 |

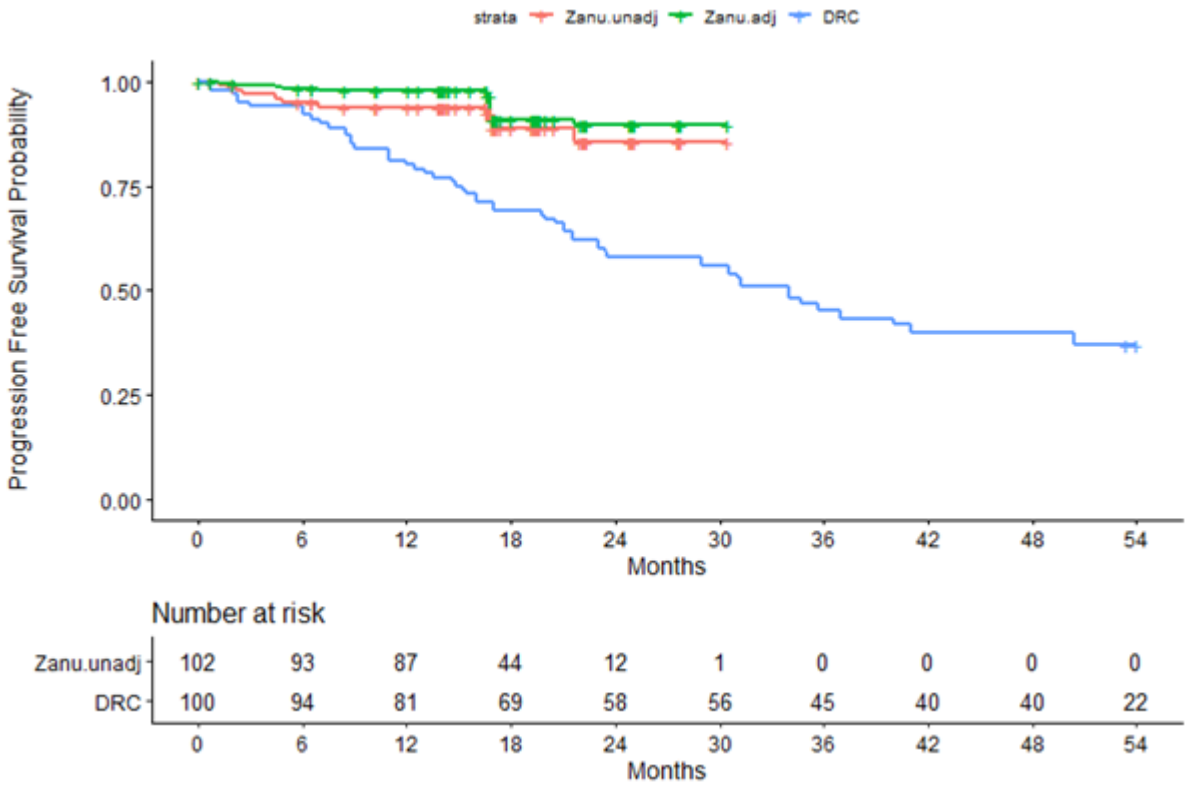
Note: The heading of "B-R" was a typographic error, which should be "DRC".
Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; IgM = immunoglobulin M; IPSS = International Prognostic Scoring System

**Table 15. PFS (zanubrutinib [overall] vs DRC) – Paludo 2017/2018[13,15]**

| | Before matching adjustment | | After matching adjustment | |
|---|---|---|---|---|
| | **P-value, log-rank test** | **HR of zanubrutinib vs comparator (95% CI)** | **P-value, log-rank test** | **HR of zanubrutinib vs comparator (95% CI)** |
| PFS | 0.0003 | 0.298 (0.148, 0.599) | <0.0001 | 0.198 (0.077, 0.512) |
| OS | Not reported | Not reported | Not reported | Not reported |

Note: Due to a lack of publicly reported number of patients at risk at baseline and the markers for censoring for the comparator arm from the associated study publication, it was no censoring until the end of the follow-up (see response to clarification question A23).
Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; CI = confidence interval; HR = hazard ratio; OS = overall survival; PFS = progression-free survival

**Figure 12. PFS (zanubrutinib [overall] vs DRC) – Paludo 2017/2018[13,15]**




Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; PFS = progression-free survival

5. Adjusting the zanubrutinib arm, Relapsed/Refractory Analysis Set (N=83) to the DRC arm, relapsed/refractory subgroup (n=50) of Paludo 2017/2018 (including 50 treatment-naïve patients *suitable* for chemo-immunotherapies and 50 relapsed/refractory patients from either a single centre or multiple centres in the US)

**Figure 13. Baseline characteristics (zanubrutinib [Relapsed/Refractory Analysis Set] vs DRC) – Paludo 2017/2018[13,15]**

| | Zanubrutinib unadj | Zanubrutinib adj | DRC |
|---|---|---|---|
| age<=68 | 0.494 | 0.500 | 0.500 |
| platelet <196x10^9/L | 0.325 | 0.500 | 0.500 |
| Hemoglobin <105 g/L | 0.518 | 0.500 | 0.500 |
| IPSS_high | 0.446 | 0.550 | 0.550 |
| IPSS_intermediate | 0.361 | 0.263 | 0.262 |
| IgM<=40.85 g/L | 0.675 | 0.500 | 0.500 |
| beta2 microglobulin <=3.4 mg/L | 0.361 | 0.500 | 0.500 |

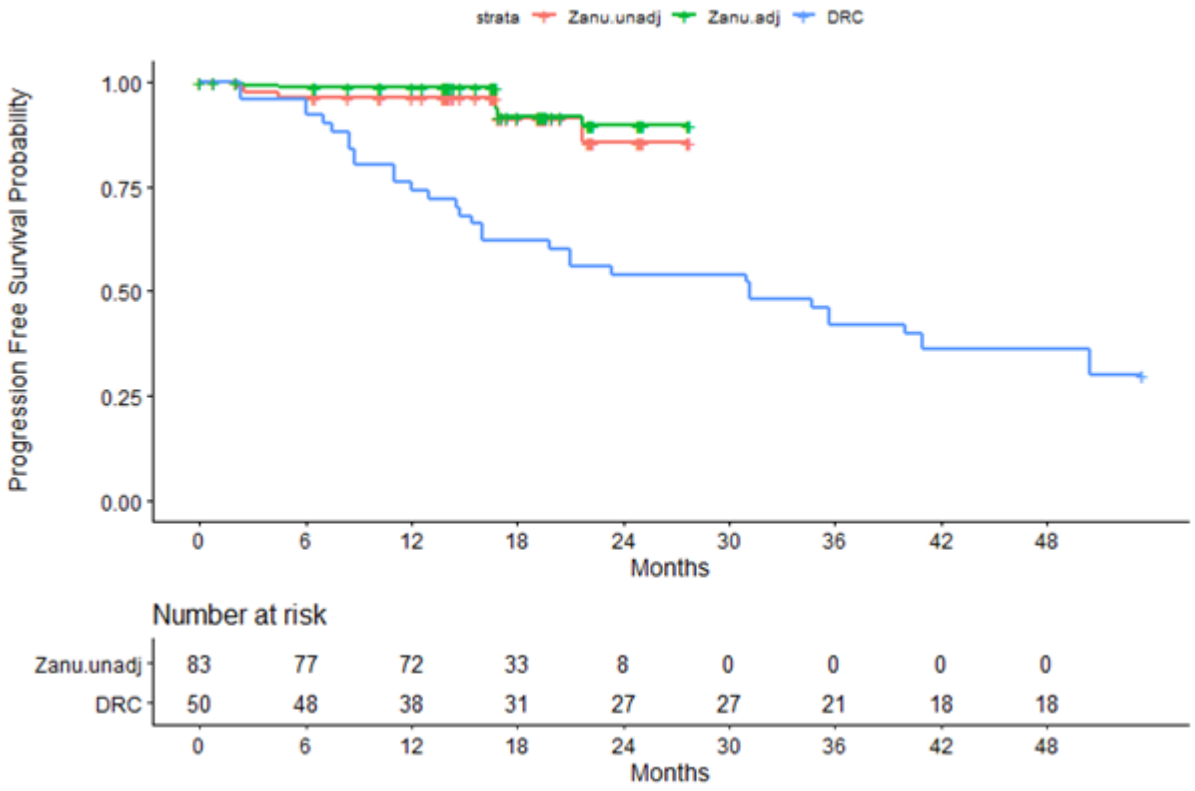
Of note, the patient baseline characteristics are only available for the overall population, whereas the PFS KM is available for the relapsed/refractory subgroup. In order for exploratory analyses, the patient baseline characteristics of the overall population were used for this matching adjustment, assuming that the patient profile were the same between the relapsed/refractory subgroup (n=50) and the overall population (including 50 treatment-naïve patients suitable for chemo-immunotherapies and 50 relapsed/refractory patients).
Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; IgM = immunoglobulin M; IPSS = International Prognostic Scoring System; KM = Kaplan-Meier; PFS = progression-free survival

**Table 16. PFS (zanubrutinib [Relapsed/Refractory Analysis Set] vs DRC) – Paludo 2017/2018[13,15]**

| | Before matching adjustment | | After matching adjustment | |
|---|---|---|---|---|
| | P-value, log-rank test | HR of zanubrutinib vs comparator (95% CI) | P-value, log-rank test | HR of zanubrutinib vs comparator (95% CI) |
| PFS | <0.0001 | 0.183 (0.078, 0.430) | <0.0001 | 0.134 (0.044, 0.413) |
| OS | Not reported | | | |

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; CI = confidence interval; HR = hazard ratio; OS = overall survival; PFS = progression-free survival

**Figure 14. PFS (zanubrutinib [Relapsed/Refractory Analysis Set] vs DRC) – Paludo 2017/2018[13,15]**




Note: Due to a lack of publicly reported number of patients at risk at baseline and the markers for censoring for the comparator arm from the associated study publication, it was no censoring until the end of the follow-up (see response to clarification question A23).

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; PFS = progression-free survival

6. Adjusting the overall zanubrutinib arm (N=102) to the DRC arm (N=38) of Castillo 2018 (including 38 treatment-naïve patients *suitable* for chemo-immunotherapies from a single centre in the US)

Of note, in Table B.5.7 in the Appendix D of the company submission, “treatment line not reported” was specified. However, based on the description of the patient

population below, it is likely that these were treatment-naïve patients *suitable* for chemo-immunotherapies.

"*We searched our database for WM patients who received primary therapy with Benda-R, BDR or CDR between January 2005 and December 2016. All patients met diagnostic criteria for WM and criteria for treatment initiation based on the recommendations made by the 2nd International Workshop for WM (IWWM) (Kyle et al, 2003; Owen et al, 2003).*"[16]

The company contacted the author, Dr. Castillo, who confirmed that "in our paper, all the patients were in the frontline setting and all patients were fit enough to be good candidates for chemo-immunotherapy. I think it would not be fair to compare a mixed pool of patients (treatment naïve and relapsed/refractory) treated with zanubrutinib versus a purely treatment-naïve group."

**Figure 15. Baseline characteristics (zanubrutinib vs DRC) – Castillo 2018[14]**

| | Zanubrutinib unadj | Zanubrutinib adj | DRC |
|---|---|---|---|
| age<=65 | 0.402 | 0.500 | 0.500 |
| platelet <100x10^9/L | 0.118 | 0.053 | 0.053 |
| Hemoglobin <115 g/L | 0.706 | 0.263 | 0.263 |
| IgM<=40 g/L | 0.657 | 0.605 | 0.605 |
| beta2 microglobulin <=3 mg/L | 0.265 | 0.579 | 0.579 |

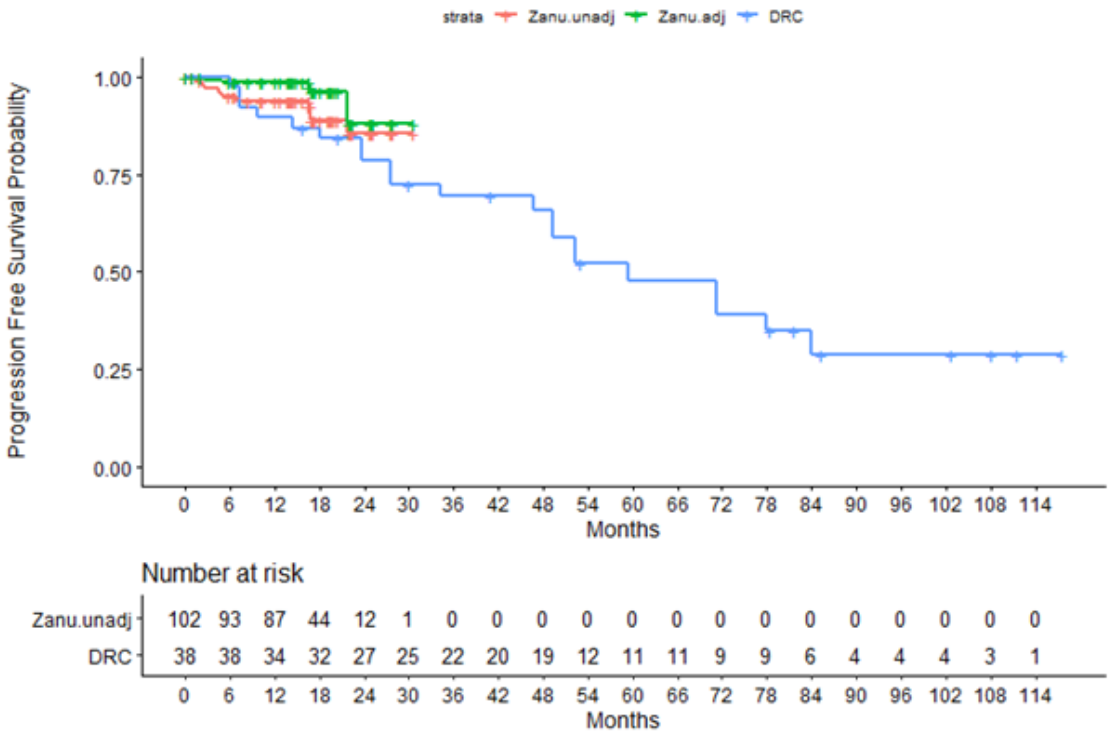
Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; IgM = immunoglobulin M; IPSS = International Prognostic Scoring System

**Table 17. PFS and OS (zanubrutinib vs DRC) – Castillo 2018[14]**

| | Before matching adjustment | | After matching adjustment | |
|---|---|---|---|---|
| | **P-value, log-rank test** | **HR of zanubrutinib vs comparator (95% CI)** | **P-value, log-rank test** | **HR of zanubrutinib vs comparator (95% CI)** |
| PFS | 0.427 | 0.681 (0.265, 1.751) | 0.131 | 0.328 (0.088, 1.221) |
| OS | 0.545 | 1.645 (0.324, 8.361) | 0.936 | 0.905 (0.083, 9.916) |

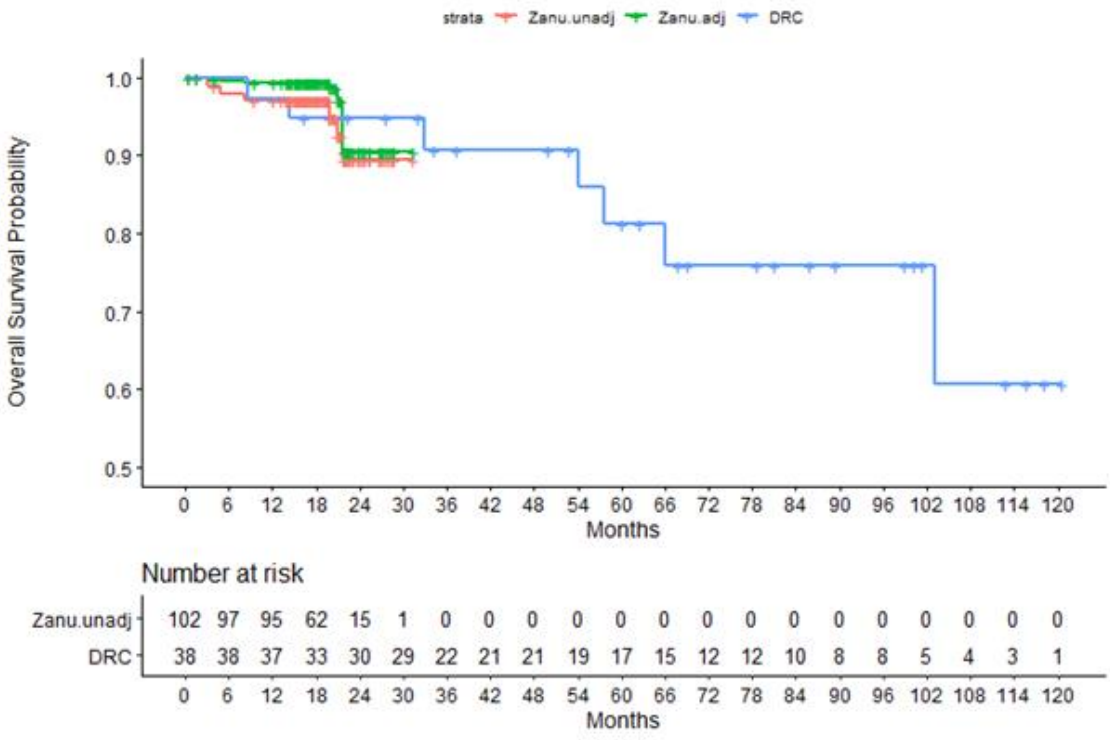
Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; CI = confidence interval; HR = hazard ratio; OS = overall survival; PFS = progression-free survival

**Figure 16. PFS (zanubrutinib vs DRC) – Castillo 2018[14]**




Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; PFS = progression-free survival

**Figure 17. OS (zanubrutinib vs DRC) – Castillo 2018[14]**




Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; PFS = progression-free survival

7. Adjusting the overall zanubrutinib arm (N=102) to the FCR/FR arm (N=43) of Treon 2009 (including 27 treatment-naïve patients *suitable* for chemo-immunotherapies and 16 relapsed/refractory patients from a multi-national prospective study)

**Figure 18. Baseline characteristics (zanubrutinib vs FCR/FR) – Treon 2009[17]**

| | Zanubrutinib unadj | Zanubrutinib adj | FCR/FR |
|---|---|---|---|
| age<=61 | 0.225 | 0.500 | 0.500 |
| platelet <100x10^9/L | 0.118 | 0.093 | 0.093 |
| No prior therapy | 0.186 | 0.628 | 0.628 |
| IgM<=30 g/L | 0.480 | 0.372 | 0.372 |

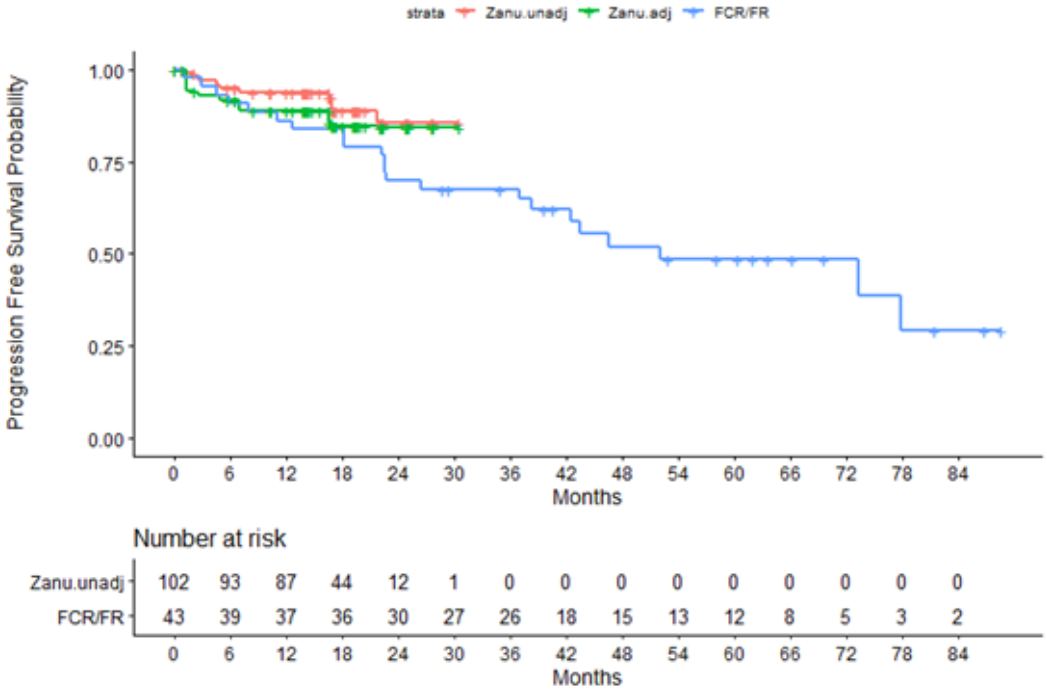
Abbreviations: FCR = fludarabine, cyclophosphamide and rituximab; FR = fludarabine and rituximab; IgM = immunoglobulin M

**Table 18. PFS (zanubrutinib vs FCR/FR) – Treon 2009[17]**

| | Before matching adjustment | | After matching adjustment | |
|---|---|---|---|---|
| | **P-value, log-rank test** | **HR of zanubrutinib vs comparator (95% CI)** | **P-value, log-rank test** | **HR of zanubrutinib vs comparator (95% CI)** |
| PFS | 0.102 | 0.494 (0.210, 1.163) | 0.314 | 0.588 (0.194, 1.785) |
| OS | Not reported | | | |

Abbreviations: CI = confidence interval; FCR = fludarabine, cyclophosphamide and rituximab; FR = fludarabine and rituximab; HR = hazard ratio; OS = overall survival; PFS = progression-free survival

**Figure 19. PFS (zanubrutinib vs FCR/FR) – Treon 2009[17]**




Abbreviations: FCR = fludarabine, cyclophosphamide and rituximab; FR = fludarabine and rituximab; PFS = progression-free survival

8. Adjusting the overall zanubrutinib arm (N=102) to the FCR/FR arm (N=43) of Tedeschi 2012 (including 28 treatment-naïve patients suitable for chemo-immunotherapies and 15 relapsed/refractory patients from multiple centres in an Italy-based prospective study)

**Figure 20. Baseline characteristics (zanubrutinib vs FCR/FR) – Tedeschi 2012[18]**

| | Zanubrutinib unadj | Zanubrutinib adj | FCR |
|---|---|---|---|
| age<=65 | 0.402 | 0.500 | 0.500 |
| platelet <100x10^9/L | 0.118 | 0.023 | 0.023 |
| Hemoglobin <100 g/L | 0.471 | 0.500 | 0.500 |
| IPSS_high | 0.461 | 0.279 | 0.279 |
| IPSS_intermediate | 0.373 | 0.372 | 0.372 |
| IgM<=43.8 g/L | 0.706 | 0.500 | 0.500 |
| Adenopathy/splenomegaly/extranodal involvement | 0.618 | 0.581 | 0.581 |

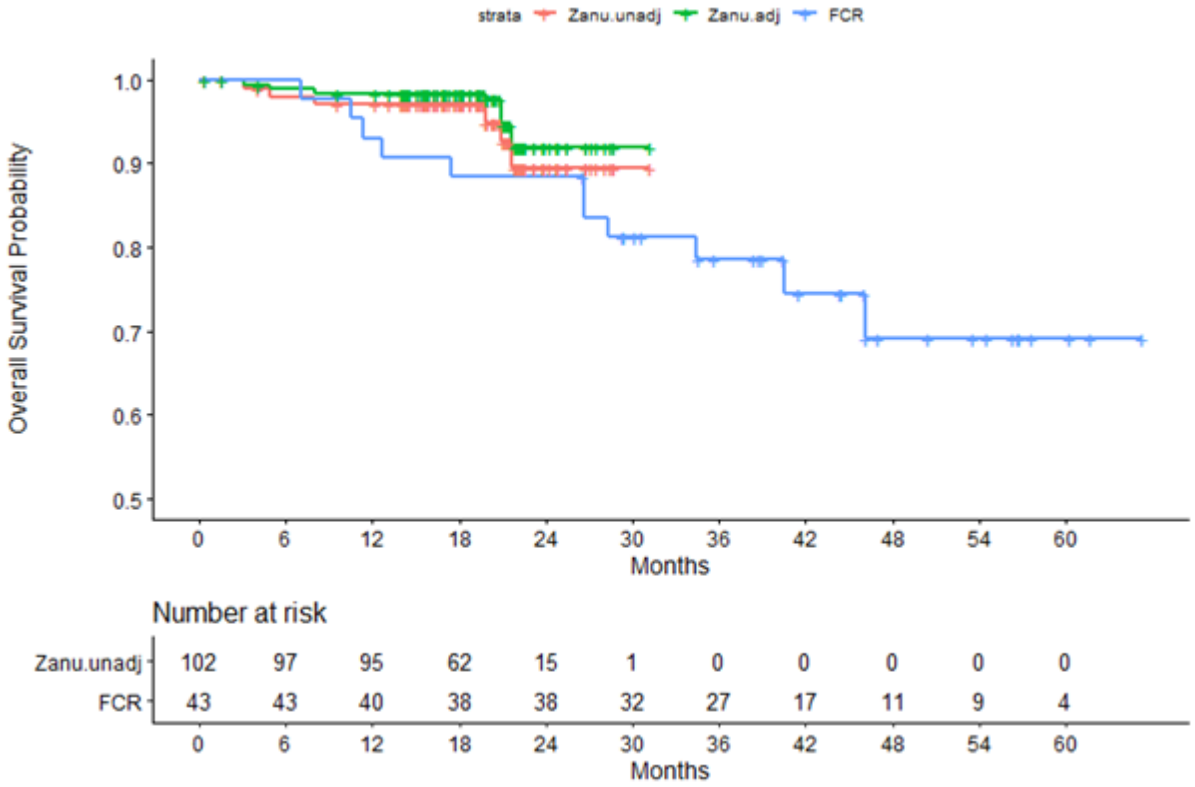
Abbreviations: FCR = fludarabine, cyclophosphamide and rituximab; FR = fludarabine and rituximab; IgM = immunoglobulin M; IPSS = International Prognostic Scoring System

**Table 19. OS (zanubrutinib vs FCR/FR) – Tedeschi 2012[18]**

| | Before matching adjustment | | After matching adjustment | |
|---|---|---|---|---|
| | **P-value, log-rank test** | **HR of zanubrutinib vs comparator (95% CI)** | **P-value, log-rank test** | **HR of zanubrutinib vs comparator (95% CI)** |
| PFS | Not reported | | | |
| OS | 0.374 | 0.598 (0.191, 1.834) | 0.204 | 0.397 (0.105, 1.504) |

Abbreviations: CI = confidence interval; FCR = fludarabine, cyclophosphamide and rituximab; FR = fludarabine and rituximab; HR = hazard ratio; OS = overall survival; PFS = progression-free survival

**Figure 21. OS (zanubrutinib vs FCR/FR) – Tedeschi 2012[18]**




Abbreviations: FCR = fludarabine, cyclophosphamide and rituximab; FR = fludarabine and rituximab; OS = overall survival;

9. Adjusting the overall zanubrutinib arm (N=102) to the FCR/FR arm (N=82) of Souchet 2016 (including 25 treatment-naïve patients suitable for chemo-immunotherapies and 57 relapsed/refractory patients from multiple centres in France)

**Figure 22. Baseline characteristics (zanubrutinib vs FCR/FR) – Souchet 2016[19]**

| | Zanubrutinib unadj | Zanubrutinib adj | FCR |
|---|---|---|---|
| age<=61.3 | 0.225 | 0.5 | 0.5 |
| platelet <179.5x10^9/L | 0.294 | 0.5 | 0.5 |
| Hemoglobin <95 g/L | 0.333 | 0.5 | 0.5 |
| 0-1 prior therapy | 0.647 | 0.5 | 0.5 |
| beta2 microglobulin <=3.26 mg/L | 0.284 | 0.5 | 0.5 |

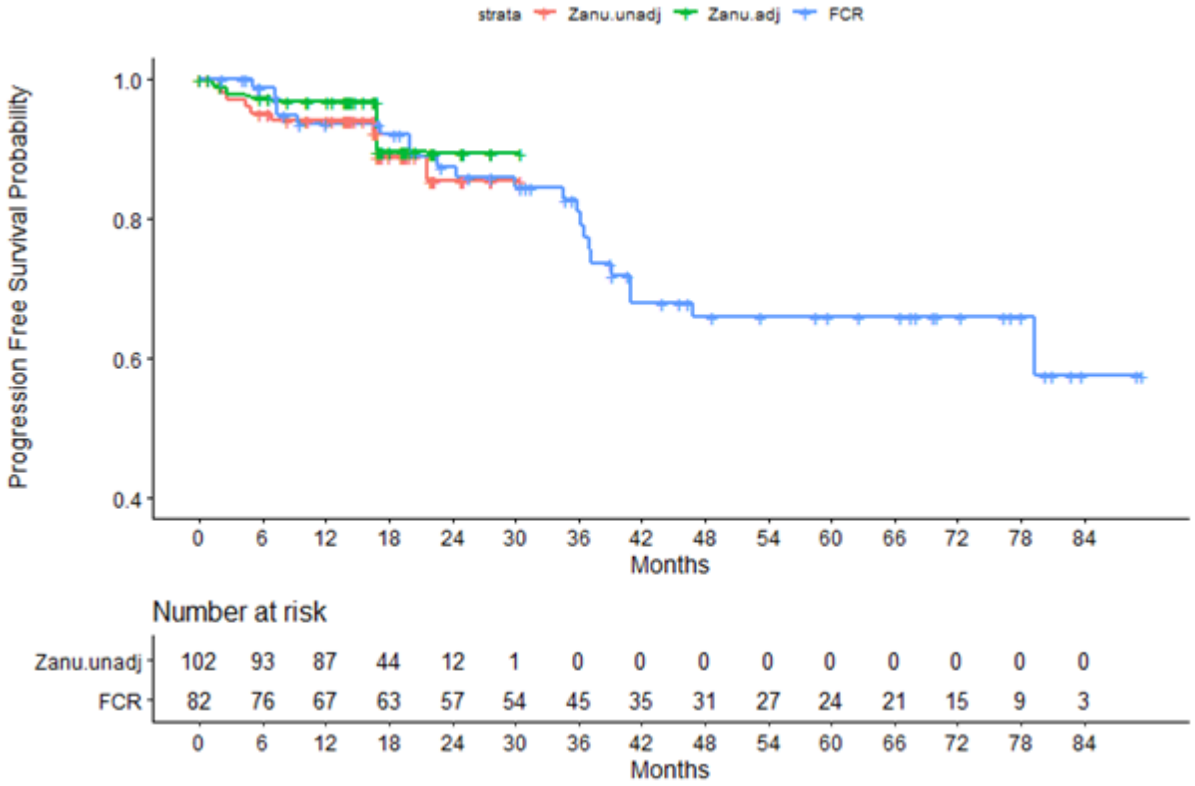
Note: The “0.5” were all exactly 41/82.
Abbreviations: FCR = fludarabine, cyclophosphamide and rituximab; FR = fludarabine and rituximab

**Table 20. PFS (zanubrutinib vs FCR/FR) – Souchet 2016[19]**

| | Before matching adjustment | | After matching adjustment | |
|---|---|---|---|---|
| | P-value, log-rank test | HR of zanubrutinib vs comparator (95% CI) | P-value, log-rank test | HR of zanubrutinib vs comparator (95% CI) |
| PFS | 0.672 | 1.215 (0.490, 3.010) | 0.743 | 0.811 (0.225, 2.919) |
| OS | Not reported | | | |

Abbreviations: CI = confidence interval; FCR = fludarabine, cyclophosphamide and rituximab; FR = fludarabine and rituximab; HR = hazard ratio; OS = overall survival; PFS = progression-free survival

**Figure 23. PFS (zanubrutinib vs FCR/FR) – Souchet 2016[19]**




Abbreviations: FCR = fludarabine, cyclophosphamide and rituximab; FR = fludarabine and rituximab; PFS = progression-free survival

10. Adjusting the overall zanubrutinib arm, Relapsed/Refractory Analysis Set (n=83) to the FCR/FR arm, relapsed/refractory subgroup (n=57) of Souchet 2016 (including 25 treatment-naïve patients suitable for chemo-immunotherapies and 57 relapsed/refractory patients from multiple centres in France)

**Figure 24. Baseline characteristics (zanubrutinib [Relapsed/Refractory Analysis Set] vs FCR/FR) – Souchet 2016[19]**

| | Zanubrutinib unadj | Zanubrutinib adj | FCR |
|---|---|---|---|
| age<=61.3 | 0.229 | 0.5 | 0.5 |
| platelet <179.5x10^9/L | 0.289 | 0.5 | 0.5 |
| Hemoglobin <95 g/L | 0.289 | 0.5 | 0.5 |
| 0-1 prior therapy | 0.566 | 0.5 | 0.5 |
| beta2 microglobulin <=3.26 mg/L | 0.277 | 0.5 | 0.5 |

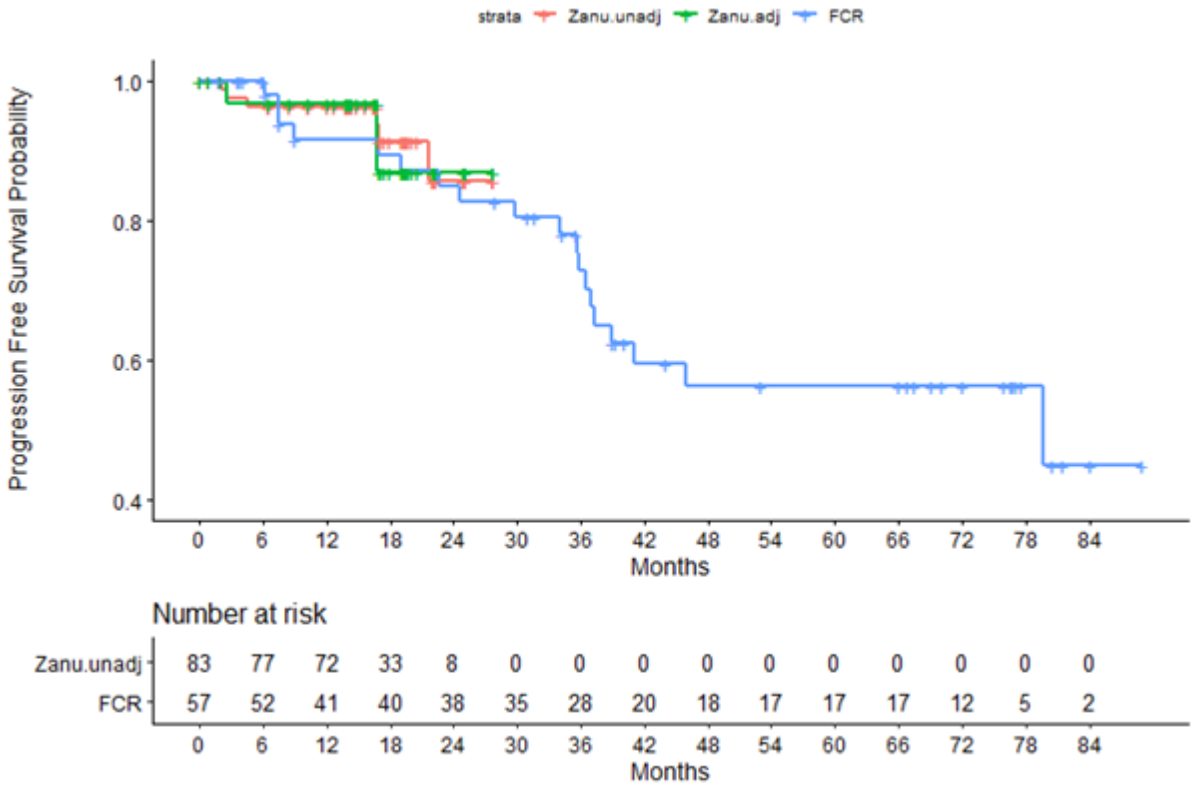
Note: The "0.5" were all exactly 41/82. Of note, the patient baseline characteristics are only available for the overall population, whereas the PFS KM is available for the relapsed/refractory subgroup. In order for exploratory analyses, the patient baseline characteristics of the overall population were used for this matching adjustment, assuming that the patient profile were the same between the relapsed/refractory subgroup (n=57) and the overall population (including 25 treatment-naïve patients suitable for chemo-immunotherapies and 57 relapsed/refractory patients).
Abbreviations: FCR = fludarabine, cyclophosphamide and rituximab; FR = fludarabine and rituximab

**Table 21. PFS (zanubrutinib [Relapsed/Refractory Analysis Set] vs FCR/FR) – Souchet 2016[19]**

| | Before matching adjustment | | After matching adjustment | |
|---|---|---|---|---|
| | P-value, log-rank test | HR of zanubrutinib vs comparator (95% CI) | P-value, log-rank test | HR of zanubrutinib vs comparator (95% CI) |
| PFS | 0.666 | 0.783 (0.258, 2.372) | 0.722 | 0.744 (0.135, 4.101) |
| OS | Not reported | | | |

Abbreviations: CI = confidence interval; FCR = fludarabine, cyclophosphamide and rituximab; FR = fludarabine and rituximab; HR = hazard ratio; OS = overall survival; PFS = progression-free survival

**Figure 25. PFS (zanubrutinib [Relapsed/Refractory Analysis Set] vs FCR/FR) – Souchet 2016[19]**




Abbreviations: FCR = fludarabine, cyclophosphamide and rituximab; FR = fludarabine and rituximab; PFS = progression-free survival

11. Adjusting the overall zanubrutinib arm (N=102) to the chlorambucil arm (N=46) of Kyle 2000 (including 46 patients with unknown prior treatment history, likely from a single centre in the US)

**Figure 26. Baseline characteristics (zanubrutinib vs chlorambucil) – Kyle 2000[20]**

| | Zanubrutinib unadj | Zanubrutinib adj | Chlorambucil |
|---|---|---|---|
| age<=63 | 0.304 | 0.50 | 0.50 |
| platelet <211x10^9/L | 0.392 | 0.50 | 0.50 |
| Hemoglobin <99 g/L | 0.451 | 0.50 | 0.50 |
| Lymphadenopathy | 0.598 | 0.15 | 0.15 |

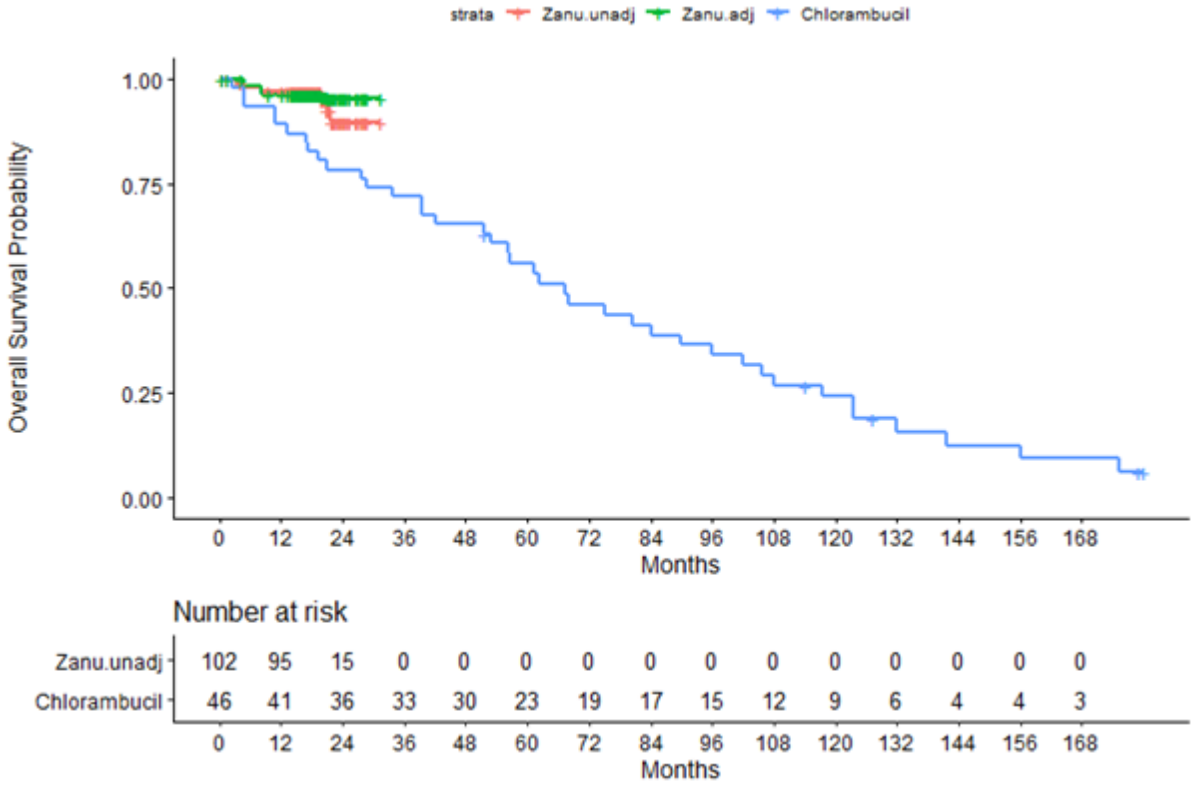
Note: The “0.50” were all exactly 23 out of 46, whereas the “0.15” were exactly “0.150” as directly reported in the trial publication with a lack of reporting of n, N, or missing data.

**Table 22. OS (zanubrutinib vs chlorambucil) – Kyle 2000[20]**

| | Before matching adjustment | | After matching adjustment | |
|---|---|---|---|---|
| | P-value, log-rank test | HR of zanubrutinib vs comparator (95% CI) | P-value, log-rank test | HR of zanubrutinib vs comparator (95% CI) |
| OS | Not reported | | | |
| PFS | 0.029 | 0.336 (0.121, 0.931) | 0.017 | 0.206 (0.055, 0.775) |

Abbreviations: CI = confidence interval; HR = hazard ratio; OS = overall survival; PFS = progression-free survival

**Figure 27. OS (zanubrutinib vs chlorambucil) – Kyle 2000[20]**




Abbreviation: OS = overall survival

12. Adjusting the overall zanubrutinib arm (N=102) to the rituximab monotherapy arm (N=69) of Gertz 2004/2009 (including 34 treatment-naïve patients and 35 relapsed/refractory patients from multiple centres in the US)

Of note, Gertz 2004 and Gertz 2009 were the same study. It is unknown whether the 34 treatment-naïve patients were suitable or unsuitable for chemo-immunotherapies.

**Figure 28. Baseline characteristics (zanubrutinib vs rituximab monotherapy) – Gertz 2004/2009[21,22]**

| | Zanubrutinib unadj | Zanubrutinib adj | Chlorambucil |
|---|---|---|---|
| age<70 | 0.471 | 0.551 | 0.551 |
| platelet <197.5x10^9/L | 0.343 | 0.500 | 0.500 |
| Hemoglobin <96 g/L | 0.373 | 0.500 | 0.500 |
| IgM<=44 g/L | 0.706 | 0.500 | 0.500 |
| beta2 microglobulin <=3.5 mg/L | 0.382 | 0.500 | 0.500 |
| ECOG=0-1 | 0.941 | 0.899 | 0.899 |

Abbreviations: ECOG = Eastern Cooperative Oncology Group; IgM = immunoglobulin M

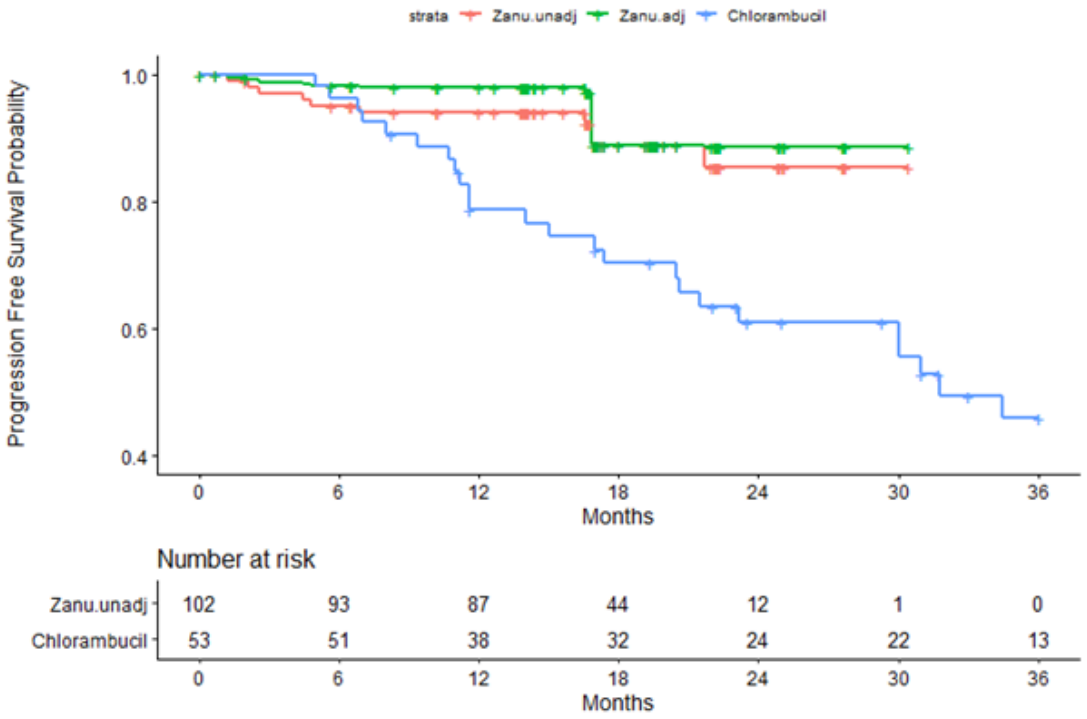
**Table 23. PFS and OS (zanubrutinib vs rituximab monotherapy) – Gertz 2004/2009[21,22]**

| | **Before matching adjustment** | | **After matching adjustment** | |
|---|---|---|---|---|
| | **P-value, log-rank test** | **HR of zanubrutinib vs comparator (95% CI)** | **P-value, log-rank test** | **HR of zanubrutinib vs comparator (95% CI)** |
| PFS | 0.003 | 0.332 (0.153, 0.719) | 0.004 | 0.237 (0.071, 0.793) |
| OS | 0.447 | 0.665 (0.232, 1.909) | 0.036 | 0.232 (0.070, 0.763) |

Notes: As shown in the KM curves below with the number of patients at risk, the PFS and OS KM curves reported in the trial publication covered 53 and 67 patients, respectively, out of the overall 69 patients with evaluable response status and outcomes. It was unknow to which extent the results would be biased, based on the publicly available information.

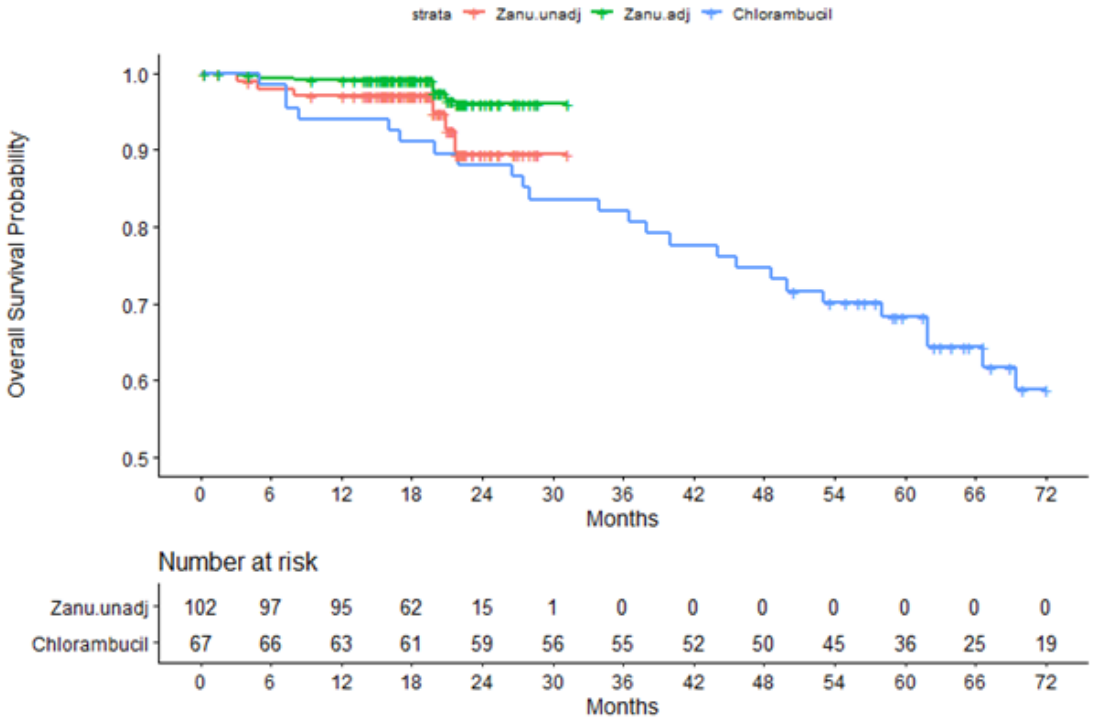
Abbreviations: CI = confidence interval; HR = hazard ratio; KM = Kaplan-Meier; OS = overall survival; PFS = progression-free survival

**Figure 29. PFS (zanubrutinib vs rituximab monotherapy) – Gertz 2004/2009[21,22]**




Abbreviation: PFS = progression-free survival

**Figure 30. OS (zanubrutinib vs rituximab monotherapy) – Gertz 2004/2009[21,22]**




Abbreviation: OS = overall survival

Of note, not all the studies presented in Table B.5.7 in Appendix D of the initial company submission were included for the MAIC. Tam 2005 and Ngan 2003 for FCR/FR were not included due to the extremely low sample size of the WM population (3 and 5, respectively). Tedeschi 2013 for FCR/FR was not included due to a lack of survival outcomes for the overall population but only for treatment responders. Ngan 2003 for FCR/FR was not included due to the lack of reporting of any survival KM curves. Dimopoulos 2002a, Dimopoulos 2002b, Byrd 1999, and Treon 2001 for rituximab monotherapy were also excluded due to the lack of reporting of any survival KM curves.

**A27. Priority Question: Please provide the individual participant data from ASPEN and all relevant code to recreate the MAIC analyses.**

With regard to the programming statistical codes, please refer to the R codes for the MAIC submitted along with the response. However, the ASPEN trial is ongoing and the subject of multiple parallel, international regulatory submissions. Consistent with BeiGene company policy, as aligned with the International Committee of Medical Journal Editors data transparency guidelines, individual de-identified participant data from BeiGene-sponsored clinical studies can only be provided upon regulatory authority request; for indications that have been approved; or in programmes that have been terminated.

# Section B: Clarification on cost-effectiveness data

## *Intervention and comparators*

**B1. Priority Question: In contrast to the NICE scope, the model does not include fludarabine and rituximab (FR), cyclophosphamide and rituximab (FCR), cladribine and rituximab (Clad-R) and autologous stem cell transplantation (ASCT) (for patients who have had at least one prior therapy), chlorambucil, rituximab monotherapy and BSC (for patients for whom chemo-immunotherapy is unsuitable) as comparators. The company stated that "Other than BR and DRC, it was not possible to conduct comparisons with chemotherapy regimens or BSC, due to a lack of data in the literature to enable comparison of zanubrutinib with the comparators of interest". Although the ERG agrees that there is a lack of data for these comparators, the company could estimate the efficacy of the above comparators based on expert opinion or adjust the BR or DRC efficacy estimates to meet clinical expectations. Full incremental analyses should be provided where there is more than one comparator.**

**Please include FR, FCR, Clad-R, ASCT, chlorambucil, rituximab monotherapy and BSC as a comparator in the model and provide a full incremental analysis.**

As detailed in Appendix D, Section D.1.2 of the company submission, there was a lack of data to inform the inclusion of non-BR/DRC comparators in the cost-effectiveness analysis. In addition, for BR and DRC (which were included in the cost-effectiveness analysis), a full incremental analysis was not applicable because the treatment comparisons relied on pairwise comparisons using an MAIC approach, in which the patient populations vary by treatment comparison. Furthermore, the BR population and the DRC population were different in terms of prior line of treatment.

The company considered conducting exploratory analyses during the clarification stage by relying on certain assumptions, such as (1) assuming equivalent clinical outcomes between BR and other chemo-immunotherapies (e.g., FR/FCR/Clad-R, chlorambucil) specified for patients with relapsed or refractory disease (i.e. adults with WM who have received at least one prior therapy), and (2) applying actual drug

costs specific to each comparator regimen (e.g., FR/FCR/Clad-R, chlorambucil). However, such analyses were not conducted due to a lack of evidence (i.e., a lack of randomised trials directly comparing these regimens) to justify assumption (1) above. Although there was some evidence from the MAIC comparing zanubrutinib to some of these comparator regimens (see the response to clarification question A26), because the MAICs were conducted in a pairwise manner, the results were not informative for a full incremental analysis for the same patient population. In addition, the limitations of MAIC regarding known discrepancies between the zanubrutinib population and the comparator population, regarding, for example, prior treatment history (i.e., previously untreated versus treated) and unknown discrepancies in unobserved variables also apply. As such, the existing clinical evidence (including the ASPEN trial for zanubrutinib and the clinical studies identified from the clinical SLR) would not be sufficient for conducting a full incremental analysis. It may also be argued that equivalent efficacy could be an inappropriate assumption, without robust evidence for or against. The company also considered referring to clinical expert opinion, but it was challenging for clinicians to make such statements or provide approximations of hazard ratios for the survival outcomes, especially when their opinions may also be subject to biases (e.g., clinicians' experience, patient profiles, facilities).

Compared with the *adults with WM who have had at least one prior therapy* population, the issue of limited data was even more obvious for the other population of interest specified in the NICE scope (*adults with WM whose disease is untreated, for whom chemo-immunotherapy is unsuitable)*.

Considering the above, no full incremental analyses were conducted due to a lack of solid clinical evidence to justify such an analyses.

In addition to the lack of clinical evidence for the non-BR/DRC comparator regimens, the company also consulted clinical experts who confirmed that, according to an upcoming 2021 WMUK registry report,[23] DRC and BR (two of the comparator regimens included in the company submission) were administered to 70.2% (n=172) of all WM patients during the past 3 years (2017–2020; 33% treated with DRC and 33.2% treated with BR). The third most common treatment administer in the UK was

ibrutinib (18.2%). Therefore, the three comparators included in the company submission (BR, DRC and ibrutinib) account for 88.4% of UK patients with WM treated between 2017 and 2020. In contrast, the treatments administered to the remaining 11.6% of patients (including FR, FCR, Clad-R, chlorambucil and autologous HSCT), each accounted for very few patients (n=4–7 [<4%]) rendering them unsuitable for use as comparators. In addition, some comparators (such as F and FR) have not been used for the treatment of WM in the UK since 2010.[3] These data further confirm the relevance and importance of DRC and BR to UK clinical practice, compared with the other regimens listed in the NICE scope.

**B2.** In contrast to zanubrutinib (97.64%), the relative dose intensity of BR and DRC was assumed to be 100%.

a. Please provide evidence for this assumption.

b. Please provide scenario analyses in which similar dose intensity rates are assumed for zanubrutinib and the comparators.

Unlike the zanubrutinib arm where patient-level data (including relative dose intensity) were directly available from ASPEN, there was a lack of reported relative dose intensities for BR and DRC in the corresponding publications. The company agrees that a (common) alternative would be to use the same estimate between zanubrutinib and the comparators; but given the discrepancies in the drug class (i.e., BTK inhibitors versus chemo-immunotherapies), this alternative option also requires assumptions that are difficult to verify based on publicly available information from comparator publications. However, a scenario analysis was conducted using the same relative dose intensity for zanubrutinib, BR and DRC. The results are presented in the table below, which showed that this assumption had minimal impact on the incremental cost-effectiveness ratio (ICER).

**Table 24. Scenario analysis assuming the same relative dose intensity for zanubrutinib, BR and DRC**

| | ICER (pairwise comparison with BR) | ICER (pairwise comparison with DRC) |
|---|---|---|
| CS base-case analysis (97.64% for zanubrutinib; 100% for BR and DRC) | [redacted] | [redacted] |
| Scenario analysis (97.64% for zanubrutinib, BR and DRC) | [redacted] | [redacted] |

Abbreviations: BR = rituximab and bendamustine; CS = company submission; DRC = dexamethasone, rituximab and cyclophosphamide; ICER = incremental cost-effectiveness ratios
Note: For transparency and simplicity, the results above were based on the base-case analysis in the initial company submission, revised only according to this specific clarification question, without accounting for the revisions of model inputs/programming in response to other clarification questions. To generate the results for different combinations of revisions, please refer to the updated Excel model

**B3. Priority Question: Ibrutinib is included as comparator but is also included as subsequent treatment to BR and DRC. In the company submission it is stated that "Data from literature and previous HTA submissions were used to inform the subsequent treatment use and distribution in the model". However, no further justification (or specific references) has been provided.**

a. **Please provide extensive justification for the use of ibrutinib as subsequent treatment. Please note that NICE's position statement: consideration of products recommended for use in the Cancer Drugs Fund as comparators, or in a treatment sequence, in the appraisal of a new cancer product states that technologies available through the CDF should not be modelled in treatment sequences.**

b. **Please provide additional evidence (e.g. expert opinion or clinical trials) that gives insight into possible subsequent treatments in absence of ibrutinib in the UK.**

c. **Based on your response to B3b above, please provide an alternative scenario in which alternative subsequent treatments have been explored.**

Ibrutinib is considered to be clinically relevant as a subsequent treatment, given that data of the UK WMUK Rory Morrison Registry up to 2018 (a registry with a total of 579 WM patients registered from 19 hospitals across the UK) indicates that BTK

inhibitors (currently only ibrutinib is available) are an emerging standard of care in patients who have had ≥1 prior therapy, with ibrutinib being the most frequently used treatment in clinical practice.

According to the upcoming 2021 WMUK registry report (see response to clarification question B1),[3] ibrutinib use during the last 3-year period (2017-2020) was 18.2% (n=45), positioning it as the third most common treatment in UK after DRC (37% [n=71]) and BR (33.2% [n=63]). Hence, ibrutinib is considered a clinically valid comparator alongside BR and DRC, and the inclusion of ibrutinib as a third comparator increases the scale of the comparison to 88.4% of patients with WM, compared with 70.2% for BR and DRC alone. Moreover, the number of patients treated with DRC, BR and ibrutinib in UK during the past three years (71, 63 and 45, respectively) are adequate and balanced for comparison with the WM patients treated with zanubrutinib in the ASPEN study.

## *Model structure*

**B4.** In the economic model, OS and PFS from the ASPEN trial were used to inform the proportion of patients per health state over time. However, PFS and OS were respectively secondary and exploratory endpoints in the ASPEN trial. The primary endpoint in this study was achieving a very good partial or complete response. In addition, partitioned survival models (PSMs) are often used for diseases with a relatively short PFS and OS. However, in the company model patients remain progression-free relatively long and as a result health related quality of life and cost and resource use are stable over a relatively long period. The ERG considers a state transition model (STM) including health states based on response status (primary endpoint in the ASPEN trial) may be more suitable. Furthermore, NICE Decision Support Unit (DSU) technical support document (TSD) 19 recommended the use of

STMs alongside PSMs to verify the plausibility of PSM extrapolations and explore key clinical uncertainties in the extrapolation period.

a. Please elaborate on the potential limitations of informing model health state occupancy based on secondary and exploratory trial endpoints and justify why the primary trial endpoint (response status) was not used to inform the model.

b. Please elaborate on the plausibility of HRQoL and cost and resource use being stable over a relatively long period in the current model and comment on the applicability of a state transition model in which HRQoL and cost and resource use can be included conditional on response status.

c. Please justify the use of a PSM given the issues highlighted in NICE DSU TSD 19, particularly regarding the extrapolation of PFS and OS while assuming structural independence between these endpoints.

d. Please provide a STM to

    I. inform health state occupancy, HRQoL, and cost and resource use based on response status from the ASPEN trial.

    II. assist in verifying the plausibility of the PSM extrapolations and to address uncertainties in the extrapolation period (NICE DSU TSD 19, recommendation 11).

Although the company acknowledges that there are limitations to PSMs, as there are with any modelling approach, a STM approach was deemed less appropriate for this submission for the reasons detailed below.

To develop a response-based STM, at least four states are required, even in the simplest scenario (such as “stable disease” [the initial health state], “response”, “relapsed or subsequent treatment” and “death”). As detailed in NICE DSU TSD 19, unlike a PSM that requires only PFS and OS, an STM would require time-to-event data on each individual transition probability for each individual treatment. Using this

four-state model structure as an example, at least six sets of transition probabilities would be required:

1. “Stable disease” to “response”
2. “Stable disease” to “relapsed or subsequent treatment”
3. “Stable disease” to “death”
4. “Response” to “relapsed or subsequent treatment”
5. “Response” to “death”
6. “Relapsed or subsequent treatment” to “death”.

However, the above transition probabilities are not sufficiently available for any of the comparators specified in NICE’s final scope. More specifically, as partially included in Appendix D of the company submission and in response to clarification question A16, very few published studies reported PFS and OS KM to enable the development of a PSM. Additional time-to-event data was not available to inform cause-specific hazards or post-progression (or subsequent treatment) that are required for a response-based STM. As acknowledged in NICE DSU TSD 19, a common challenge of developing a STM was that many cancer-related clinical studies report PFS and OS only, which are insufficient for the development of a STM in a straightforward manner.

In summary, based on the available data for the comparators listed in the final scope, a PSM relying on an MAIC was developed to compare zanubrutinib to the comparators of interest, relying heavily on clinical expert opinion on the validity of extrapolated long-term survival (see Section B.3.3.2 of the company submission and the response to clarification question B5), whereas a STM was deemed unfeasible. Although an STM was deemed unfeasible for the comparison of zanubrutinib versus the compactors listed in the final scope, with the patient-level ASPEN data for zanubrutinib, an STM is potentially feasible for the zanubrutinib arm alone. However, the company has concerns about the development of even the simplest three-state STM, given the reliance on data from only 8 patients in the zanubrutinib arm (N=102)

who progressed to inform post-progression survival. Therefore, a response-based STM was not developed.

### *Treatment effectiveness*

**B5. Priority Question: Given that the ASPEN trial data are extremely immature, it is difficult to meaningfully extrapolate OS and PFS beyond the available study data.**

a. **Please provide evidence that the modelled OS and PFS beyond the trial period is plausible. For example, please explore any British or EU-based real world databases to examine OS and PFS estimates and smoothed hazard plots, if available.**

b. **Please consider performing survival analysis using external data from BGB-3111-AU-003 (long-term follow-up from phase 1/2 study), for example using the method described by Soikkeli et al (Extrapolating survival data using historical trial based a priori distributions Value Health 2019 Sep;22(9):1012-1017 that was also mentioned in TSD 21. This would assume that long-term hazards (not absolute survival) are comparable between the two studies. Please also discuss whether this assumption may be appropriate given any potential differences between studies, for instance in population and treatment. Please provide an updated model file based on these analyses.**

It is acknowledged that the immaturity of ASPEN survival data necessitates assessments of external validity. Therefore, several assessments were conducted, including (1) comparison of modelled landmark survival versus the observed survival in BGB-3111-AU-003, the Phase 1/2 trial for zanubrutinib with longer median follow-up of 48 months (compared with the median follow-up of 19 months for ASPEN), (2) review of external literature and technology appraisals (including clinical trials for other BTK inhibitors in the WM population, NICE TA491 [ibrutinib for treating WM], and other published literature [e.g., the ESMO guideline for WM]), and (3) clinical expert opinion on the clinical plausibility of modelled survival and hazard patterns

(see Section B.3.3.2.1 of the company submission). Further details for each of these criteria are provided below.

For the assessment relying on BGB-3111-AU-003 data, landmark OS rates observed in BGB-3111-AU-003 (48-month OS rates of 78.7% in a total of 73 patients, 75.9% in 49 relapsed/refractory patients, and 83.5% in 24 treatment-naïve patients, after a median follow-up of 48 months) were compared to landmark OS from the extrapolated OS curves to assess the validity of the latte15r. However, it should be acknowledged that the issue of immature data (with relatively short follow-up) also exists for BGB-3111-AU-003, with or without ASPEN data. As a result, despite BGB-3111-AU-003 being considered in the company submission, more weight was placed on other assessments (e.g., clinical expert opinion), as discussed below.

In addition to the Phase 1/2 BGB-3111-AU-003 trial, survival results of the previously published clinical trial for ibrutinib (Phase 2 Study 1118E of ibrutinib monotherapy) were also reviewed. However, given the immaturity of publicly available survival data (e.g., OS rate of 90% after a median follow-up of 37 months),{© National Institute for Health and Care Excellence, November 2017 #65} these results are not informative for the validation of long-term survival extrapolation.

Given the general immaturity of survival data in the clinical trials of BTK inhibitors in WM, the long-term OS estimates based on less recent studies (in which BTK inhibitors were not an available treatment option) were reviewed. According to NICE TA491 and the 2018 ESMO clinical practice guidelines for WM,{Buske, 2013 #16;© National Institute for Health and Care Excellence, November 2017 #65} patients not treated with BTK inhibitors had a median OS of approximately 10 years. Although these studies may not be completely informative for validation of the exact OS with BTK inhibitors, given the data limitations in WM, it may still be informative to rely on all available data to inform the plausible range of OS in patients treated with BTK inhibitors. For example, it was reported in NICE TA491 that the median OS in WM ranged from <4–12 years and that median OS in the European chart review study was 123 months (i.e., approximately 10 years) for patients receiving a mix of physicians' choice of therapy (second-line: 47% BR, 31% DRC, 11% FCR, 0% Clad-R, 11% other; third- or fourth-line: 43% BR, 15% DRC, 9% FCR, 30% Clad-R, 3%

other).{© National Institute for Health and Care Excellence, November 2017 #65} However, considerable country-specific OS differences were noted (e.g., UK-specific median OS was reported to be 5 years; exact estimates for other EU countries were not publicly reported). ESMO clinical practice guidelines report that the median OS exceeded 10 years for younger patients and is relatively shorter for elderly patients.{Buske, 2013 #16}

In addition to the published estimates above, expert opinion on the clinical plausibility of the modelled OS estimates and the hazard patterns is summarised in Table 25 and Table 26 (and in meeting minutes referenced in the company submission), and detailed further in Section B.3.3.2 of the company submission. The parametric models considered to be clinically plausible based on the mean OS and hazard patterns were included in either the base-case or scenario analyses in the company submission.

**Table 25. Plausible survival estimates based on clinical expert opinion**

| Treatment | Population/setting | Plausible range of survival |
|---|---|---|
| BTK inhibitor | 85% R/R and 15% TN unsuitable for chemo-immunotherapy | ~15 years |
| Chemo-immunotherapy | TN suitable for chemo-immunotherapy; ~15 years ago | 9–11 years |
| Chemo-immunotherapy | TN suitable for chemo-immunotherapy; present day | 12–15 years |
| Chemo-immunotherapy | R/R - 2L; approximately 15 years ago | 6–7 years |
| Chemo-immunotherapy | R/R - 2L; present day | 8–10 years |
| Chemo-immunotherapy | R/R - 3L; approximately 15 years ago | 2–4 years |
| Chemo-immunotherapy | R/R - 3L; present day | 4–6 years |

Abbreviations: 2L = second-line; 3L = third-line; BTK = Bruton's tyrosine kinase; R/R = relapsed/refractory; TN = treatment naïve

**Table 26. Clinical expert opinion on survival in WM**

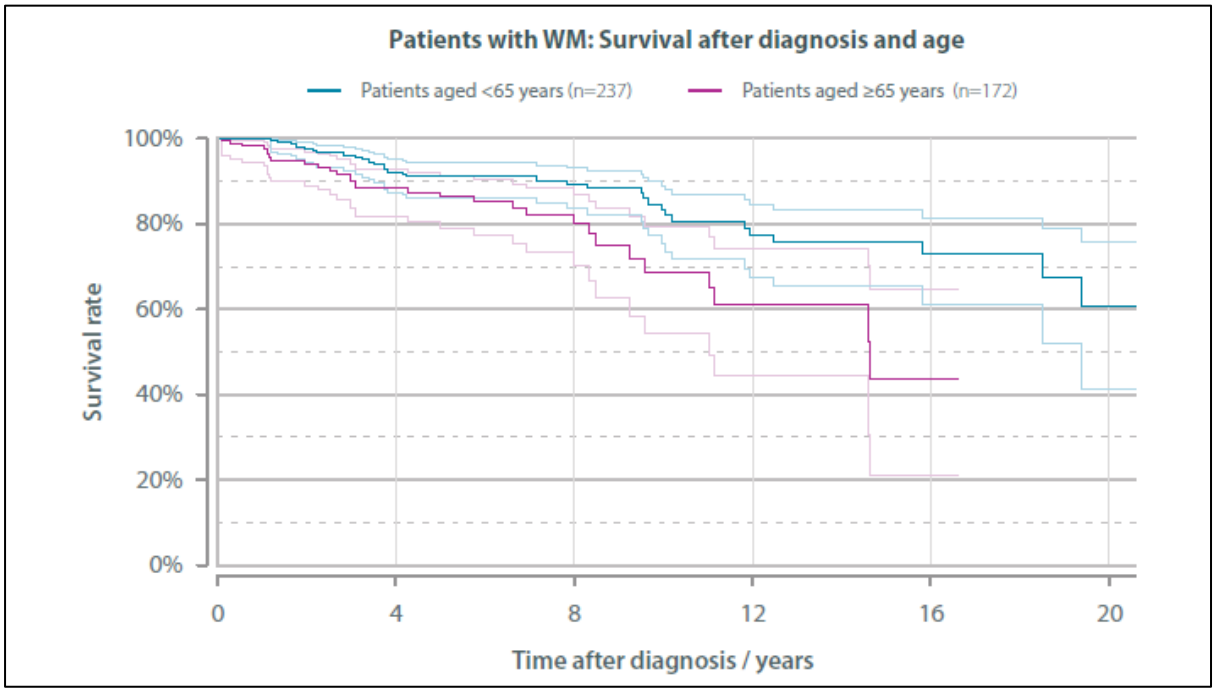
| Outcome | Population/treatment regimen/ subsequent treatment | Clinical experts' comments |
|---|---|---|
| Worst OS | R/R patients treated with chemo-immunotherapy, followed by subsequent (different) chemo-immunotherapy | Monotonically increasing |
| | TN patients treated with BTK inhibitor unsuitable for chemo-immunotherapy | Monotonically increasing; increasing faster than R/R patients treated with BTK inhibitor, because (as discussed above) these patients were expected to soon run out of active treatment option after progressing on BTK inhibitor and then rituximab monotherapy |

| Outcome | Population/treatment regimen/ subsequent treatment | Clinical experts' comments |
|---|---|---|
| Best OS | Similar survival among<br>R/R patients treated with BTK inhibitor, followed by subsequent chemo-immunotherapy<br>R/R patients treated with chemo-immunotherapy, followed by subsequent BTK inhibitor<br>TN patients treated with chemo-immunotherapy, followed by subsequent (different) chemo-immunotherapy | Monotonically increasing<br>Monotonically increasing; increasing in a same rate as R/R patients treated with BTK inhibitor, as these patients were expected to be treated with BTK inhibitor as subsequent treatment (provided that these patients were not previously treated with BTK inhibitor)<br>Monotonically increasing; increasing in similar rate as that in R/R patients treated with BTK inhibitor and R/R patients treated with chemo-immunotherapy |
| | TN patients treated with chemo-immunotherapy, followed by subsequent BTK inhibitor | Monotonically increasing; increasing in the slowest rate than all the 3 patient groups from above |

Abbreviations: BTK = Bruton's tyrosine kinase; OS = overall survival; R/R = relapsed/refractory; TN = treatment naïve

In addition to the above, further validations were performed during the clarification stage by comparing extrapolated survival to the observed survival reported in the first WMUK registry report from the Rory Morrison Registry (N=579 from 19 hospitals across the UK).{WMUK, 2018 #29} OS in patients in the WMUK report is presented in Figure 31, stratified by age group (<65 versus ≥65 years). The median OS for patients diagnosed at <65 years was 29.5 years, compared with 14.6 years for ≥65 years.{WMUK, 2018 #29}

**Figure 31. KM plot of OS in patients with WM, stratified by age at diagnosis**



Abbreviations: KM = Kaplan-Meier; OS = overall survival; WM = Waldenström's macroglobulinemia
Source: WMUK, 2018{WMUK, 2018 #29}

Considering that patients in the zanubrutinib arm of ASPEN had a mean/median age of 69.2/70.0 years and that the BR and DRC populations had a median age of 72 and 69 years, respectively, the OS of the population aged ≥65 years in the WMUK report was considered to be more comparable to the trial populations, and therefore more relevant for the assessment of external validity.

As shown in Table 27, the median OS (based on the extrapolated curves selected for the base-case analyses, after adjusting for background mortality) for the BTK inhibitors (13.15–15.29 years, including the zanubrutinib arms after matching adjustments) was broadly aligned with the observed median OS (14.6 years) from the Rory Morrison Registry, both of which were higher than the median OS for the BR (5.94 years) and DRC (7.78 years) arms in the model. However, despite the reporting of age, gender and treatment patterns in the overall population in the WMUK report, there was a lack of information in the report on patient characteristics (e.g., year of diagnosis, number of prior lines of treatment, other key prognostic factors) and treatment pattern (e.g., proportion of patients treated with BTK inhibitors) specifically for patients aged ≥65 years. As a result, further assessments of comparability of the populations between the Rory Morrison Registry and the clinical studies for zanubrutinib, BR and DRC were not possible. Therefore, the observations above regarding long-term OS should be interpreted with caution.

**Table 27. Median OS (based on the extrapolated curves selected for the base-case analyses, after adjusting for background mortality**

| | Zanubrutinib | Ibrutinib | Zanubrutinib (match BR) | BR | Zanubrutinib (match DRC) | DRC |
|---|---|---|---|---|---|---|
| Median per the base case model | 15.06 | 13.15 | 15.29 | 5.94 | 14.60 | 7.78 |

Abbreviations: BR = rituximab and bendamustine; DRC = dexamethasone, rituximab and cyclophosphamide

In addition to the above, given that the ERG also proposed using the method described by Soikkeli et al. 2019 as part of the validation process of the long-term survival extrapolation using the ASPEN trial data, during the clarification stage, the company conducted a comparison of the populations of the ASPEN and BGB-3111-

AU-003 trials (Table 28), and fitted independent parametric models to zanubrutinib using ASPEN and BGB-3111-AU-003 trial data separately (Table 29–Table 32).

For this exploratory analysis, the PFS and OS data of 69 patients treated with zanubrutinib 160 mg BID or 320 mg OD in the BGB-3111-AU-003 dose expansion part (part 2) were extracted and included. As shown in Table 28, the BGB-3111-AU-003 trial population (N=69) included a statistically significantly (chi-square test) higher proportion of treatment-naïve patients (unsuitable for chemo-immunotherapy), a lower proportion of patients with *MYD88*$^{MUT}$, and a higher proportion of patients with unknown *CXCR4* status, compared with the zanubrutinib arm of the ASPEN trial (N=102). Other patient characteristics were relatively comparable.

**Table 28. Comparison of ASPEN and BGB-3111-AU-003 trial populations**

| | **BGB-3111-AU-003 (N=69)** | **ASPEN Cohort 1 ITT set, zanubrutinib arm (N=102)** | **p value** |
|---|---|---|---|
| Patient group | | | |
| Relapsed or refractory | 45 (65.2) | 83 (81.4) | 0.027 |
| Treatment-naïve (unsuitable for chemo-immunotherapy) | 24 (34.8) | 19 (18.6) | |
| Age | | | |
| Mean (SD), years | 67.03 (10.89) | 69.16 (10.26) | 0.202 |
| Median | 67 | 70 | 0.244 |
| Gender, n (%) | | | |
| Female | 14 (20.3) | 33 (32.4) | 0.119 |
| Male | 55 (79.7) | 69 (67.6) | |
| ECOG, n (%) | | | |
| 0 | 25 (36.2) | 46 (45.1) | 0.407 |
| 1 | 41 (59.4) | 50 (49.0) | |
| 2 | 3 (4.3) | 6 (5.9) | |
| *MYD88* status, n (%) | | | |
| MUT | 33 (47.8) | 102 (100.0) | <0.001 |
| Unknown | 32 (46.4) | 0 (0.0) | |
| WT | 4 (5.8) | 0 (0.0) | |
| *CXCR4* status, n (%) | | | |
| Unknown | 42 (60.9) | 3 (2.9) | <0.001 |
| WHIM | 7 (10.1) | 14 (13.7) | |
| WT | 20 (29.0) | 85 (83.3) | |
| Number of prior therapies, n (%) | | | |
| 0 | 24 (34.8) | 19 (18.6) | 0.098 |
| 1 | 21 (30.4) | 47 (46.1) | |
| 2 | 8 (11.6) | 15 (14.7) | |
| 3 | 6 (8.7) | 14 (13.7) | |
| 4 | 5 (7.2) | 4 (3.9) | |
| 5 | 2 (2.9) | 0 (0.0) | |
| 6 | 0 (0.0) | 1 (1.0) | |
| 7 | 1 (1.4) | 1 (1.0) | |
| 8 | 2 (2.9) | 1 (1.0) | |

| | | | |
|---|---|---|---|
| Baseline extramedullary disease, n (%) | | | |
| No | 25 (36.2) | 39 (38.2) | 0.917 |
| Yes | 44 (63.8) | 63 (61.8) | |
| Splenomegaly or adenopathy, n (%) | | | |
| No | 26 (37.7) | 40 (39.2) | 0.966 |
| Yes | 43 (62.3) | 62 (60.8) | |
| Baseline platelet count, 109/L | | | |
| Mean (SD) | 214(98) | 241 (108) | 0.094 |
| Median | 206 | 237 | 0.094 |
| Baseline haemoglobin, g/L | | | |
| Mean (SD) | 104.61 (18.77) | 104.39 (19.24) | 0.942 |
| Median | 103 | 102.5 | 0.864 |
| Baseline IgM, g/L | | | |
| Mean (SD) | 33.94 (20.21)* | 33.19 (18.27) | 0.809 |
| Median | 32.4* | 31.8 | 0.991 |
| Baseline Beta-2 microglobulin, mg/L | | | |
| Mean (SD) | 4.72 (3.22) | 4.92 (2.91) | 0.771 |
| Median | 4.02 | 4.25 | 0.395 |

*4 missing
Abbreviations: CXCR4 = C-X-C Motif Chemokine Receptor 4; IgM = immunoglobulin M; ITT = intention-to-treat; SD = standard deviation; WHIM = warts, hypogammaglobulinemia, infections, myelokathexis; WT = wild-type

Figure 32 and Figure 33 present the OS and PFS curves of ASPEN and BGB-3111-AU-003. Table 29–Table 32 present the results of independently fitted parametric models of zanubrutinib using the ASPEN and BGB-3111-AU-003 trial data separately. As discussed in Section B.3.3.2.1.2 of the CS, to avoid over-fitting, the goodness-of-fit was assessed based on BIC statistics, which showed that the exponential model provided better fits to the KM curves of both ASPEN and BGB-3111-AU-003 trial data. Such results were consistent with all the results of fit statistics in Section B.3.3.2.1.2, B.3.3.2.2.2, and B.3.3.2.3.2 of the CS (based on ASPEN trial data), which showed that the exponential distribution was associated with the lowest BIC for all the parametric models for zanubrutinib. Of note, although the results of fit statistics supported the use of the exponential model (using either the ASPEN trial data or the BGB-3111-AU-003 trial data with relatively longer follow-up), given the immaturity of survival data, the model selection (discussed throughout Section B.3.3.2 of the CS) relied heavily on the external validity assessment. In addition, given the consistency in the results of fit statistics between the ASPEN and BGB-3111-AU-003 data, using historical data-based a priori distributions (per Soikkeli et al. 2019) to update the Excel model was not applicable.

**Figure 32. KM curves of OS – ASPEN Cohort 1, zanubrutinib arm vs BGB-3111-AU-003**

Abbreviations: KM = Kaplan-Meier; OS = overall survival; vs = versus

**Figure 33. KM curves of PFS – ASPEN cohort 1, zanubrutinib arm vs BGB-3111-AU-003**

Abbreviations: KM = Kaplan-Meier; PFS = progression-free survival; vs = versus

**Table 29. Fit statistics and model parameters for independent models of OS of zanubrutinib (without matching adjustment), using ASPEN trial data**

| Parametric distribution | AIC | BIC | Parameter | | Estimate |
|---|---|---|---|---|---|
| Exponential | 83.2615 | 85.8864 | λ | rate | 0.0031 |
| Weibull | 84.5425 | 89.7924 | p | shape | 1.4189 |
| | | | λ | scale | 141.96 |
| Gompertz | 84.2064 | 89.4563 | p | shape | 0.0624 |
| | | | λ | scale | 0.0015 |
| Log-normal | 84.7351 | 89.985 | μ | meanlog | 5.6279 |
| | | | σ | sdlog | 1.6978 |
| Log-logistic | 84.5908 | 89.8408 | p | shape | 1.4366 |
| | | | λ | scale | 135.7831 |
| Gamma | 84.5784 | 89.8284 | k | shape | 1.4438 |
| | | | λ | rate | 0.0091 |

Abbreviations: AIC = Akaike information criterion; BIC = Bayesian information criterion; OS = overall survival

**Table 30. Fit statistics and model parameters for independent models of OS of zanubrutinib, using BGB-3111-AU-003 trial data**

| Parametric distribution | AIC | BIC | Parameter | | Estimate |
|---|---|---|---|---|---|
| Exponential | 132.3906 | 134.6247 | λ | rate | -5.5195 |
| Weibull | 133.7354 | 138.2036 | p | shape | 0.2516 |
| | | | λ | scale | 5.1068 |
| Gompertz | 134.2963 | 138.7645 | p | shape | 0.0073 |
| | | | λ | scale | -5.6716 |
| Log-normal | 133.2013 | 137.6695 | μ | meanlog | 5.2525 |
| | | | σ | sdlog | 0.4143 |
| Log-logistic | 133.5615 | 138.0297 | p | shape | 0.3057 |
| | | | λ | scale | 4.9622 |
| Gamma | 133.6502 | 138.1185 | k | shape | 0.3123 |
| | | | λ | rate | -4.7995 |

Abbreviations: AIC = Akaike information criterion; BIC = Bayesian information criterion; OS = overall survival

**Table 31. Fit statistics and model parameters for independent models of PFS of zanubrutinib (without matching adjustment), using ASPEN trial data**

| Parametric distribution | AIC | BIC | Parameter | | Estimate |
|---|---|---|---|---|---|
| Exponential | 124.7975 | 127.4224 | λ | rate | 0.0059 |
| Weibull | 126.7598 | 132.0097 | p | shape | 0.9438 |
| | | | λ | scale | 194.9886 |
| Gompertz | 126.7899 | 132.0398 | p | shape | -0.0044 |
| | | | λ | scale | 0.0061 |
| Log-normal | 126.472 | 131.722 | μ | meanlog | 5.7651 |
| | | | σ | sdlog | 2.2557 |
| Log-logistic | 126.7734 | 132.0234 | p | shape | 0.9679 |
| | | | λ | scale | 173.6035 |
| Gamma | 126.7613 | 132.0112 | k | shape | 0.9406 |
| | | | λ | rate | 0.0049 |

Abbreviations: AIC = Akaike information criterion; BIC = Bayesian information criterion; PFS = progression-free survival

**Table 32. Fit statistics and model parameters for independent models of PFS of zanubrutinib, using BGB-3111-AU-003 trial data**

| Parametric distribution | AIC | BIC | Parameter | | Estimate |
|---|---|---|---|---|---|
| Exponential | 201.0495 | 203.2836 | λ | rate | -4.8544 |
| Weibull | 202.9231 | 207.3913 | p | shape | -0.0791 |
| | | | λ | scale | 4.9603 |
| Gompertz | 202.116 | 206.5842 | p | shape | -0.0192 |
| | | | λ | scale | -4.5184 |
| Log-normal | 200.9382 | 205.4064 | μ | meanlog | 4.8052 |
| | | | σ | sdlog | 0.5937 |
| Log-logistic | 202.4355 | 206.9037 | p | shape | 0.004 |
| | | | λ | scale | 4.6849 |
| Gamma | 202.9856 | 207.4538 | k | shape | -0.0664 |
| | | | λ | rate | -4.9941 |

Abbreviations: AIC = Akaike information criterion; BIC = Bayesian information criterion; PFS = progression-free survival

**B6. Priority Question: In case the assessment of MYD88 is not standard practice in the England & Wales NHS (see clarification question A17), either both cohorts would be treated irrespective of their MYD88 status, or the assessment of MYD88 would have to be added to the treatment pathway. Please provide an updated economic model with two scenarios:**

- **analysis using data from both zanubrutinib cohorts 1 & 2 (survival analysis for OS, PFS, TTD using merged KM data, health-related quality of life).**
- **add MYD88 assessment costs to the treatment pathway in the zanubrutinib arm.**

The response to this question is split into several components, including:

1. updated MAICs and clinical inputs, using the pooled data of cohort 1 (zanubrutinib arm) (n=102) and cohort 2 (in which all the patients received zanubrutinib, without ibrutinib) (n=28) for zanubrutinib,
2. updated HRQoL inputs, using the pooled data of cohort 1 (including both zanubrutinib and ibrutinib arms) and cohort 2 (in which all the patients received zanubrutinib), and
3. updated cost-effectiveness model outputs.

Of note, unlike cohort 1 (i.e., the ITT population of ASPEN), where patients were randomised to zanubrutinib or ibrutinib, cohort 2 included patients treated with zanubrutinib only. Therefore, in case of potential biases, this scenario analysis using the pooled data of cohorts 1 and 2 was only conducted for the comparisons with BR and DRC, for both the MAIC and cost-effectiveness analysis. However, for the HRQoL analysis, to maintain the sample size of the analysis, data from the ibrutinib arm of cohort 1 were still captured.

The response below focuses on the differences in inputs and outputs between this scenario analysis (using the pooled data of cohorts 1 and 2) and the base-case analysis of the CS (using cohort 1 data only), with brief discussion only of the contents that are the same as the CS base-case analysis.

*Part 1: Updated MAICs and clinical inputs*

Baseline patient characteristics before and after matching adjustments are shown below.

**Figure 34. Baseline patient characteristics before and after adjusting zanubrutinib to match BR (effective sample size for zanubrutinib after matching adjustment = [redacted])**

Abbreviations: BR = rituximab and bendamustine; IPSS = International Prognostic Scoring System

**Figure 35. Baseline patient characteristics before and after adjusting zanubrutinib to match DRC (effective sample size for zanubrutinib after matching adjustment = [redacted])**

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide

KM curves of PFS, OS and TTD before and after matching adjustments are shown below.

**Figure 36. KM curves of PFS before and after adjusting zanubrutinib to match BR**

Abbreviations: BR = rituximab and bendamustine; KM = Kaplan-Meier; PFS = progression-free survival

**Figure 37. KM curves of OS before and after adjusting zanubrutinib to match BR**

Abbreviations: BR = rituximab and bendamustine; KM = Kaplan-Meier; OS = overall survival

**Figure 38. KM curves of TTD before and after adjusting zanubrutinib to match BR**

Abbreviations: BR = rituximab and bendamustine; KM = Kaplan-Meier; TTD = time to discontinuation

**Figure 39. KM curves of PFS before and after adjusting zanubrutinib to match DRC**

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; PFS = progression-free survival

**Figure 40. KM curves of OS before and after adjusting zanubrutinib to match DRC**

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; OS = overall survival

**Figure 41. KM curves of TTD before and after adjusting zanubrutinib to match DRC**

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; TTD = time to discontinuation

*Zanubrutinib (match BR) vs. BR*

Results of the extrapolations of PFS, OS and TTD for the comparison of zanubrutinib (after matching BR) to BR are shown below, including: (1) summary of model selection, (2) fit statistics and visual inspection, and (3) mean (undiscounted) survival and hazard patterns.

Of note, because the conclusions of the model selection are the same as those in the initial company submission, only summary information are provided below without repeating rationales from B.3.3.2 of the CS.

**Table 33. Summary of model selection for the pairwise comparison of zanubrutinib (after matching BR) vs BR**

| Outcome | Treatment | Base Case Setting | Justification for Model Selection in Base Case | Scenario Analysis Settings | Justification for Model Selection in Scenario Analyses |
|---|---|---|---|---|---|
| OS | Zanubrutinib (matching BR) | Independent exponential model | Clinically plausible mean OS for both treatments<br>Clinically plausible hazard patterns for BR | Dependent Weibull model; dependent gamma model | Clinically plausible mean OS and hazard patterns for zanubrutinib (matching BR)<br>Clinically plausible hazard pattern for both treatment arms |
| | BR | Independent Weibull model | | | |
| PFS | Zanubrutinib (matching BR) | Dependent exponential model | The lowest BIC<br>Alignment with TTD in parametric distribution (specific for zanubrutinib) | None | For both PFS and TTD, the exponential distribution was consistently associated with obviously lower BIC compared to the other distributions. |
| | BR | | | | |
| TTD | Zanubrutinib (matching BR) | Independent exponential model | The lowest BIC<br>Alignment with PFS in parametric distribution (specific for zanubrutinib) | None | |
| | BR | NA | NA | NA | NA |

Abbreviations: BIC, Bayesian information criteria; BR, bendamustine-rituximab; NA, not applicable; OS, overall survival; PFS, progression-free survival; TTD, time to treatment discontinuation

**Table 34. Fit statistics for jointly fitted PFS, OS and TTD; zanubrutinib (after matching BR) vs BR**

| Parametric distribution | PFS | | OS | | TTD | |
|---|---|---|---|---|---|---|
| | AIC | BIC | AIC | BIC | AIC | BIC |
| Exponential | *374.702* | *381.2987* | *235.2915* | *241.8882* | Not applicable | |
| Weibull | 376.6546 | 386.5496 | 237.0649 | 246.9598 | | |
| Gompertz | 376.351 | 386.2459 | 237.2338 | 247.1288 | | |
| Log-normal | 375.0266 | 384.9215 | 236.0111 | 245.906 | | |
| Log-logistic | 375.6711 | 385.5661 | 236.3517 | 246.2466 | | |
| Gamma | 376.5863 | 386.4813 | 236.9461 | 246.841 | | |

Abbreviations: AIC, Akaike information criteria; BIC, Bayesian information criteria; BR, bendamustine rituximab; OS, overall survival; PFS, progression-free survival; TTD, time to discontinuation
Note: The lowest AIC or BIC is in *italics*.

**Table 35. Fit Statistics for independently fitted PFS, OS and TTD; zanubrutinib (after matching BR)**

| Parametric distribution | PFS | | OS | | TTD | |
|---|---|---|---|---|---|---|
| | AIC | BIC | AIC | BIC | AIC | BIC |
| Exponential | 148.4211 | 151.2809 | 74.6711 | 77.5309 | 218.5605 | 221.4125 |
| Weibull | 150.417 | 156.1367 | 75.9889 | 81.7086 | 220.2769 | 225.981 |
| Gompertz | 150.4042 | 156.1238 | 75.0354 | 80.755 | 220.4211 | 226.1252 |
| Log-normal | 149.9417 | 155.6614 | 76.5507 | 82.2703 | 220.024 | 225.7281 |
| Log-logistic | 150.4499 | 156.1695 | 76.0773 | 81.797 | 220.3292 | 226.0332 |
| Gamma | 150.4155 | 156.1351 | 76.0676 | 81.7873 | 220.2867 | 225.9908 |

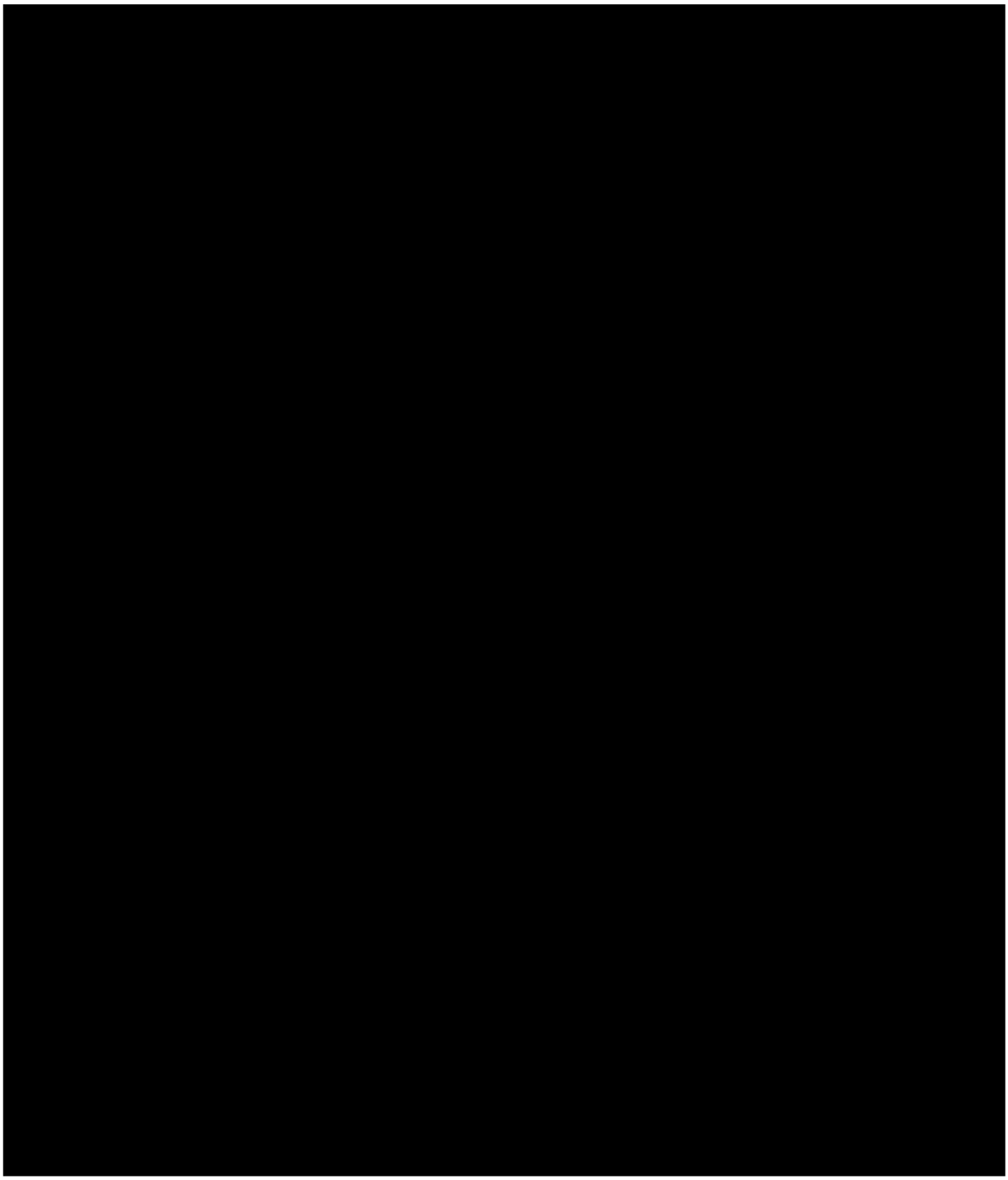
Abbreviations: AIC, Akaike information criteria; BIC, Bayesian information criteria; BR, bendamustine rituximab; OS, overall survival; PFS, progression-free survival; TTD, time to discontinuation
Note: The lowest AIC or BIC is in *italics*.

**Table 36. Fit Statistics for independently fitted PFS, OS and TTD; BR**

| Parametric distribution | PFS | | OS | | TTD | |
|---|---|---|---|---|---|---|
| | AIC | BIC | AIC | BIC | AIC | BIC |
| Exponential | 226.281 | 228.5436 | 160.6204 | 162.8831 | Not applicable | |
| Weibull | 228.2315 | 232.7569 | 162.6073 | 167.1327 | | |
| Gompertz | 227.7723 | 232.2976 | 162.1063 | 166.6317 | | |
| Log-normal | 226.3579 | 230.8833 | 161.4533 | 165.9787 | | |
| Log-logistic | 227.0411 | 231.5664 | 162.0177 | 166.543 | | |
| Gamma | 228.1377 | 232.6631 | 162.5711 | 167.0965 | | |

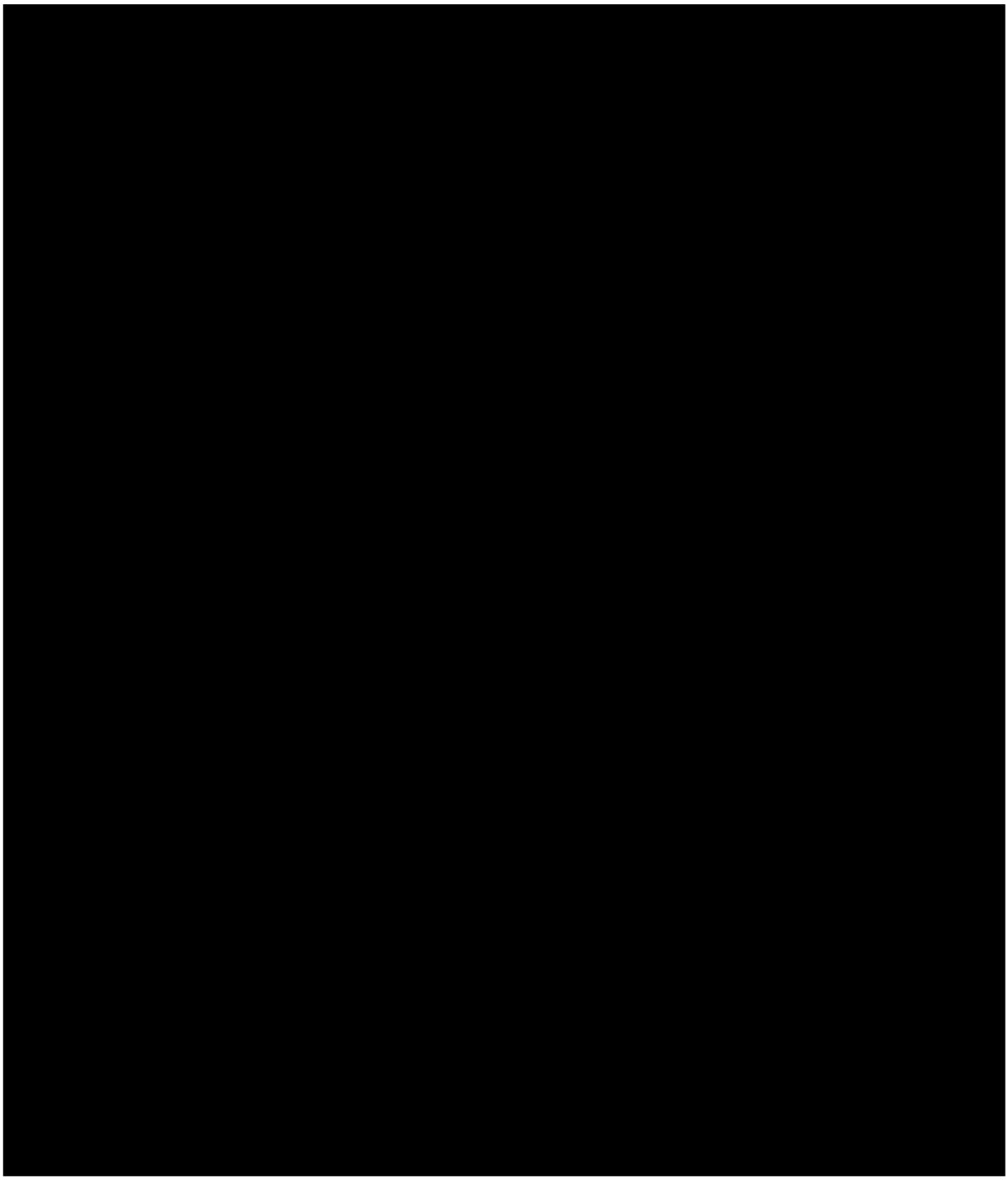
Abbreviations: AIC, Akaike information criteria; BIC, Bayesian information criteria; BR, bendamustine rituximab; OS, overall survival; PFS, progression-free survival; TTD, time to discontinuation
Note: The lowest AIC or BIC is in *italics*.

**Figure 42. Visual inspection of jointly fitted parametric vs KM Curves for PFS of zanubrutinib (After Matching BR) vs BR**



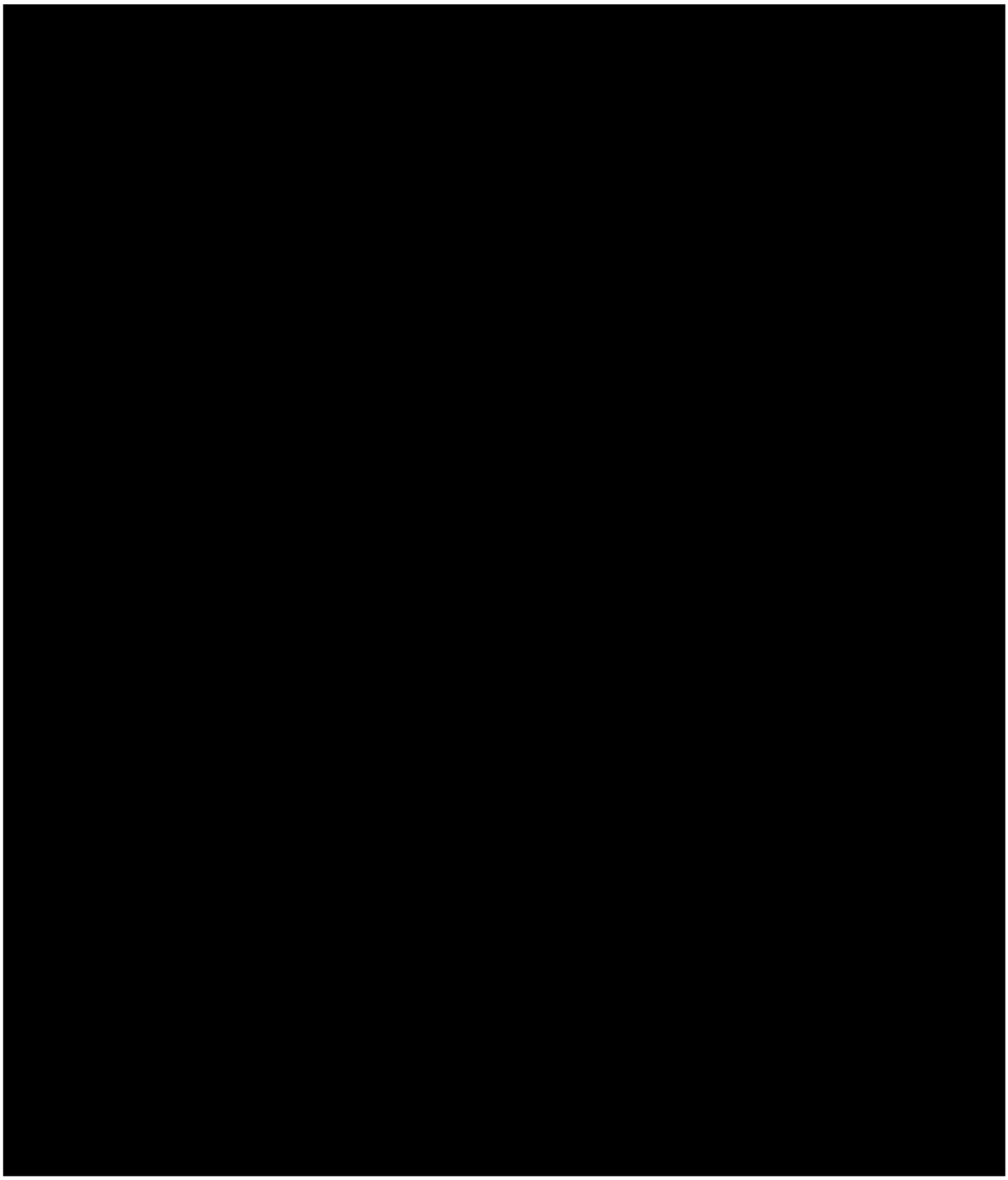
Abbreviations: BR = rituximab and bendamustine; KM = Kaplan-Meier; PFS = progression-free survival

**Figure 43. Visual inspection of jointly fitted parametric vs KM curves for OS of zanubrutinib (after matching BR) vs BR**



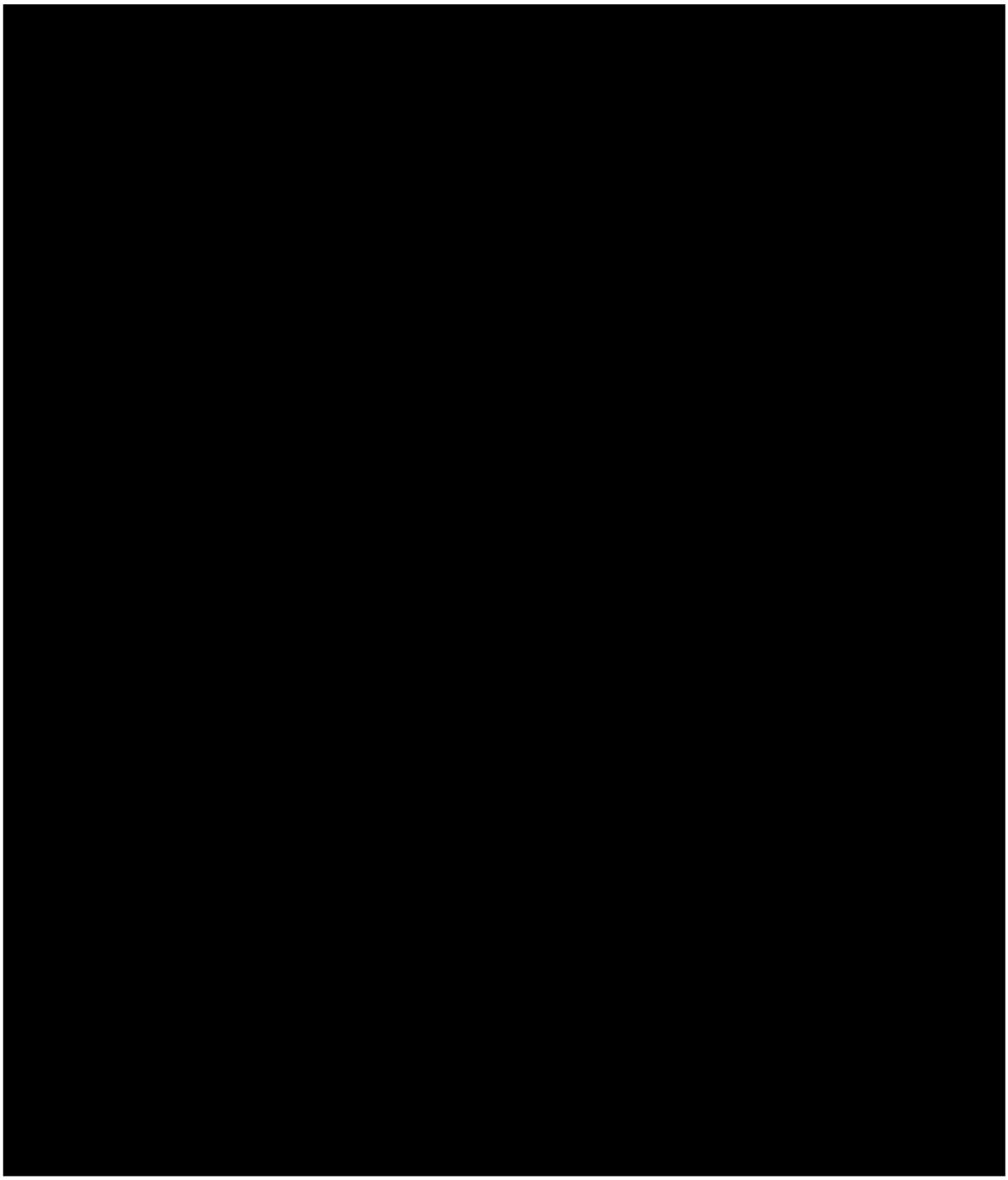
Abbreviations: BR = rituximab and bendamustine; KM = Kaplan-Meier; OS = overall survival

**Figure 44. Visual inspection of independently fitted parametric vs KM curves for PFS of zanubrutinib (after matching BR)**



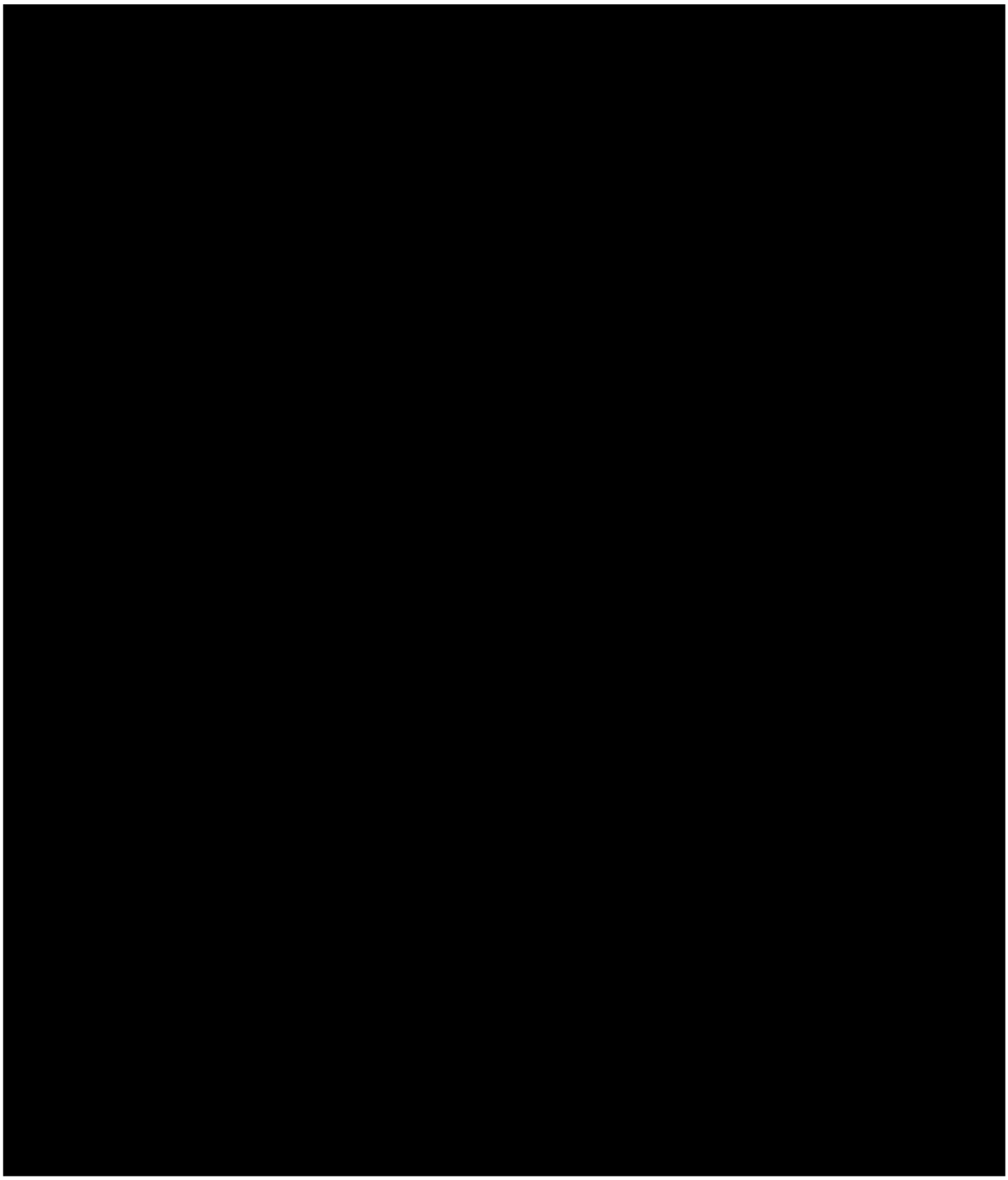
Abbreviations: BR = rituximab and bendamustine; KM = Kaplan-Meier; PFS = progression-free survival

**Figure 45. Visual inspection of independently fitted parametric vs KM curves for PFS of BR (same as the CS)**



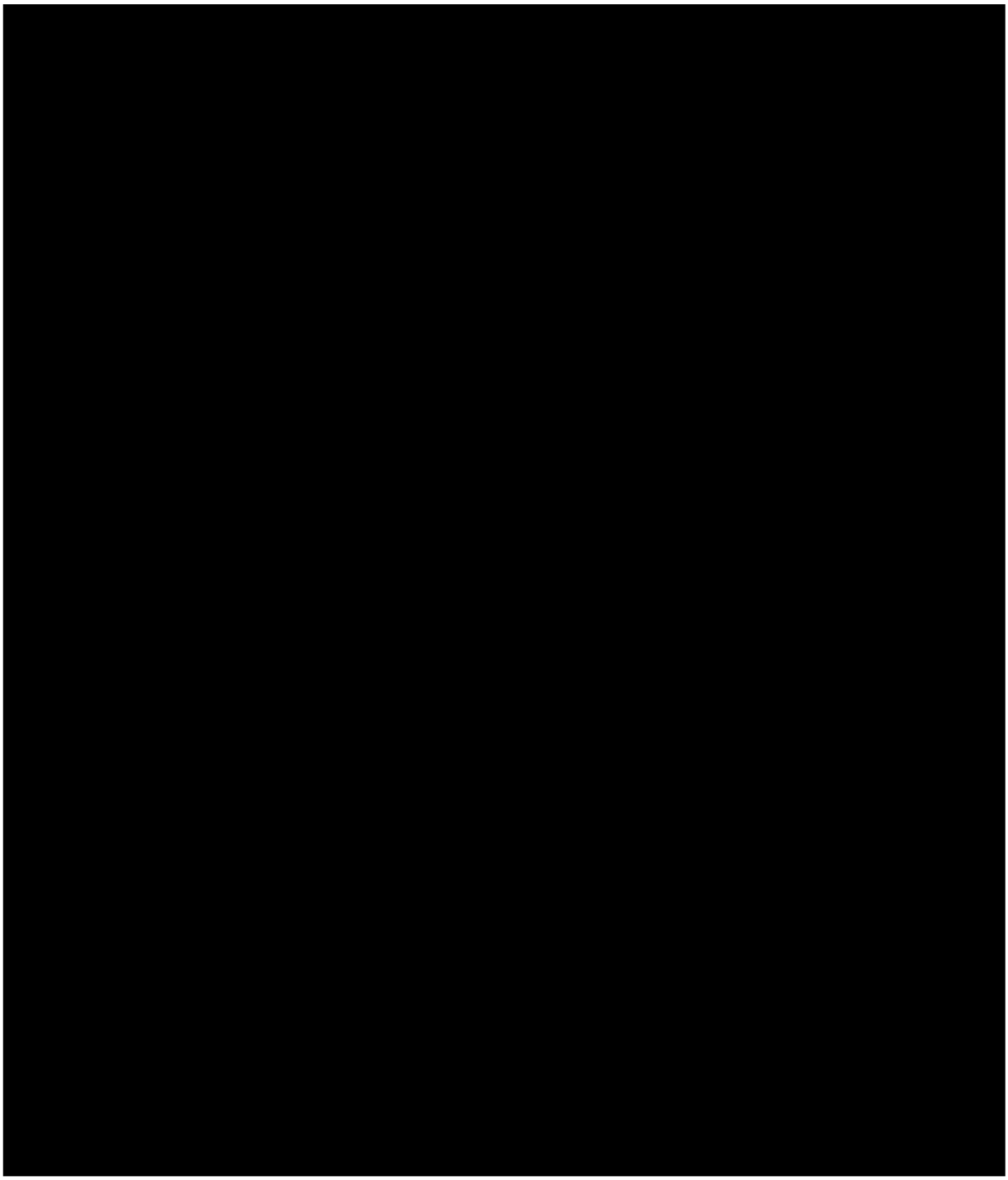
Abbreviations: BR = rituximab and bendamustine; CS = company submission; KM = Kaplan-Meier; PFS = progression-free survival

**Figure 46. Visual inspection of independently fitted parametric vs KM curves for OS of zanubrutinib (after matching BR)**



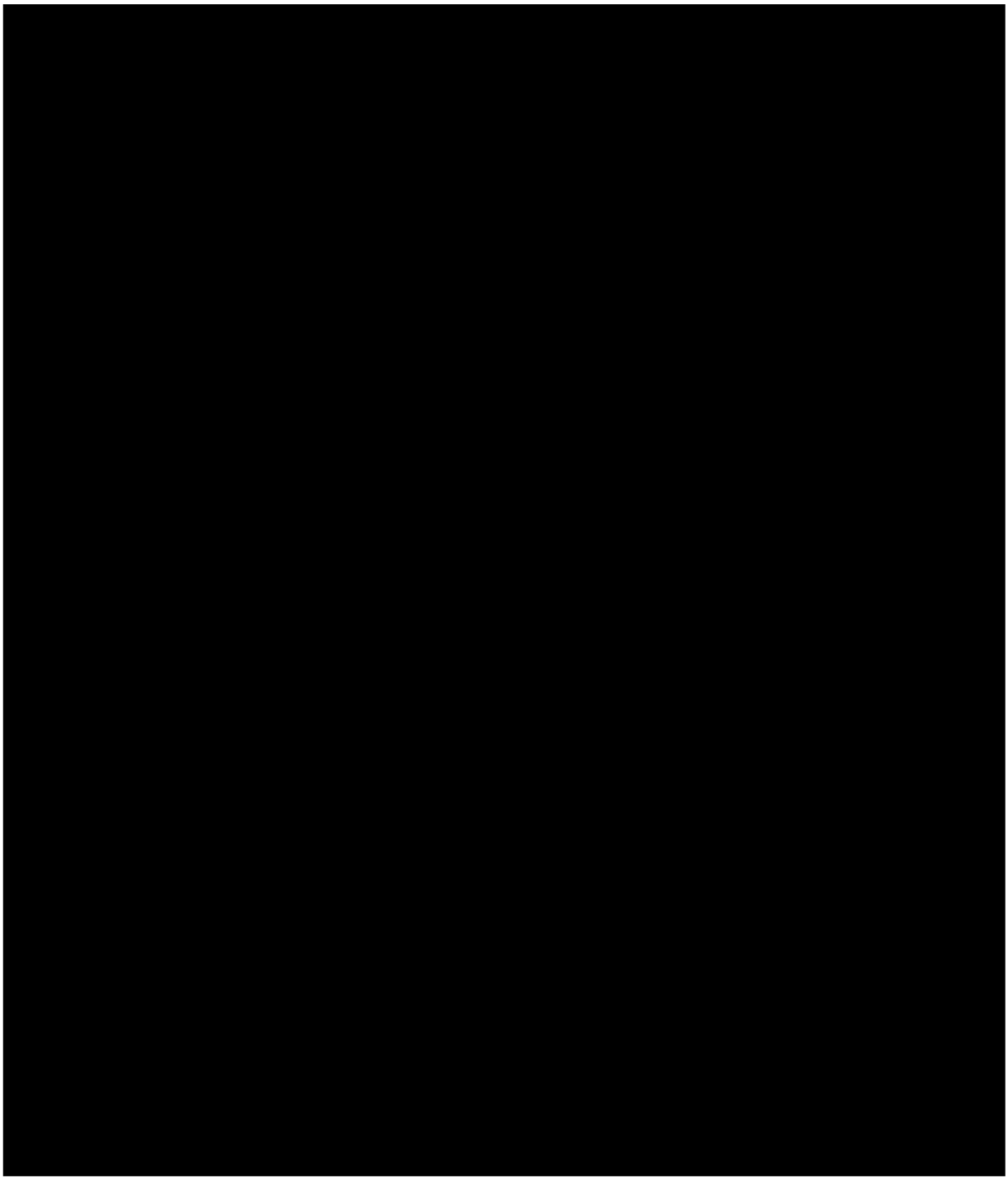
Abbreviations: BR = rituximab and bendamustine; KM = Kaplan-Meier; OS = overall survival

**Figure 47 Visual inspection of independently fitted parametric vs KM curves for OS of BR (same as the CS)**



Abbreviations: BR = rituximab and bendamustine; CS = company submission; KM = Kaplan-Meier; OS = overall survival

**Figure 48 Visual inspection of independently fitted parametric vs KM curves for TTD for zanubrutinib (after matching BR)**



Abbreviations: BR = rituximab and bendamustine; KM = Kaplan-Meier; TTD = time to discontinuation

**Table 37. Mean and hazard patterns of OS; zanubrutinib (after matching BR)[a]**

| | Exponential | Weibull | Gompertz | Log-normal | Log-logistic | Gamma |
|---|---|---|---|---|---|---|
| Jointly fitted models | | | | | | |
| Mean (year) | 14.19 | 13.44 | 14.68 | 14.17 | 13.64 | 13.30 |
| Hazard pattern | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically increasing | Monotonically decreasing | Increasing for first 1 year and decreasing | Increasing for first 4 years and decreasing | Monotonically increasing |
| After adjusting for background mortality | Constant for first 8 years and then increasing | Constant for first 12 years and increasing | Decreasing in first 6 years, and then increasing | Increasing in first 1.5 years and decreasing until 7th year and increasing | Increasing for 2 years and constant until 10th year and then increasing | Increasing for first 2 years and constant until 12th year and then increasing |
| Independently fitted models | | | | | | |
| Mean (year) | 14.19 | 9.68 | 3.63 | 14.29 | 11.97 | 11.29 |
| Hazard pattern | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically increasing | Increasing for first 8 years and decreasing at a certain level and constant | Increasing for first 1 year and decreasing | Increasing for first 6 years and decreasing | Monotonically increasing |
| After adjusting for background mortality | Constant for first 8 years and increasing | Monotonically increasing | Increasing for first 8 years and decreasing at a certain level and constant | Increasing for first 1 year and decreasing until 6th year, then increasing | Increasing for 6 years and decreasing until 12th year and then increasing | Monotonically increasing |

Abbreviations: BR = bendamustine rituximab; OS = overall survival

a Unless otherwise specified, the survival estimates were after adjustment by background mortality such that in any time during the model, the morality rate of modelled population would not be lower than that of general population.

**Table 38. Mean and hazard patterns of OS; BR[a]**

| | Exponential | Weibull | Gompertz | Log-normal | Log-logistic | Gamma |
|---|---|---|---|---|---|---|
| Jointly fitted models | | | | | | |
| Mean (year) | 8.34 | 7.35 | 9.92 | 10.52 | 9.06 | 7.39 |
| Hazard pattern | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically increasing | Monotonically decreasing | Increasing for first 1 year and decreasing | Increasing for first 2 years and decreasing | Monotonically increasing |
| After adjusting for background mortality | Constant for first 18 years and then increasing | Increasing for first 4 years and constant until 20$^{th}$ year, and then increasing | Decreasing in first 12 years, and then increasing | Increasing in first 0.5 year and decreasing until 11$^{th}$ year and increasing | Increasing for 2 years and decreasing until 13$^{th}$ year and then increasing | Increasing for first 2 years and constant until 20$^{th}$ year and then increasing |
| Independently fitted models | | | | | | |
| Mean (year) | 8.36 | 8.07 | 11.64 | 10.46 | 9.46 | 7.91 |
| Hazard pattern | | | | | | |
| Before adjusting for background mortality | Constant for first 18 years and increasing | Constant for first 18 years and increasing | Decreasing for first 6 years and increasing | Increasing for first 1 year and decreasing until 10$^{th}$ year and then increasing | Increasing for first 1 year and decreasing until 12$^{th}$ year and then increasing | Constant for first 18 years and increasing |
| After adjusting for background mortality | Constant | Monotonically increasing | Monotonically decreasing | Increasing for first 1 year and then decreasing | Increasing for first 1 year and then decreasing | Increasing for first 12 years and then constant |

Abbreviations: BR = bendamustine rituximab; OS = overall survival

[a] Unless otherwise specified, the survival estimates were after adjustment by background mortality such that in any time during the model, the morality rate of modelled population would not be lower than that of general population.

*Zanubrutinib (match DRC) vs DRC*

Results of the extrapolations of PFS, OS and TTD for the comparison of zanubrutinib (after matching DRC) to DRC are shown below, including: (1) summary of model selection, (2) fit statistics and visual inspection, and (3) mean (undiscounted) survival and hazard patterns. As the conclusions of the model selection are the same as those in the CS, only summary information are provided below without repeating rationales from the B.3.3.2 of the CS.

**Table 39. Summary of model selection for the pairwise comparison of zanubrutinib (after matching DRC) vs DRC**

| Outcome | Treatment | Base case setting | Justification for model selection in base case | Scenario analysis settings | Justification for model selection in scenario analyses |
|---|---|---|---|---|---|
| OS | Zanubrutinib (matching DRC) | Dependent gamma model | Clinically plausible mean OS for both treatments<br>Clinically plausible hazard patterns for both treatments<br>The second lowest BIC and close to the lowest BIC | Dependent Weibull model; dependent Gompertz model | Clinically plausible mean OS for both treatments<br>Clinically plausible hazard patterns for both treatments<br>The third and fourth lowest BIC |
| | DRC | | | | |
| PFS | Zanubrutinib (matching DRC) | Dependent exponential model | The lowest BIC<br>Alignment with TTD in parametric distribution | None | For both PFS and TTD, the exponential distribution was consistently associated with obviously lower BIC compared to the other distributions. |
| | DRC | | | | |
| TTD | Zanubrutinib (matching DRC) | Independent exponential model | The lowest BIC<br>Alignment with PFS in parametric distribution | | |
| | DRC | NA | NA | NA | NA |

Abbreviations: BIC = Bayesian information criteria; DRC = dexamethasone, rituximab, cyclophosphamide; NA = not applicable; OS = overall survival; PFS = progression-free survival; TTD = time to treatment discontinuation

**Table 40. Fit statistics for jointly fitted PFS, OS and TTD; zanubrutinib (after matching DRC) vs DRC**

| Parametric distribution | Investigator-assessed PFS | | OS | | TTD | |
|---|---|---|---|---|---|---|
| | AIC | BIC | AIC | BIC | AIC | BIC |
| Exponential | 651.158 | 657.7646 | 486.8123 | 493.419 | Not applicable | |
| Weibull | 652.6724 | 662.5823 | 488.4383 | 498.3482 | | |
| Gompertz | 651.7494 | 661.6593 | 488.5406 | 498.4506 | | |
| Log-normal | 654.8508 | 664.7607 | 489.3848 | 499.2947 | | |
| Log-logistic | 652.4242 | 662.3341 | 488.8507 | 498.7606 | | |
| Gamma | 652.8314 | 662.7413 | 488.4045 | 498.3145 | | |

Source: ASPEN patient-level data, August 2019 data cut
Abbreviations: AIC = Akaike information criteria; BIC = Bayesian information criteria; DRC = dexamethasone, rituximab, cyclophosphamide; OS = overall survival; PFS = progression-free survival; TTD = time to discontinuation
Note: The lowest AIC or BIC is in *italics*.

**Table 41. Fit statistics for independently fitted PFS, OS and TTD; zanubrutinib (after matching DRC)**

| Parametric distribution | Investigator-assessed PFS | | OS | | TTD | |
|---|---|---|---|---|---|---|
| | AIC | BIC | AIC | BIC | AIC | BIC |
| Exponential | 137.2143 | 140.0741 | 69.9872 | 72.8471 | 178.5063 | 181.3584 |
| Weibull | 138.2842 | 144.0038 | 71.9821 | 77.7017 | 179.6665 | 185.3706 |
| Gompertz | 136.6418 | 142.3615 | 71.5014 | 77.221 | 179.0451 | 184.7491 |
| Log-normal | 136.938 | 142.6576 | 71.3973 | 77.1169 | 178.3706 | 184.0746 |
| Log-logistic | 138.0506 | 143.7702 | 71.9265 | 77.6461 | 179.4422 | 185.1462 |
| Gamma | 138.3948 | 144.1144 | 71.9857 | 77.7053 | 179.7615 | 185.4656 |

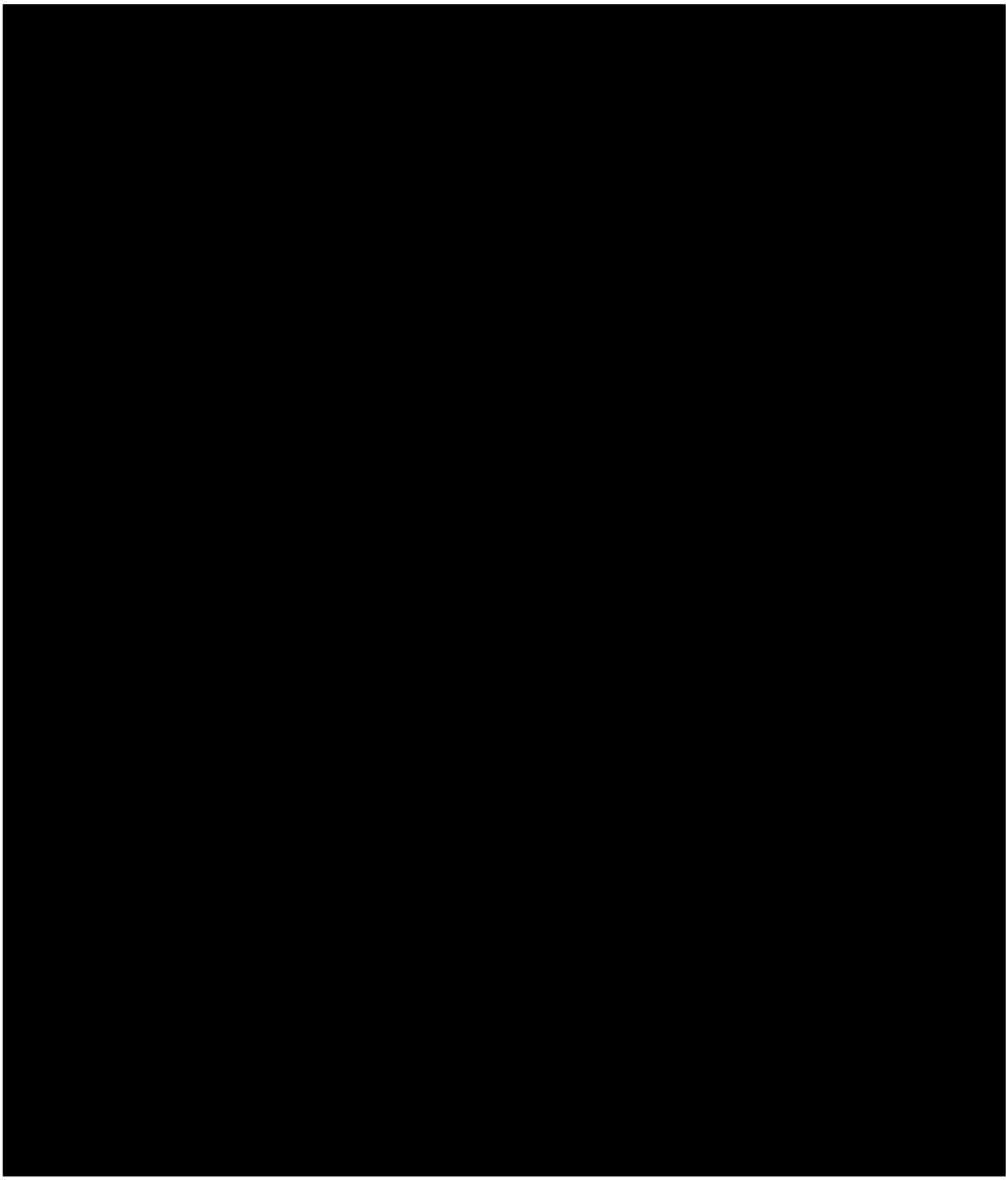
Source: ASPEN patient-level data, August 2019 data cut
Abbreviations: AIC = Akaike information criteria; BIC = Bayesian information criteria; DRC = dexamethasone, rituximab, cyclophosphamide; OS = overall survival; PFS = progression-free survival; TTD = time to discontinuation
Note: The lowest AIC or BIC is in *italic*s.

**Table 42. Fit statistics for independently fitted PFS, OS and TTD; DRC (same as the CS)**

| Parametric distribution | Investigator-assessed PFS | | OS | | TTD | |
|---|---|---|---|---|---|---|
| | AIC | BIC | AIC | BIC | AIC | BIC |
| Exponential | 513.9437 | 516.2204 | 416.8251 | 419.1018 | Not applicable | |
| Weibull | 515.8478 | 520.4011 | 418.3617 | 422.915 | | |
| Gompertz | 514.9402 | 519.4935 | 418.4837 | 423.0371 | | |
| Log-normal | 515.409 | 519.9624 | 418.4984 | 423.0518 | | |
| Log-logistic | 514.1105 | 518.6639 | 418.5687 | 423.122 | | |
| Gamma | 515.9287 | 520.482 | 418.3081 | 422.8614 | | |

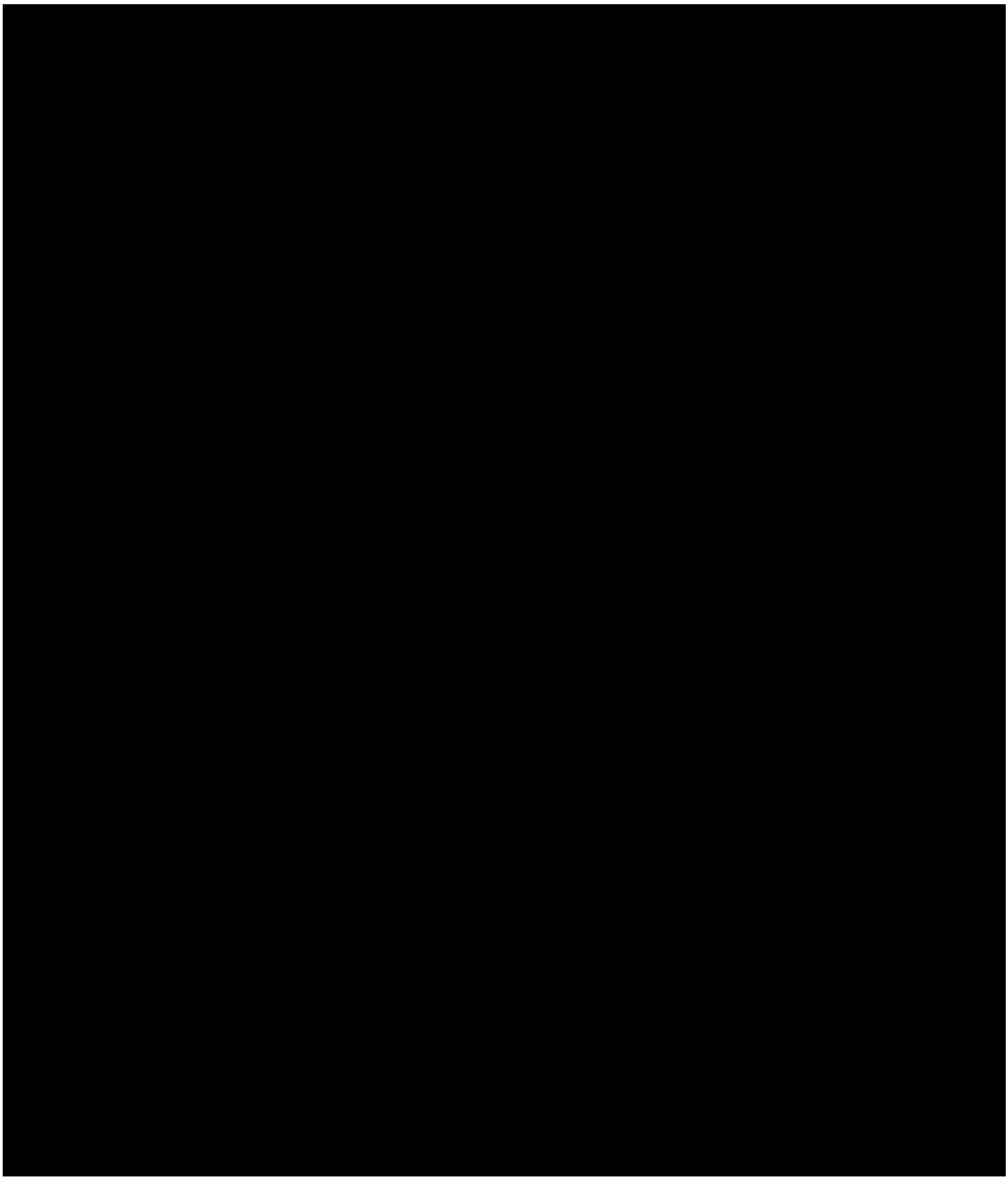
Source: ASPEN patient-level data
Abbreviations: AIC = Akaike information criteria; BIC = Bayesian information criteria; CS = company submission; DRC = dexamethasone, rituximab, cyclophosphamide; OS = overall survival; PFS = progression-free survival; TTD = time to discontinuation
Note: The lowest AIC or BIC is in *italics.*

**Figure 49. Visual inspection of jointly fitted parametric vs KM curves for PFS; zanubrutinib (after matching DRC) vs DRC**



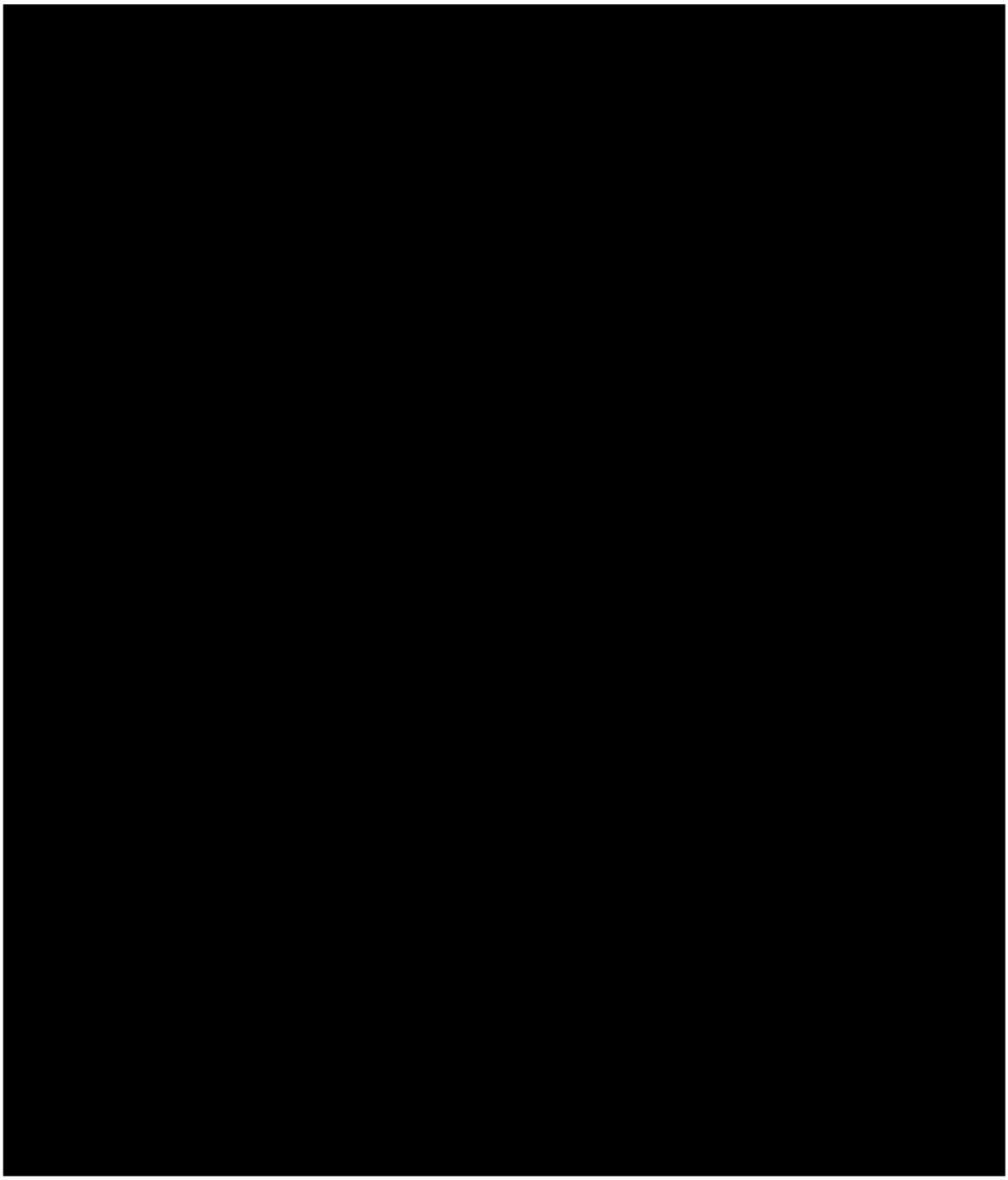
Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; PFS = progression-free survival

**Figure 50. Visual inspection of independently fitted parametric vs KM curves for PFS; zanubrutinib (after matching DRC)**



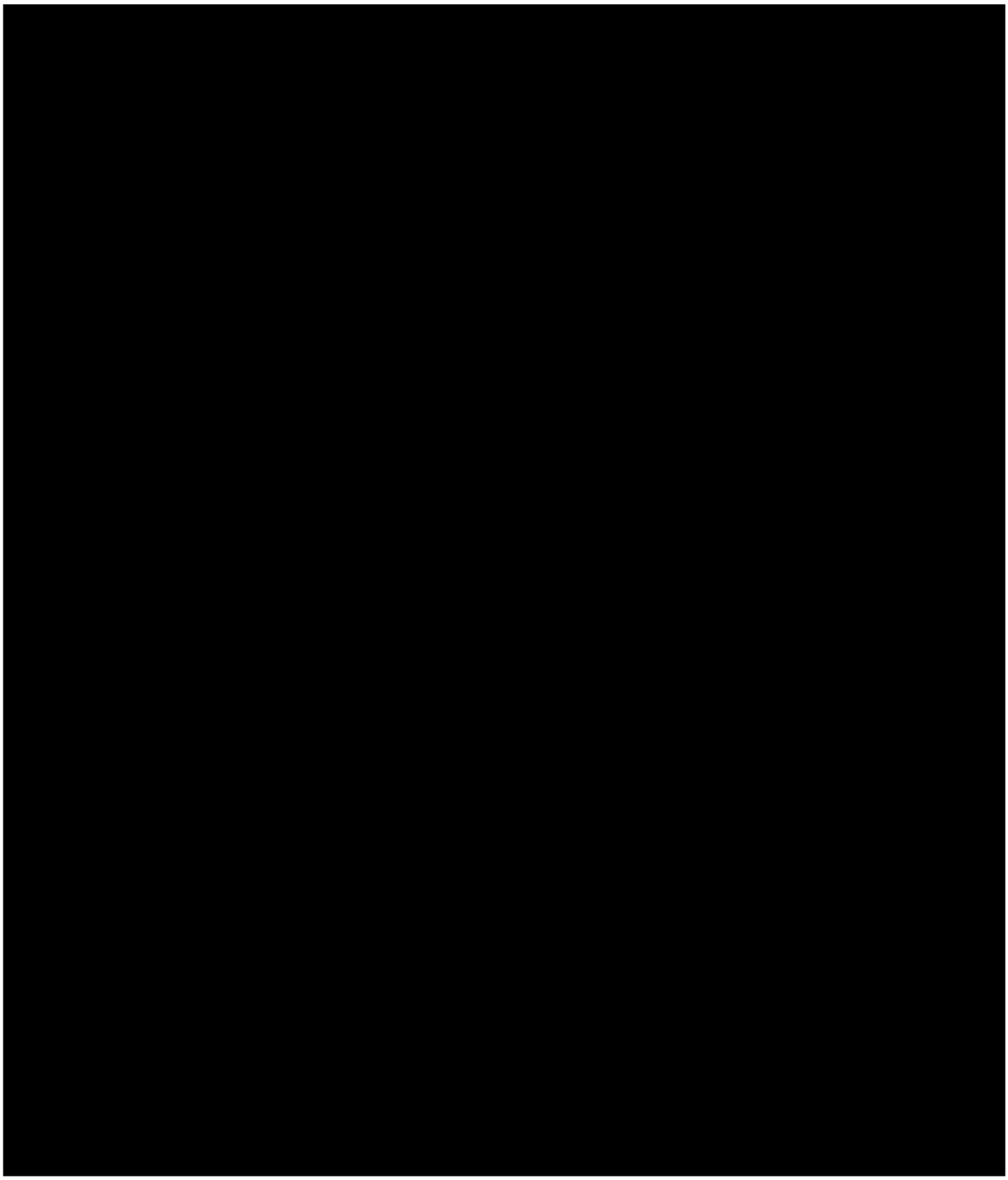
Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; PFS = progression-free survival

**Figure 51. Visual inspection of independently fitted parametric vs KM curves for PFS, DRC (same as the CS)**



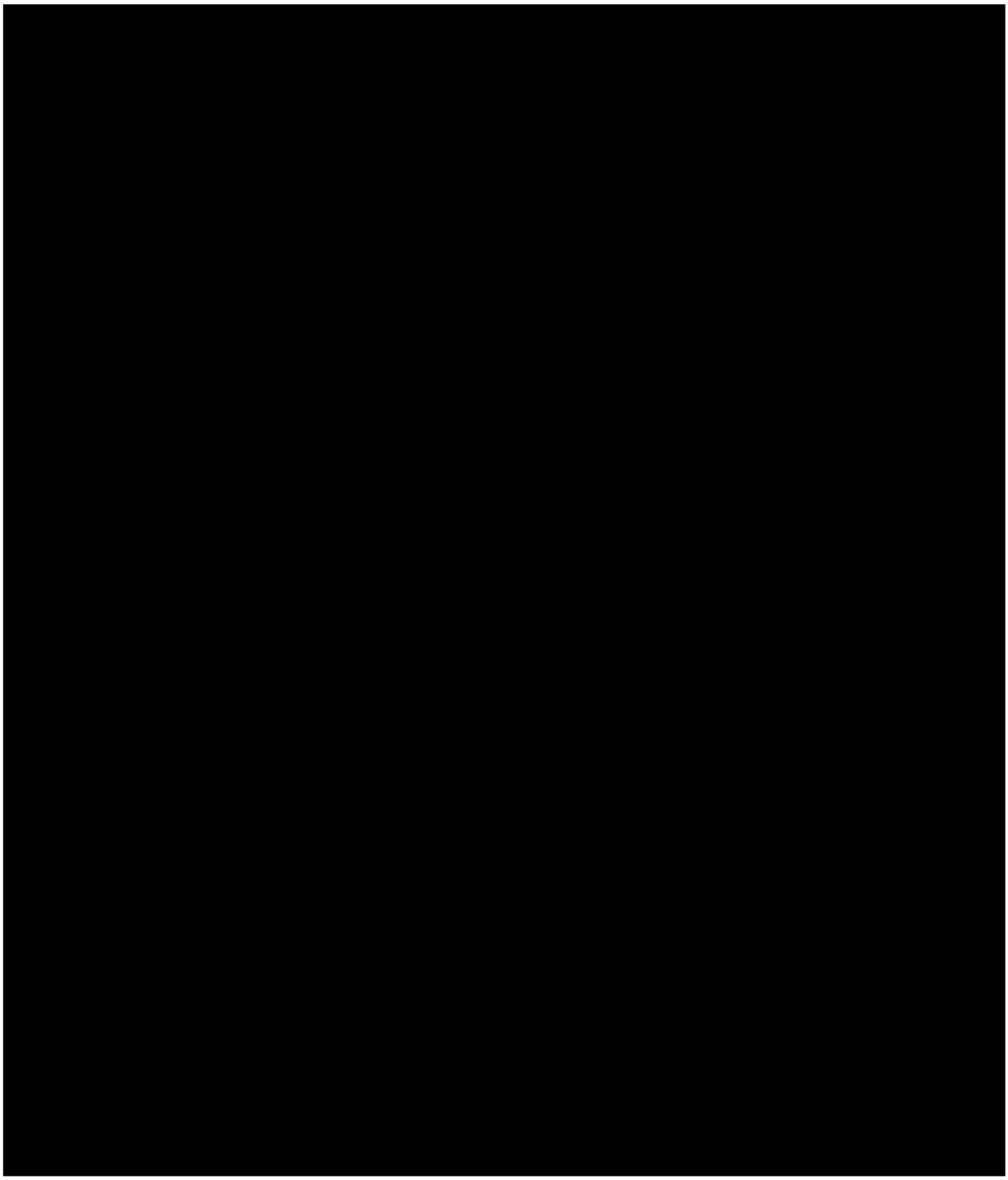
Abbreviations: CS = company submission; DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; PFS = progression-free survival

**Figure 52. Visual inspection of jointly fitted parametric vs KM curves for OS; zanubrutinib (after matching DRC) vs DRC**



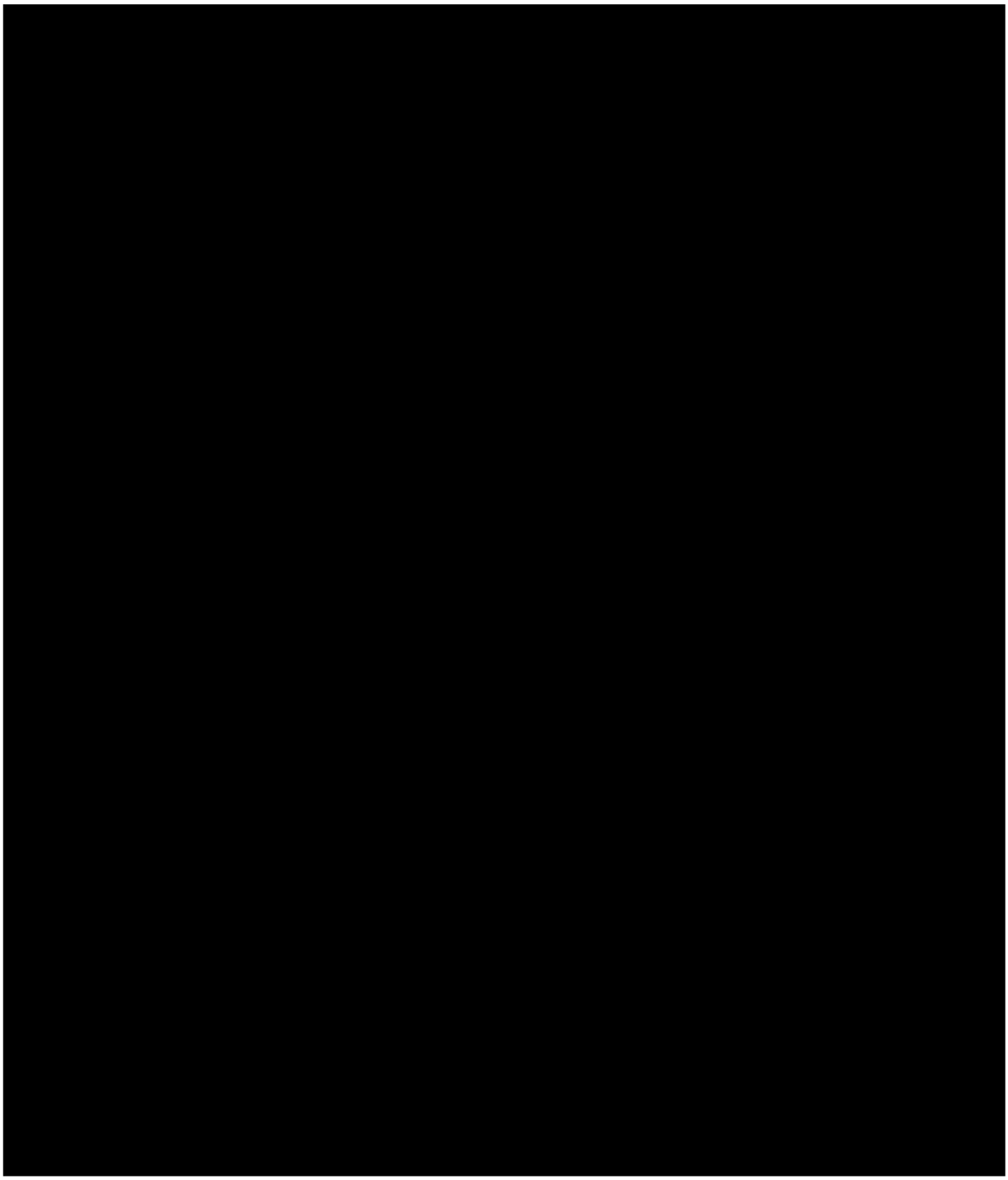
Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; OS = overall survival

**Figure 53. Visual inspection of independently fitted parametric vs KM curves for OS; zanubrutinib (after matching DRC)**



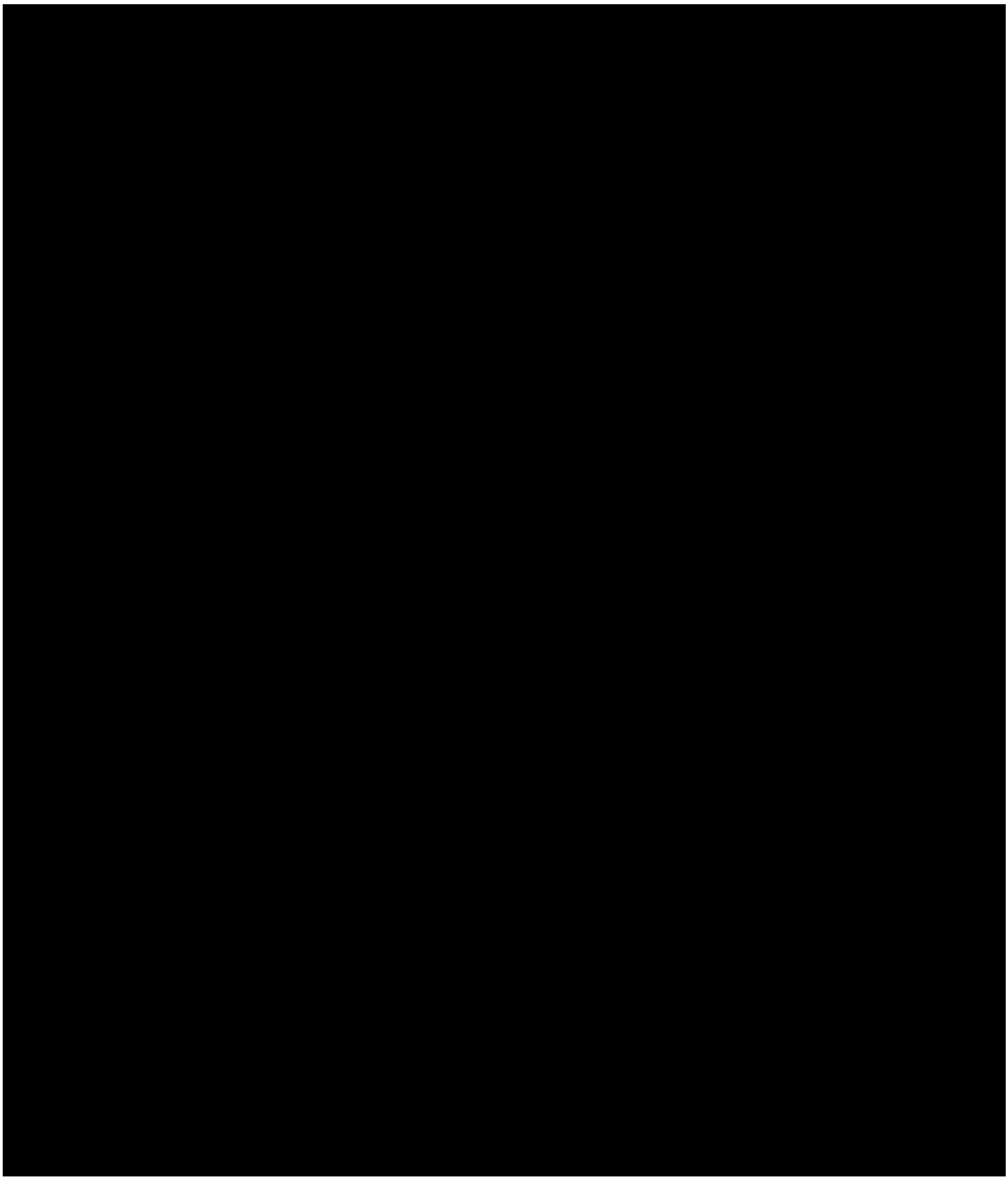
Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; OS = overall survival

**Figure 54. Visual inspection of independently fitted parametric vs KM curves for OS; DRC (same as the CS)**



Abbreviations: CS = company submission; DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; OS = overall survival

**Figure 55. Visual inspection of independently fitted parametric vs KM curves for TTD; zanubrutinib (after matching DRC)**



Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; TTD = time to discontinuation

**Table 43. Mean and hazard patterns of OS; zanubrutinib (after matching DRC)[a]**

| | Exponential | Weibull | Gompertz | Log-normal | Log-logistic | Gamma |
|---|---|---|---|---|---|---|
| Jointly fitted models | | | | | | |
| Mean (year) | 14.47 | 13.89 | 13.89 | 13.35 | 13.60 | 13.83 |
| Hazard pattern | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically increasing | Monotonically increasing | Increasing for first 2 years and then decreasing | Increasing for first 5 years and then decreasing | Monotonically increasing but very slowly after 10th year |
| After adjusting for background mortality | Constant for first 7 years, then increasing | Increasing for first 2 years and stable until 10th year, then increasing | Increasing slowly for first 12 years and then increasing steeply | Increasing for first 2 years and then decreasing slowly until 10th year, then increasing | Increasing for first 2 years, then constant until 10th year, and increasing | Increasing very slowly for first 2 years, and constant until 10th year, and increasing |
| Independently fitted models | | | | | | |
| Mean (year) | 14.47 | 14.61 | 15.30 | 14.98 | 14.75 | 14.55 |
| Hazard pattern | | | | | | |
| Before adjusting for background mortality | Constant | Decreasing monotonically | Decreasing monotonically and reaching 0% after 6th year | Increasing steeply at beginning and decreasing | Decreasing monotonically | Decreasing monotonically |
| After adjusting for background mortality | Constant for first 8 years, then increasing | Decreasing very slowly in first 6 years, and then increasing | Decreasing until 2nd year, and then increasing | Decreasing until 4th year, and then increasing | Decreasing slowly until 6th year, and increasing | Decreasing slowly until 6th year, and increasing |

Source: ASPEN patient-level data, August 2019 data cut
Abbreviations: DRC = dexamethasone, rituximab, cyclophosphamide; OS = overall survival
[a] Unless otherwise specified, the survival estimates were after adjustment by background mortality such that in any time during the model, the morality rate of modelled population would not be lower than that of general population.

**Table 44. Mean and hazard patterns of OS; DRC[a]**

| | Exponential | Weibull | Gompertz | Log-normal | Log-logistic | Gamma |
|---|---|---|---|---|---|---|
| Jointly fitted models | | | | | | |
| Mean (year) | 10.09 | 9.81 | 9.39 | 10.80 | 10.51 | 9.88 |
| Hazard pattern | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically increasing | Monotonically increasing | Increasing for first 1 year and then decreasing | Increasing for first 3 years and then decreasing | Monotonically increasing but very slowly after 12th year |
| After adjusting for background mortality | Constant for first 16 years, then increasing | Increasing for first 2 years and then constant until 17th | Monotonically increasing | Increasing steeply for first 1 year, then decreasing until 12th | Increasing for first 2 years, then decreasing until 12th year, and increasing | Increasing for first 2 years, then constant until 16th year, and increasing |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | year, then increasing | | year, then increasing | | |
| Independently fitted models | | | | | | |
| Mean (year) | 10.05 | 9.76 | 9.32 | 10.58 | 10.42 | 9.79 |
| Hazard pattern | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically increasing | Monotonically increasing | Increasing in the first 1.5 years; then decreasing | Increasing in the first 3 years; then decreasing | Monotonically increasing |
| After adjusting for background mortality | Constant in the first 15 years; then increasing | Monotonically increasing | Monotonically increasing | Increasing in the first 1 year; then decreasing for 12 years; then increasing | Increasing in the first 3 years; then decreasing for 11 years; then increasing | Monotonically increasing |

Source: ASPEN patient-level data, August 2019 data cut
Abbreviations: DRC = dexamethasone, rituximab, cyclophosphamide; OS = overall survival
[a] Unless otherwise specified, the survival estimates were after adjustment by background mortality such that in any time during the model, the morality rate of modelled population would not be lower than that of general population.

Of note, this scenario analysis of using pooled data of cohorts 1 and 2 did not include the updates of the AE incidences and duration, given (1) the results of the base-case and sensitivity analyses from the initial CS and (2) the conclusion from the response to clarification question B10, which showed that AEs had very minimal impact on the cost-effectiveness model results.

Similar to the safety inputs, this scenario analysis of using pooled data of cohorts 1 and 2 did not include the updates to baseline patient characteristics, as the results from the CS showed that these inputs had very minimal impact on the cost-effectiveness model results.

*Part 2: Updated HRQoL inputs*

The updated HRQoL analyses included 201 patients (102 patients in the zanubrutinib arm, 99 in the ibrutinib arm) from cohort 1 and 28 patients (all treated with zanubrutinib) from cohort 2. Patients without at least one complete measurement are excluded from the analysis. Additional exclusion includes one patient not treated but having baseline measurement at screening. A total of 998 observations from 220 patients are used in the following modelling.

Model 1 reports the lower AIC and BIC among all the regression models and hence the pre-progression health state utility value from Model 1 (0.7841) is recommended for use in the cost-effectiveness model.

**Table 45. Summary of pre-progression health state utility values (using pooled data of cohorts 1 and 2)**

| | Model 1 | | Model 2 | | Model 3 | |
|---|---|---|---|---|---|---|
| | **Zanubrutinib (N=127)** | **Ibrutinib (N=93)** | **Zanubrutinib (N=127)** | **Ibrutinib (N=93)** | **Zanubrutinib (N=127)** | **Ibrutinib (N=93)** |
| LS Mean (SE)* | 0.7805 (0.0150) | 0.7891 (0.0175) | 0.7827 (0.0061) | 0.7905 (0.0070) | 0.7825 (0.0061) | 0.7903 (0.0070) |
| AIC | -985.7 | | -978.2 | | -983.7 | |
| BIC | -975.5 | | -968.0 | | -973.6 | |
| Weighted LS Mean across treatment (SE)** | 0.7841 (0.0114) | | 0.7860 (0.0046) | | 0.7858 (0.0046) | |

*LS Means are adjusted at mean age of 69.16 years old, and average of 65.5% male in the population. And for model 2 and 3, the LS Means are weighted average of the LS Means at scheduled time points, with weights as the number of observations at each time point divided by the total number of observations within each treatment arm.
**Weights are the proportions of patients in each treatment arm.

**Table 46. Detailed information of model 1 (using pooled data of cohorts 1 and 2)**

| | **Coefficient** | **SE** | **Df** | **t statistics** | **p-value** |
|---|---|---|---|---|---|
| Intercept | 1.0248 | 0.08570 | 213 | 11.96 | <.0001 |
| Tx_zanu | -0.00858 | 0.02309 | 215 | -0.37 | 0.7107 |
| Age | -0.00401 | 0.001146 | 215 | -3.50 | 0.0006 |
| Gender (male) | 0.06332 | 0.02414 | 215 | 2.62 | 0.0094 |
| | | | | | |
| **Number of observations** | | 998 | **Number of patients** | | 220 |
| **AIC** | | -985.7 | **BIC** | | -975.5 |

**Table 47. Detailed information of model 2 (using pooled data of cohorts 1 and 2)**

| | **Coefficient** | **SE** | **Df** | **t statistics** | **p-value** |
|---|---|---|---|---|---|
| Intercept | 0.9955 | 0.08595 | 218 | 11.58 | <.0001 |
| Tx_zanu | -0.00708 | 0.02304 | 216 | -0.31 | 0.7588 |
| Age | -0.00389 | 0.001144 | 216 | -3.40 | 0.0008 |
| Gender (male) | 0.06287 | 0.02409 | 216 | 2.61 | 0.0097 |
| Days from treatment initiation ($\boldsymbol{day_t}$) – numerical | 0.000084 | 0.000024 | 318 | 3.48 | 0.0006 |
| | | | | | |
| **Number of observations** | | 998 | **Number of patients** | | 220 |
| **AIC** | | -978.2 | **BIC** | | -968.0 |

**Table 48. Detailed information of model 3 (using pooled data of cohorts 1 and 2)**

| | **Coefficient** | **SE** | **Df** | **t statistics** | **p-value** |
|---|---|---|---|---|---|
| Intercept | 0.9962 | 0.08597 | 218 | 11.59 | <.0001 |
| Tx_zanu | -0.00713 | 0.02304 | 216 | -0.31 | 0.7572 |
| Age | -0.00389 | 0.001144 | 216 | -3.40 | 0.0008 |
| Gender (male) | 0.06283 | 0.02409 | 216 | 2.61 | 0.0097 |
| Completed treatment cycles ($visit_t$) – numerical | 0.002302 | 0.000696 | 313 | 3.31 | 0.0011 |
| | | | | | |
| **Number of observations** | | 998 | **Number of patients** | | 220 |
| **AIC** | | -983.7 | **BIC** | | -973.6 |

**Table 49. ASPEN EQ-5D-5L results by Cycle/Day (using pooled data of cohorts 1 and 2)**

| | Zanubrutinib (N = 130) | Ibrutinib (N = 99) |
|---|---|---|
| **Screening** | | |
| n | 79 | 61 |
| Mean (SD) | 0.7301 (0.18498) | 0.7335 (0.20686) |
| Median | 0.7350 | 0.7360 |
| Q1, Q3 | 0.6470, 0.8770 | 0.6350, 0.8790 |
| Min, Max | 0.167, 1.000 | 0.064, 1.000 |
| **Cycle 4 Day 1** | | |
| n | 91 | 63 |
| Mean (SD) | 0.8117 (0.18699) | 0.7820 (0.22808) |
| Median | 0.8370 | 0.8370 |
| Q1, Q3 | 0.6980, 1.0000 | 0.6790, 1.0000 |
| Min, Max | 0.294, 1.000 | -0.032, 1.000 |
| **Cycle 7 Day 1** | | |
| n | 101 | 72 |
| Mean (SD) | 0.7984 (0.21643) | 0.7788 (0.20681) |
| Median | 0.8370 | 0.7680 |
| Q1, Q3 | 0.7110, 1.0000 | 0.6535, 1.0000 |
| Min, Max | 0.057, 1.000 | 0.169, 1.000 |
| **Cycle 10 Day 1** | | |
| n | 99 | 74 |
| Mean (SD) | 0.8085 (0.19286) | 0.8083 (0.23029) |
| Median | 0.8360 | 0.8425 |
| Q1, Q3 | 0.6930, 1.0000 | 0.7080, 1.0000 |
| Min, Max | -0.173, 1.000 | -0.202, 1.000 |
| **Cycle 13 Day 1** | | |
| n | 99 | 79 |
| Mean (SD) | 0.8127 (0.18154) | 0.8151 (0.18320) |
| Median | 0.8360 | 0.8370 |
| Q1, Q3 | 0.6950, 1.0000 | 0.7270, 1.0000 |
| Min, Max | 0.155, 1.000 | 0.231, 1.000 |
| **Cycle 19 Day 1** | | |
| n | 72 | 60 |
| Mean (SD) | 0.7759 (0.18760) | 0.8149 (0.21722) |
| Median | 0.7730 | 0.8625 |
| Q1, Q3 | 0.6590, 1.0000 | 0.7155, 1.0000 |
| Min, Max | 0.155, 1.000 | -0.098, 1.000 |
| **Cycle 25 Day 1** | | |
| n | 23 | 17 |

| | Zanubrutinib (N = 130) | Ibrutinib (N = 99) |
|---|---|---|
| Mean (SD) | 0.7958 (0.17036) | 0.8119 (0.26089) |
| Median | 0.8370 | 0.9060 |
| Q1, Q3 | 0.6790, 0.9060 | 0.7400, 1.0000 |
| Min, Max | 0.304, 1.000 | 0.057, 1.000 |
| **Cycle 31 Day 1** | | |
| n | 3 | 4 |
| Mean (SD) | 0.8513 (0.13543) | 0.9290 (0.08376) |
| Median | 0.8190 | 0.9395 |
| Q1, Q3 | 0.7350, 1.0000 | 0.8580, 1.0000 |
| Min, Max | 0.735, 1.000 | 0.837, 1.000 |
| **End of Treatment** | | |
| n | 10 | 7 |
| Mean (SD) | 0.4883 (0.29488) | 0.7170 (0.18429) |
| Median | 0.5525 | 0.7110 |
| Q1, Q3 | 0.1550, 0.6830 | 0.5550, 0.8770 |
| Min, Max | 0.090, 0.906 | 0.451, 1.000 |

*Part 3: Updated cost-effectiveness model outputs*

The results presented in Table 50 and Table 51 demonstrate that the ICERs were very close between using cohort 1 data and pooled data of cohorts 1 and 2.

**Table 50. Scenario analysis: use of cohort 1 data versus pooled cohorts 1 and 2 for the comparison with BR**

| | **ICER (pairwise comparison with BR)** |
|---|---|
| Cohort 1 data (CS base-case analysis) – independent exponential for zanubrutinib (match BR), independent Weibull for BR | [redacted] |
| Cohort 1 data (CS scenario analysis) – dependent Weibull | [redacted] |
| Cohort 1 data (CS scenario analysis) – dependent gamma | [redacted] |
| Pooled data of cohorts 1 and 2 (Clarification-base-case analysis) – independent exponential for zanubrutinib (match BR), independent Weibull for BR | [redacted] |
| Pooled data of cohorts 1 and 2 (Clarification-scenario analysis) - dependent Weibull | [redacted] |
| Pooled data of cohorts 1 and 2 (Clarification-scenario analysis) - dependent gamma | [redacted] |

Abbreviations: BR = rituximab and bendamustine; ICER = incremental cost-effectiveness ratio
Notes: For transparency and simplicity, the results above were based on the base-case analysis in the CS, revised for this specific clarification question, without accounting for the revisions of model inputs/programming in response to other clarification questions. To generate the results according to different combinations of revisions, please refer to the updated Excel model

**Table 51. Scenario analysis: use of cohort 1 data versus pooled cohorts 1 and 2 for the comparison with DRC**

| | ICER (pairwise comparison with DRC) |
|---|---|
| Cohort 1 data (CS base-case analysis) –dependent gamma | [redacted] |
| Cohort 1 data (CS scenario analysis) – dependent Gompertz | |
| Cohort 1 data (CS scenario analysis) – dependent Weibull | |
| Pooled data of cohorts 1 and 2 (Clarification-base-case analysis) – dependent gamma | |
| Pooled data of cohorts 1 and 2 (Clarification-scenario analysis) - dependent Gompertz | [redacted] |
| Pooled data of cohorts 1 and 2 (Clarification-scenario analysis) - dependent Weibull | [redacted] |

Abbreviations: AE = adverse event; DRC = dexamethasone, rituximab and cyclophosphamide; ICER = incremental cost-effectiveness ratio

Notes: For transparency and simplicity, the results above were based on the base-case analysis in the CS, revised for this specific clarification question, without accounting for the revisions of model inputs/programming in response to other clarification questions. To generate the results according to different combinations of revisions, please refer to the updated Excel model

Of note, the results above did not capture additional costs of the MYD88 assessment, because according to clinical experts, testing for MYD88 mutation is the standard of care at the majority of the 24 British WM centres (covering 90% of all WM patients since 2016 in the UK).

**B7. Priority Question: The company submission provides detailed description of the company's methods used for survival analysis and their validation efforts, and steps undertaken are largely in line with NICE DSU TSD 14 and 21. However:**

a. **Please explain why the generalised gamma distribution was not explored in any comparison. Please also explore the use of the**

**generalised gamma and compare its statistical fit and validity to those of the other distributions (and incorporate it in the model).**

b. **Please provide all KM plot figures for OS, PFS and TTD in all comparisons with numbers of patients at risk included for the full duration of follow-up.**

c. **Please provide figures showing the fit of all survival distributions in one plot for OS, PFS and TTD for all comparisons. This will help appreciate differences between the different distributions.**

d. **In the comparison with DRC, differential distributions are used to extrapolate PFS and OS. Please justify that this is reflective of the disease, supporting this with expert opinion.**

e. **For the comparison with BR, the company bases model selection (and whether to use dependent or independent modelling) for the comparator arm on OS expectations from England. The ERG questions whether this is appropriate because for this comparison, the ASPEN study is matched to the Tedeschi et al study and extrapolations should therefore reflect the population in the Tedeschi study, not expectations from the England population. The Tedeschi study is an EU-based study and according to the company, OS varies between countries. Furthermore, UK based OS expectations would be based on a mix of DRC, BR and others, which means that not all individual comparators would be required to fit any average OS expectations. The ERG therefore questions whether it was appropriate to rule out dependent models for this comparison with the given reasoning. Please explain the rationale for the company's approach and comment on its appropriateness.**

f. **If possible, please provide expert opinion for estimated OS for patients treated with BR and patients treated with DRC in the second-line setting.**

The generalised gamma distribution was considered for inclusion at an earlier stage of trial data analyses and model development but failed to converge for several treatments and outcomes (e.g., ibrutinib OS, zanubrutinib PFS after matching DRC, zanubrutinib time to treatment discontinuation (TTD) after matching DRC), likely driven by (1) the immaturity of the survival data, (2) relatively lower (effective) sample size after matching adjustment (on top of the data immaturity), and (3) relatively higher number of parameters (compared with other parametric distributions such as exponential, Weibull, etc.). Given the above, the generalised gamma distribution was not further assessed.

The KM figures presented in Section B.3.3 of the initial company submission are reproduced and presented below with the addition of the number of patients at risk.

**Figure 56. KM curves of PFS – zanubrutinib vs ibrutinib (Figure B.3.3)**

Abbreviations: KM = Kaplan-Meier; PFS = progression-free survival

**Figure 57. KM curves of OS – zanubrutinib vs ibrutinib (Figure B.3.4)**

Abbreviations: KM = Kaplan-Meier; OS = overall survival

**Figure 58. KM curves of TTD – zanubrutinib vs ibrutinib (Figure B.3.5)**

Abbreviations: KM = Kaplan-Meier; TTD = time to discontinuation

**Figure 59. KM curves of PFS – zanubrutinib vs DRC (Figure B.3.12)**

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; PFS = progression-free survival

**Figure 60. KM curves of OS – zanubrutinib vs DRC (Figure B.3.13)**

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; OS = overall survival

**Figure 61. KM curves of TTD – zanubrutinib (Figure B.3.14)**

Abbreviations: KM = Kaplan-Meier; TTD = time to discontinuation

**Figure 62. KM curves of PFS – zanubrutinib vs BR (Figure B.3.24)**

Abbreviations: BR = rituximab and bendamustine; KM = Kaplan-Meier; PFS = progression-free survival

**Figure 63. KM curves of OS – zanubrutinib vs BR (Figure B.3.25)**

Abbreviations: BR = rituximab and bendamustine; KM = Kaplan-Meier; OS = overall survival

**Figure 64. KM curves of TTD – zanubrutinib (Figure B.3.26)**

Abbreviations: KM = Kaplan-Meier; TTD = time to discontinuation

Figures showing the fit of all survival distributions in one plot are presented below for OS, PFS and TTD for all independently fitted curves (jointly fitted curves are not shown due to difficulties in visualisation of one plot). It should be noted that some plots were generated at an earlier stage of trial data analysis and hence include the generalised gamma distribution, but were not presented in the company submission as it was later decided later to present the 95% CI of each individual curve, in order to assess the structural stability of each parametric model in the case of immature data.

**Figure 65. Visual inspection of independently fitted parametric vs KM curves for PFS – zanubrutinib (match DRC; Figure B.3.18)**

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; PFS = progression-free survival

**Figure 66. Visual inspection of independently fitted parametric vs KM curves for PFS – DRC (Figure B.3.19)**

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; PFS = progression-free survival

**Figure 67. Visual inspection of independently fitted parametric versus KM curves for OS – zanubrutinib (match DRC; Figure B.3.21)**

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; OS = overall survival

**Figure 68. Visual inspection of independently fitted parametric versus KM curves for OS – DRC (Figure B.3.22)**

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; OS = overall survival

**Figure 69. Visual inspection of independently fitted parametric versus KM curves for TTD – zanubrutinib (match DRC; Figure B.3.23)**

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide; KM = Kaplan-Meier; TTD = time to discontinuation

**Figure 70. Visual inspection of independently fitted parametric vs KM curves for PFS – zanubrutinib (match BR; Figure B.3.31)**

Abbreviations: BR = rituximab and bendamustine; KM = Kaplan-Meier; PFS = progression-free survival

**Figure 71. Visual inspection of independently fitted parametric versus KM curves for PFS – BR (Figure B.3.32)**

Abbreviations: BR = rituximab and bendamustine; KM = Kaplan-Meier; PFS = progression-free survival

**Figure 72. Visual inspection of independently fitted parametric versus KM curves for OS – zanubrutinib (match BR; Figure B.3.33)**

Abbreviations: BR = rituximab and bendamustine; KM = Kaplan-Meier; OS = overall survival

**Figure 73. Visual inspection of independently fitted parametric versus KM curves for OS – BR (Figure B.3.34)**

Abbreviations: BR = rituximab and bendamustine; KM = Kaplan-Meier; OS = overall survival

**Figure 74. Visual inspection of independently fitted parametric versus KM curves for TTD – zanubrutinib (match BR; Figure B.3.35)**

Abbreviations: BR = rituximab and bendamustine; KM = Kaplan-Meier; TTD = time to discontinuation

For the “justifications of differential distributions of PFS and OS for DRC”, it is important to highlight that “consistency in the distributions between PFS and OS” was not a criterion of the model selection. As discussed in Section B.3.3.2 of the company submission, the alignment in distributions between PFS and TTD was indeed taken into consideration to reflect the common clinical practice that disease progression usually results in a treatment discontinuation. However, between PFS and OS, the distributions or hazard patterns are not necessarily aligned, depending on specific diseases or treatment patterns (including both primary treatment and subsequent treatments). More specifically, as discussed in response to clarification question B5, due to a lack of treatment options for patients with WM after these patients progress on the current treatment, the hazard of death was expected be monotonically increasing over time. However, such a pattern was less relevant for PFS given the clinical association between PFS and TTD, as discussed above.

For the PFS and OS of DRC, as summarised in Table B.3.6 in the company submission, different assessment criteria were applied to inform the model selection for PFS and OS separately. The ERG questioned whether it was appropriate to rule

out dependent models for this comparison. It is important to note that dependent models were not ruled out for the comparison with BR. Instead, due to the known uncertainties of survival extrapolation, despite that the independent models were used in the base case, the dependent models were included as scenario analyses (with rationales presented in Table B.3.5 in the company submission).

The ERG also questioned whether the expected OS estimates in one country could be used to inform the OS in another country. It is acknowledged that, ideally, a UK-based study directly comparing zanubrutinib to standard-of-care therapy would be available. However, as discussed partially in the response to other clarification questions (e.g., A22, A25), due to a lack of existing evidence, further assessments of the studies identified from the SLR were conducted to identify studies likely to be most representative of the UK population (e.g. conducted in the EU, sufficient sample size, etc.).

The ERG also stated that “according to the company, OS varies between countries”. The company would like to clarify that it was reported in NICE TA491 that the median OS in WM ranged <4–12 years and that in a European chart review study (for which the manufacturer of ibrutinib has full access to the patient-level data),[24] considerable country-specific OS differences were noted. However, due to a lack of reporting of details in NICE TA491 (e.g., exact country-specific estimates), it was not possible to fully examine the findings of the chart review (e.g., country-specific sample size, patient characteristics, treatment patterns, relevance of this study [conducted a number of years ago] to more recent UK clinical practice) or rely on such findings to inform the decisions for the model produced for zanubrutinib. Instead, the company considered not only the findings reported in NICE TA491,[24] but also other published estimates (e.g., ESMO clinical practice guidelines for WM and the Phase 2 Study 1118E of ibrutinib)[24,25] to identify relevant information wherever available, and relied on multiple criteria (e.g., clinical expert opinion and the Phase 1/2 BGB-3111-AU-003 trial of zanubrutinib)[26] for model selection.

**B8. Priority Question: The OS hazards estimated using survival analysis only based on the ASPEN trial and MAICs fall below those experienced for the**

**general population (background mortality) during the patients' modelled lifetime.**

a. **Please provide an overview of time points at which the hazards estimated with the extrapolated distributions are lower than those of background mortality for each comparison and each distribution.**

b. **Please justify the assumption that after these time points (at which the estimated hazards are lower than those of background mortality) patients with WM do not die anymore from their disease but only due to background mortality and provide clinical expert opinion to support this.**

c. **Please enable in the model a scenario that adds the hazards estimated from the survival analysis to those of background mortality (instead of using a max function) and present results as well as providing the updated model. Please comment on how the hazards change over time and the clinical plausibility.**

The time points at which the hazards estimated with the extrapolated distributions are lower than those of background mortality are summarised in the tables below by treatment and parametric distribution (see the 'Patient distribution' tab in the Excel model). These tables are based on those included in Section B.3.3.2 of the company submission, with the addition of a row for the time points.

Unfortunately, unlike some other lymphomas (e.g., diffuse large B-cell lymphoma) with available literature to inform the time points when the patient population would have the same mortality as the general population, the company did not identify similar studies specific to the WM population. Therefore, it is not possible to rely on this criterion for the model selection.

**Table 52. Summary of model selection – zanubrutinib**

| | Exponential | Weibull | Gompertz | Log-normal | Log-logistic | Gamma |
|---|---|---|---|---|---|---|
| **Hazard pattern** | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically decreasing | Monotonically decreasing | Monotonically decreasing | Monotonically decreasing | Monotonically decreasing |
| After adjusting for background mortality | Constant in the first 8 years; then increasing | Monotonically decreasing in the first 5 years; then increasing | Monotonically decreasing in the first 5 years; then increasing | Monotonically decreasing in the first 4 years; then increasing | Monotonically decreasing in the first 5 years; then increasing | Monotonically decreasing in the first 6 years; then increasing |
| Time points at which the hazards estimated with the extrapolated distributions are lower than those of background mortality | Approximately 8 years | Approximately 5-6 years | Approximately 5 years | Approximately 4 years | Approximately 5 years | Approximately 6 years |
| Applied in company submission | Base case | - | - | - | - | - |

**Table 53. Summary of model selection – ibrutinib**

| | Exponential | Weibull | Gompertz | Log-normal | Log-logistic | Gamma |
|---|---|---|---|---|---|---|
| **Hazard pattern** | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically decreasing | Monotonically decreasing | Monotonically decreasing | Monotonically decreasing | Monotonically decreasing |
| After adjusting for background mortality | Constant in the first 11 years; then increasing | Monotonically decreasing in the first 8 years; then increasing | Monotonically decreasing in the first 5 years; then increasing | Monotonically decreasing in the first 5 years; then increasing | Monotonically decreasing in the first 6 years; then increasing | Monotonically decreasing in the first 8 years; then increasing |
| Time points at which the hazards estimated with the extrapolated distributions are lower than those of background mortality | Approximately 11 years | Approximately 8 years | Approximately 5 years | Approximately 4-5 years | Approximately 6-7 years | Approximately 8-9 years |
| Applied in company submission | Base case | - | - | - | - | - |

**Table 54. Summary of model selection – zanubrutinib (match DRC)**

| | Exponential | Weibull | Gompertz | Log-normal | Log-logistic | Gamma |
|---|---|---|---|---|---|---|
| **Jointly fitted models** | | | | | | |
| **Hazard pattern** | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically increasing | Monotonically increasing | Increasing in the first 1 year; then decreasing | Increasing in the first 5 years; then decreasing | Monotonically increasing |
| After adjusting for background mortality | Constant in the first 7 years; then increasing | Monotonically increasing | Monotonically increasing | Increasing in the first 2 years; then decreasing for 8 years; then increasing | Increasing in the first 3 years; then stable for 10 years; then increasing | Monotonically increasing |
| Time points at which the hazards estimated with the extrapolated distributions are lower than those of background mortality | Approximately 7 years | Approximately 9 years | Approximately 10 years | Approximately 10 years | Approximately 10 years | Approximately 9-10 years |
| Applied in company submission | - | Scenario analysis | Scenario analysis | - | - | Base case |
| **Independently fitted models** | | | | | | |
| **Hazard pattern** | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically decreasing | Monotonically decreasing | Increasing in the first 3 months; then decreasing | Monotonically decreasing | Monotonically decreasing |
| After adjusting for background mortality | Constant in the first 8 years; then increasing | Stable in the first 7 years; then increasing | Decreasing in the first 2 years; then increasing | Increasing in the first 3 months; then decreasing for 4 years; then increasing | Stable in the first 6 years; then increasing | Stable in the first 8 years; then increasing |
| Time points at which the hazards estimated with the extrapolated | Approximately 7 years | Approximately 5-6 years | Approximately 2 years | Approximately 4 years | Approximately 5 years | Approximately 6-7 years |

| | Exponential | Weibull | Gompertz | Log-normal | Log-logistic | Gamma |
|---|---|---|---|---|---|---|
| distributions are lower than those of background mortality | | | | | | |
| Applied in company submission | - | - | - | - | - | - |

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide

**Table 55. Summary of model selection – DRC**

| | Exponential | Weibull | Gompertz | Log-normal | Log-logistic | Gamma |
|---|---|---|---|---|---|---|
| **Jointly fitted models** | | | | | | |
| **Hazard pattern** | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically increasing | Monotonically increasing | Increasing in the first 1 year; then decreasing | Increasing in the first 3 years; then decreasing | Monotonically increasing |
| After adjusting for background mortality | Constant in the first 15 years; then increasing | Monotonically increasing | Monotonically increasing | Increasing in the first 1 year; then decreasing for 11 years; then increasing | Increasing in the first 3 years; then decreasing for 10 years; then increasing | Monotonically increasing |
| Time points at which the hazards estimated with the extrapolated distributions are lower than those of background mortality | Approximately 15 years | Approximately 17 years | Approximately 21 years | Approximately 21 years | Approximately 13 years | Approximately 17 years |
| Applied in company submission | - | Scenario analysis | Scenario analysis | - | - | Base case |
| **Independently fitted models** | | | | | | |
| **Hazard pattern** | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically increasing | Monotonically increasing | Increasing in the first 1.5 years; then decreasing | Increasing in the first 3 years; then decreasing | Monotonically increasing |
| After adjusting for background mortality | Constant in the first 15 years; then increasing | Monotonically increasing | Monotonically increasing | Increasing in the first 1 year; then decreasing for 12 years; then increasing | Increasing in the first 3 years; then decreasing for 11 years; then increasing | Monotonically increasing |
| Time points at which the hazards estimated with the extrapolated | Approximately 15 years | Approximately 17 years | Approximately 22-23 years | Approximately 12 years | Approximately 13 years | Approximately 17 years |

| | Exponential | Weibull | Gompertz | Log-normal | Log-logistic | Gamma |
|---|---|---|---|---|---|---|
| distributions are lower than those of background mortality | | | | | | |
| Applied in company submission | - | - | - | - | - | - |

Abbreviations: DRC = dexamethasone, rituximab and cyclophosphamide

**Table 56. Summary of model selection – zanubrutinib (match BR)**

| | Exponential | Weibull | Gompertz | Log-normal | Log-logistic | Gamma |
|---|---|---|---|---|---|---|
| **Jointly fitted models** | | | | | | |
| **Hazard pattern** | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically increasing | Monotonically decreasing | Increasing in the first 2 years; then decreasing | Increasing in the first 5 years; then decreasing | Monotonically increasing |
| After adjusting for background mortality | Constant for 7 years; then increasing | Monotonically increasing | Decreasing in the first 5 years; the increasing | Increasing in the first 2 years; then decreasing/ stable for 5 years; then increasing | Increasing in the first 5 years; then stable for 5 years; then increasing | Monotonically increasing |
| Time points at which the hazards estimated with the extrapolated distributions are lower than those of background mortality | Approximately 7 years | Approximately 11 years | Approximately 5 years | Approximately 7 years | Approximately 9 years | Approximately 11-12 years |
| Applied in company submission | - | Scenario analysis | - | - | - | Scenario analysis |
| **Independently fitted models** | | | | | | |
| **Hazard pattern** | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically increasing | Monotonically increasing | Increasing in the first 4 years; then decreasing | Increasing in the first 6 years; then decreasing | Monotonically increasing |
| After adjusting for background mortality | Constant for 7 years; then increasing | Monotonically increasing | Monotonically increasing | Increasing in the first 3 years; then decreasing for 9 years; then increasing | Increasing in the first 6 years; then decreasing for 12 years; then increasing | Monotonically increasing |
| Time points at which the hazards estimated with the extrapolated | Approximately 7 years | Since the model baseline | Since the model baseline | Approximately 12 years | Approximately 18 years | Approximately 18-19 years |

| | Exponential | Weibull | Gompertz | Log-normal | Log-logistic | Gamma |
|---|---|---|---|---|---|---|
| distributions are lower than those of background mortality | | | | | | |
| Applied in company submission | Base case | - | - | - | - | - |

Abbreviations: BR = rituximab and bendamustine

**Table 57. Summary of model selection – BR**

| | Exponential | Weibull | Gompertz | Log-normal | Log-logistic | Gamma |
|---|---|---|---|---|---|---|
| **Jointly fitted models** | | | | | | |
| **Hazard pattern** | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically increasing | Monotonically decreasing | Increasing in the first 1 year; then decreasing | Increasing in the first 2 years; then decreasing | Monotonically increasing |
| After adjusting for background mortality | Constant for 17 years; then increasing | Monotonically increasing | Decreasing in the first 13 years; the increasing | Increasing in the first 1 years; then decreasing for 10 years; then increasing | Increasing in the first 2 years; then decreasing for 11 years; then increasing | Monotonically increasing |
| Time points at which the hazards estimated with the extrapolated distributions are lower than those of background mortality | Approximately 17 years | Approximately 21-22 years | Approximately 13 years | Approximately 11 years | Approximately 13 years | Approximately 21 years |
| Applied in company submission | - | Scenario analysis | - | - | - | Scenario analysis |
| **Independently fitted models** | | | | | | |
| **Hazard pattern** | | | | | | |
| Before adjusting for background mortality | Constant | Monotonically increasing | Monotonically decreasing | Increasing in the first 1 year; then decreasing | Increasing in the first 1 year; then decreasing | Monotonically increasing |
| After adjusting for background mortality | Constant for 17 years; then increasing | Monotonically increasing | Decreasing in the first 7 years; then increasing | Increasing in the first 1 year; then decreasing for 10 years; then increasing | Increasing in the first 1 year; then decreasing for 12 years; then increasing | Monotonically increasing |
| Time points at which the hazards estimated with the extrapolated distributions are | Approximately 17 years | Approximately 17-18 years | Approximately 6 years | Approximately 11 years | Approximately 12-13 years | Approximately 25-26 years |

| | Exponential | Weibull | Gompertz | Log-normal | Log-logistic | Gamma |
|---|---|---|---|---|---|---|
| lower than those of background mortality | | | | | | |
| Applied in company submission | - | Base case | - | - | - | - |

The Excel model has been updated to allow for a scenario (see 'Life Table' tab) which *adds* the hazards estimated from the survival analysis to those of background mortality (instead of using a *max* function). The results are provided below, which suggest that this could be a potential model driver, depending on specific treatment arm or other model parameters.

**Table 58. Scenario analysis: hazards estimated from the survival analysis added to those of background mortality**

| | ICER (pairwise comparison with ibrutinib) | ICER (pairwise comparison with BR) | ICER (pairwise comparison with DRC) |
|---|---|---|---|
| Max of hazards (CS base-case analysis) | [redacted] | [redacted] | [redacted] |
| Sum of hazards | [redacted] | [redacted] | [redacted] |

Abbreviations: BR = rituximab and bendamustine; CS = company submission; DRC = dexamethasone, rituximab and cyclophosphamide; ICER = incremental cost-effectiveness ratio
Notes: For transparency and simplicity, the results above were based on the base-case analysis in the CS, revised for this specific clarification question, without accounting for the revisions of model inputs/programming in response to other clarification questions. To generate the results according to different combinations of revisions, please refer to the updated Excel model

**B9. Priority Question: In the company submission base-case no treatment waning was assumed, i.e. the PFS and OS were assumed to be different for comparators and zanubrutinib for the whole duration of the time horizon.**

a. **Please justify the assumption of no treatment waning, i.e. that there is a lifetime difference in PFS and OS based on the initial treatment, also supporting this with expert opinion.**

b. **Please provide results for scenarios assuming treatment waning for the comparisons with BR and DRC.**

It is acknowledged that there are a lack of data (including but not limited to mature long-term data for BTK inhibitors in general, prior technology appraisals, etc.) to determine the best starting point for the treatment waning. As such, the model was developed such that treatment waning could be tested, but was not implemented in the base-case analysis due to substantial uncertainties.

Several exploratory analyses have been explored, results of which are shown in Table 59. These suggests that the incremental cost-effectiveness ratios are sensitive

to the assumption of potential treatment waning, depending on the time cut-off for treatment waning. The main challenge for this and prior appraisals (e.g., TA627 for FL/MZL) was that, despite treatment waning typically being a driver of the model results, there was a lack of evidence to suggest an appropriate time point. Therefore, several scenario analyses were conducted, including:

1. The most conservative (assuming no relative treatment benefit beyond the trial period),
2. The 5-year cut-off adopted in prior appraisals (despite the uncertainties of this time cut-off, as discussed in prior appraisals)
3. Relying on the extrapolated mean TTD (assuming that the relative treatment effect remains for as long as patients are on the treatment), to
4. Relying on the extrapolated mean PFS (assuming that the relative treatment effect remains for as long as patients remaining progression free alive).

**Table 59. Scenario analyses: treatment waning**

| Time cut-off for treatment waning | ICER (pairwise comparison with ibrutinib) | ICER (pairwise comparison with BR) | ICER (pairwise comparison with DRC) |
|---|---|---|---|
| No treatment waning (i.e., base case per company submission) | [redacted] | [redacted] | [redacted] |
| 30 months (approximating ASPEN trial follow-up) | [redacted] | [redacted] | [redacted] |
| 5 years (per other NICE TAs in lymphoma indications, such as TA627 for FL/MZL) | [redacted] | [redacted] | [redacted] |
| 7 years (approximating the extrapolated mean TTD of zanubrutinib) | [redacted] | [redacted] | [redacted] |
| 10 years (approximating the extrapolated mean PFS of zanubrutinib) | [redacted] | [redacted] | [redacted] |

Abbreviations: BR = rituximab and bendamustine; CS = company submission; DRC = dexamethasone, rituximab and cyclophosphamide; FL/MZL = follicular lymphoma/marginal zone lymphoma ICER = incremental cost-effectiveness ratio; NICE = National Institute of Health and Care Excellence; PFS = progression-free survival; TA = technology appraisal; TTD = time to discontinuation

Notes: For transparency and simplicity, the results above were based on the base-case analysis in the CS, revised for this specific clarification question, without accounting for the revisions of model inputs/programming in response to other clarification questions. To generate the results according to different combinations of revisions, please refer to the updated Excel model

## *Adverse events*

**B10.** Section B 3.3.3 describes the incidence and duration of AEs for the intervention and comparators. Only AEs with a severity grade of 3 or larger and an occurrence in 5% or more of the patient population are included.

a. Please provide a justification for not including mild AEs with grade <3. Please also provide justification for excluding AEs with an incidence < 5%.
b. Please conduct scenario analyses which include all AEs with an incidence ≥ 1% in the population and provide an equivalent to table B.3.19. describing the incidence and duration of all AEs with an incidence of ≥ 1%

The inclusion criteria of “Grade ≥3” and “incidence of ≥5%” for AEs has been widely adopted in oncology models (including NICE TA491 of ibrutinib for WM), as these AEs are more likely to have an impact on costs and health-related quality of life (HRQoL), whereas mild AEs or extremely rare AEs are considered to have minimal impact on the model results.

Table 60 is an updated version of Table B.3.19 in the company submission with the inclusion criteria of Grade ≥3 and incidence of ≥*1%*. The newly added AEs (i.e., those occurring in 1–5% of patients) are indicated in italics. Of note, during the clarification stage, the company identified one AE, hypotension, that should have been included in the company submission (because it occurred in >5% of patients in the DRC arm) but was omitted in error. A comment box in the ‘AE’ tab of the updated model has been added to reflect this.

**Table 60. Incidence and duration of Grade ≥3 AEs occurring in ≥1% of patients in any treatment arm**

| | AE incidence, % | | | | | | Duration, days |
|---|---|---|---|---|---|---|---|
| | Zanubrutinib (N=101) | Ibrutinib (N=98) | Zanubrutinib adjusted to match BR ($n_{eff}$= [redacted]) | BR (N=71) | Zanubrutinib adjusted to match DRC ($n_{eff}$= [redacted]) | DRC (N=72) | ASPEN Safety Analysis Set (N=199) |
| Reference | ASPEN IPD | ASPEN IPD | ASPEN IPD (match BR) | Tedeschi et al. 2015[12] | ASPEN IPD, (match DRC) | Dimopoulos et al. 2007[10] | ASPEN IPD |
| Anaemia | 4.95 | 5.10 | [redacted] | NR | [redacted] | NR | 17.0 |
| Hypertension | 5.94 | 11.22 | [redacted] | NR | [redacted] | NR | 20.9 |
| Neutropenia | 15.84 | 8.16 | [redacted] | 35.21 | [redacted] | 10.00 | 10.9 |
| Pneumonia | 0.99 | 7.14 | [redacted] | 5.63 | [redacted] | NR | 21.3 |
| Thrombocytopenia | 5.94 | 3.06 | [redacted] | NR | [redacted] | 0.00 | 28.8 |
| *Nausea* | *0.00* | *1.02* | [redacted] | *NR* | [redacted] | *0.00* | *5.0* |
| *Vomiting* | *0.00* | *1.02* | [redacted] | *NR* | [redacted] | *0.00* | *5.0* |
| *Headache* | *0.99* | *1.02* | [redacted] | *NR* | [redacted] | *2.78* | *6.7* |
| *Hypotension* | *0.00* | *0.00* | [redacted] | *NR* | [redacted] | *5.56* | *0.0* |
| *Sepsis* | *1.98* | *3.06* | [redacted] | *1.41* | [redacted] | *NR* | *5.0* |

Abbreviations: BR = bendamustine and rituximab; DRC = dexamethasone, rituximab and cyclophosphamide; IPD = individual patient-level data; N = number of patients evaluable; $n_{eff}$ = effective sample size; NR = not reported

To enable a scenario analysis using the AE incidences from Table 60, the company also updated Table B.3.23. AE disutilties (see Table 61). AE costs were not changed.

**Table 61. AE disutilties**

| AE | Disutility | Source |
|---|---|---|
| Anaemia | 0.088 | NICE TA491[3] |
| Hypertension | 0.195 | Assumed to be the same as that for pneumonia, in line with the assumption adopted in NICE TA429 for ibrutinib in CLL[65] |
| Neutropenia | 0.185 | NICE TA491[3] |
| Pneumonia | 0.195 | NICE TA491[3] |
| Thrombocytopenia | 0.123 | NICE TA491[3] |
| *Nausea* | *0.195* | *Assumption, based on the disutilities above for other AEs* |
| *Vomiting* | *0.195* | |
| *Headache* | *0.195* | |
| *Hypotension* | *0.195* | |
| *Sepsis* | *0.195* | |

Abbreviations: AE = adverse event; CLL = chronic lymphocytic leukaemia; NICE = National Institute for Health and Care Excellence; TA = technology appraisal

Based on the inputs above, results of a scenario analysis results relying on Grade ≥3 AEs occurring in ≥1% of patients in any treatment arms are summarised Table 62.

This analysis demonstrates that the inclusion/exclusion of AEs occurring in 1–5% of patients in any treatment arms had extremely minimal impact on the ICERs.

**Table 62. Scenario analysis: inclusion of Grade ≥3 AEs occurring in ≥5% and ≥1% of patients in any treatment arms**

| | ICER (pairwise comparison with ibrutinib) | ICER (pairwise comparison with BR) | ICER (pairwise comparison with DRC) |
|---|---|---|---|
| Grade ≥3 AEs in ≥5% of patients (CS base-case analysis) | [redacted] | [redacted] | [redacted] |
| Grade ≥3 AEs in ≥1% of patients (i.e., scenario analysis) | [redacted] | [redacted] | [redacted] |

Abbreviations: AE = adverse event; BR = rituximab and bendamustine; DRC = dexamethasone, rituximab and cyclophosphamide; ICER = incremental cost-effectiveness ratio
Notes: For transparency and simplicity, the results above were based on the base-case analysis in the CS, revised for this specific clarification question, without accounting for the revisions of model inputs/programming in response to other clarification questions. To generate the results according to different combinations of revisions, please refer to the updated Excel model

## *Health-related quality of life*

**B11. Priority Question: Health state utility values are, according to Figure B.3.47, B.3.50, and B.3.53 key drivers of the cost-effectiveness results.**

a. **Table B.3.21 is part of the explanation how the progression-free utility was calculated but the values of the regression coefficients are not provided. Please provide a table in which the coefficients of the regression models are added.**

b. **Please explain, with appropriate justifications, how the utility values reported in Table B3.22 were estimated.**

c. **Section B.3.4.1 explains the EQ-5D-5L assessment schedule. Please provide, per measurement timepoint, separately for zanubrutinib and ibrutinib:**

   I. **the total number of EQ-5D-5L responses**

   II. **estimated mean utility values and standard error**

   III. **a breakdown how many patients were on and off treatment**

IV. **the extent of missing data observed**

d. **Please explain, with appropriate justifications, how missing data were handled and the implications of this approach.**

e. **Please compare patient characteristics of patients which were included and patients excluded from utility value calculations for both treatment groups separately and for the whole trial population combined (independent of treatment groups).**

f. **Please clarify what the likely causes of missing data were and what the potential impact of these missing data on the estimation of the utility scores would be, separately for patients who had completely and partially missing utility data.**

g. **Please recalculate the utility estimates reported in Table B.3.22 while imputing missing values (for the patients with completely missing utility data and patients with partially missing utility data) using multiple imputation (incorporating potential explanatory variables and using at least 10 imputations).**

   I. **Please provide in detail, the methods used to impute and pool the utility data**

   II. **Please elaborate on the plausibility of the imputed utility values**

   III. **Please provide an updated economic model as well as scenario analysis incorporating these newly calculated utility values**

h. **Please provide the table requested above (Table B.3.22 while imputing missing values) stratified for patients being on treatment or the comparator.**

i. **Please rerun the analyses performed to obtain the utility values presented in B 3.22 (i.e. original approach from the company submission) stratified for patients being on treatment (i.e. receiving Zanubrutinib) or a comparator.**

j. **Please provide an updated economic model as well as scenario analysis incorporating the estimated utility values in response to sub-questions g and h (i.e. utility values estimated stratified for patients being on treatment or not with and without imputation).**

The detailed coefficients of the regression models from Table B.3.21 are provided below.

**Table 63. Model 1**

| | Coefficient | SE | Df | t statistics | p-value |
|---|---|---|---|---|---|
| Intercept | 1.0037 | 0.09541 | 186 | 10.52 | <.0001*** |
| Tx_zanu | 0.001645 | 0.02446 | 188 | 0.07 | 0.9465 |
| Age | -0.00354 | 0.001294 | 187 | -2.73 | 0.0068*** |
| Gender (male) | 0.04566 | 0.02611 | 188 | 1.75 | 0.0819* |
| | | | | | |
| Number of observations | | 900 | Number of patients | | 192 |
| AIC | | -895.9 | BIC | | -886.2 |

Abbreviations: AIC = Akaike information criteria; BIC =Bayesian information criteria; SE = standard error
Significance level: * 10%; ** 5%; *** 1%

**Table 64. Model 2**

| | Coefficient | SE | Df | t statistics | p-value |
|---|---|---|---|---|---|
| Intercept | 0.9689 | 0.09577 | 190 | 10.12 | <.0001*** |
| Tx_zanu | 0.002176 | 0.02444 | 189 | 0.09 | 0.9292 |
| Age | -0.00339 | 0.001294 | 188 | -2.62 | 0.0095*** |
| Gender (male) | 0.04550 | 0.02609 | 189 | 1.74 | 0.0827* |
| Days from treatment initiation (dayt) – numerical | 0.000098 | 0.000025 | 306 | 3.99 | <.0001*** |
| | | | | | |
| Number of observations | | 900 | Number of patients | | 192 |
| AIC | | -892.1 | BIC | | -882.3 |

Abbreviations: AIC = Akaike information criteria; BIC =Bayesian information criteria; SE = standard error
Significance level: * 10%; ** 5%; *** 1%

**Table 65. Model 3**

| | Coefficient | SE | Df | t statistics | p-value |
|---|---|---|---|---|---|
| Intercept | 0.9694 | 0.09577 | 191 | 10.12 | <.0001*** |
| Tx_zanu | 0.002209 | 0.02444 | 189 | 0.09 | 0.9281 |
| Age | -0.00340 | 0.001294 | 188 | -2.63 | 0.0093*** |
| Gender (male) | 0.04541 | 0.02608 | 189 | 1.74 | 0.0833* |
| Days from treatment initiation ($day_t$) – numerical | 0.002718 | 0.000709 | 301 | 3.83 | 0.0002*** |
| | | | | | |
| Number of observations | | 900 | Number of patients | | 192 |
| AIC | | -897.6 | BIC | | -887.9 |

Abbreviations: AIC = Akaike information criteria; BIC =Bayesian information criteria; SE = standard error
Significance level: * 10%; ** 5%; *** 1%

For each model, details of how the utility values reported in Table B3.22 were estimated are provided below.

For model 1, least square (LS) means were estimated for each treatment group at mean values of the covariates (age=69.07 years, males=0.6756) (see SAS output in Table 66). For example, for zanubrutinib, the LS mean of the utility value was estimated by 1.0037+0.001645-0.00354*69.07+0.04566*0.6756.

**Table 66. SAS output, Model 1**

Least Squares Means

| Effect | trtp | age | sex_n | Estimate | Standard Error | DF | t Value | Pr > \|t\| |
|---|---|---|---|---|---|---|---|---|
| trtp | 1 | 69.07 | 0.68 | 0.7917 | 0.01695 | 187 | 46.71 | <.0001 |
| trtp | 0 | 69.07 | 0.68 | 0.7901 | 0.01762 | 190 | 44.84 | <.0001 |

For model 2, LS means were first estimated for each treatment group at each measurement timepoint adjusted for the mean values of the covariates ((age=69.07 years, males=0.6756) (see SAS output in Table 67). For example, for zanubrutinib, the LS mean of the utility value at cycle 4 day 1 (day = 84 in the model specification) is estimated by 0.9689+0.002176-0.00339*69.07+0.04550*0.6756+0.000098*84. After that, for each treatment, the LS mean was derived as the weighted average of the LS means at measurement timepoints using the number of observations at each timepoint divided by the total number of observations within each treatment as the weight for each timepoint.

**Table 67. SAS output, Model 2**

| Least Squares Means | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Effect | trtp | age | sex_n | ADY | Estimate | Standard Error | DF | t Value | Pr > \|t\| |
| trtp | 1 | 69.07 | 0.68 | -1.00 | 0.7674 | 0.01801 | 237 | 42.61 | <.0001 |
| trtp | 0 | 69.07 | 0.68 | -1.00 | 0.7652 | 0.01867 | 237 | 40.99 | <.0001 |
| trtp | 1 | 69.07 | 0.68 | 84.00 | 0.7758 | 0.01741 | 209 | 44.55 | <.0001 |
| trtp | 0 | 69.07 | 0.68 | 84.00 | 0.7736 | 0.01808 | 211 | 42.79 | <.0001 |
| trtp | 1 | 69.07 | 0.68 | 168.00 | 0.7840 | 0.01705 | 193 | 45.97 | <.0001 |
| trtp | 0 | 69.07 | 0.68 | 168.00 | 0.7818 | 0.01772 | 196 | 44.11 | <.0001 |
| trtp | 1 | 69.07 | 0.68 | 252.00 | 0.7922 | 0.01694 | 188 | 46.77 | <.0001 |
| trtp | 0 | 69.07 | 0.68 | 252.00 | 0.7900 | 0.01761 | 190 | 44.88 | <.0001 |
| trtp | 1 | 69.07 | 0.68 | 336.00 | 0.8005 | 0.01708 | 193 | 46.88 | <.0001 |
| trtp | 0 | 69.07 | 0.68 | 336.00 | 0.7983 | 0.01773 | 195 | 45.03 | <.0001 |
| trtp | 1 | 69.07 | 0.68 | 504.00 | 0.8169 | 0.01807 | 236 | 45.21 | <.0001 |
| trtp | 0 | 69.07 | 0.68 | 504.00 | 0.8148 | 0.01867 | 235 | 43.65 | <.0001 |
| trtp | 1 | 69.07 | 0.68 | 672.00 | 0.8334 | 0.01989 | 320 | 41.91 | <.0001 |
| trtp | 0 | 69.07 | 0.68 | 672.00 | 0.8312 | 0.02041 | 314 | 40.72 | <.0001 |
| trtp | 1 | 69.07 | 0.68 | 840.00 | 0.8499 | 0.02233 | 415 | 38.06 | <.0001 |
| trtp | 0 | 69.07 | 0.68 | 840.00 | 0.8477 | 0.02278 | 406 | 37.21 | <.0001 |

For model 3, a similar approach as that for model 2 was used, but the time variable is completed cycles of treatment at the measurement timepoint. For example, for zanubrutinib, the LS mean of the utility value at cycle 4 day 1 (visit=3 in the model specification) is estimated by 0.9694+0.002209-0.00340*69.07+0.04541*0.6756+0.002718*3. The SAS output for the LS mean at each timepoint is shown in Table 68. After that, for each treatment, the LS mean was derived by weighted average of the LS means at measurement timepoints using the number of observations at each timepoint divided by the total number of observations within each treatment as the weight for each timepoint.

**Table 68. SAS output, Model 4**

| Least Squares Means | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Effect | trtp | age | sex_n | MCYCLE | Estimate | Standard Error | DF | t Value | Pr > \|t\| |
| trtp | 1 | 69.07 | 0.68 | 0.00 | 0.7675 | 0.01808 | 240 | 42.45 | <.0001 |
| trtp | 0 | 69.07 | 0.68 | 0.00 | 0.7653 | 0.01874 | 241 | 40.84 | <.0001 |
| trtp | 1 | 69.07 | 0.68 | 3.00 | 0.7757 | 0.01745 | 211 | 44.45 | <.0001 |
| trtp | 0 | 69.07 | 0.68 | 3.00 | 0.7735 | 0.01812 | 213 | 42.68 | <.0001 |
| trtp | 1 | 69.07 | 0.68 | 6.00 | 0.7838 | 0.01706 | 194 | 45.94 | <.0001 |
| trtp | 0 | 69.07 | 0.68 | 6.00 | 0.7816 | 0.01774 | 197 | 44.07 | <.0001 |
| trtp | 1 | 69.07 | 0.68 | 9.00 | 0.7920 | 0.01693 | 188 | 46.77 | <.0001 |
| trtp | 0 | 69.07 | 0.68 | 9.00 | 0.7898 | 0.01760 | 190 | 44.87 | <.0001 |
| trtp | 1 | 69.07 | 0.68 | 12.00 | 0.8001 | 0.01707 | 193 | 46.87 | <.0001 |
| trtp | 0 | 69.07 | 0.68 | 12.00 | 0.7979 | 0.01772 | 195 | 45.03 | <.0001 |
| trtp | 1 | 69.07 | 0.68 | 18.00 | 0.8165 | 0.01811 | 239 | 45.08 | <.0001 |
| trtp | 0 | 69.07 | 0.68 | 18.00 | 0.8142 | 0.01870 | 237 | 43.54 | <.0001 |
| trtp | 1 | 69.07 | 0.68 | 24.00 | 0.8328 | 0.02002 | 326 | 41.60 | <.0001 |
| trtp | 0 | 69.07 | 0.68 | 24.00 | 0.8305 | 0.02053 | 319 | 40.45 | <.0001 |
| trtp | 1 | 69.07 | 0.68 | 30.00 | 0.8491 | 0.02258 | 423 | 37.61 | <.0001 |
| trtp | 0 | 69.07 | 0.68 | 30.00 | 0.8469 | 0.02301 | 413 | 36.80 | <.0001 |

For each model, the weighted LS mean across treatment was derived as the weighted average of the LS means of the two treatments derived above with the proportion of patients in each treatment arm included in the modelling as the weight for each treatment arm (99/192 for zanubrutinib, 93/192 for ibrutinib). For example, for model 1, the weighted LS mean across treatment= 99/192*0.7917+93/192*0.7901.

In response to question B11c, further details are provided in Table 69.

**Table 69. ASPEN EQ-5D-5L results by Cycle/Day**

| | **Zanubrutinib (N = 102)** | **Ibrutinib (N = 99)** |
|---|---|---|
| Screening | | |
| n | 61 | 62* |
| Mean (SD) | 0.7390 (0.17915) | 0.7268 (0.21193) |
| Median | 0.7350 | 0.7360 |
| Q1, Q3 | 0.6470, 0.8790 | 0.6350, 0.8790 |
| Min, Max | 0.270, 1.000 | 0.064, 1.000 |
| Cycle 4 Day 1 | | |
| n | 73 | 63 |
| Mean (SD) | 0.8165 (0.19106) | 0.7820 (0.22808) |
| Median | 0.8370 | 0.8370 |
| Q1, Q3 | 0.7250, 1.0000 | 0.6790, 1.0000 |
| Min, Max | 0.294, 1.000 | -0.032, 1.000 |
| Cycle 7 Day 1 | | |
| n | 81 | 72 |
| Mean (SD) | 0.8069 (0.20906) | 0.7788 (0.20681) |

| | Zanubrutinib (N = 102) | Ibrutinib (N = 99) |
|---|---|---|
| Median | 0.8370 | 0.7680 |
| Q1, Q3 | 0.7080, 1.0000 | 0.6535, 1.0000 |
| Min, Max | 0.073, 1.000 | 0.169, 1.000 |
| Cycle 10 Day 1 | | |
| n | 82 | 74 |
| Mean (SD) | 0.8074 (0.20238) | 0.8083 (0.23029) |
| Median | 0.8370 | 0.8425 |
| Q1, Q3 | 0.6910, 1.0000 | 0.7080, 1.0000 |
| Min, Max | -0.173, 1.000 | -0.202, 1.000 |
| Cycle 13 Day 1 | | |
| n | 84 | 79 |
| Mean (SD) | 0.8255 (0.17160) | 0.8151 (0.18320) |
| Median | 0.8370 | 0.8370 |
| Q1, Q3 | 0.7315, 1.0000 | 0.7270, 1.0000 |
| Min, Max | 0.254, 1.000 | 0.231, 1.000 |
| Cycle 19 Day 1 | | |
| n | 63 | 60 |
| Mean (SD) | 0.7937 (0.17233) | 0.8149 (0.21722) |
| Median | 0.7950 | 0.8625 |
| Q1, Q3 | 0.6660, 1.0000 | 0.7155, 1.0000 |
| Min, Max | 0.306, 1.000 | -0.098, 1.000 |
| Cycle 25 Day 1 | | |
| n | 22 | 17 |
| Mean (SD) | 0.7908 (0.17262) | 0.8119 (0.26089) |
| Median | 0.8370 | 0.9060 |
| Q1, Q3 | 0.6790, 0.9060 | 0.7400, 1.0000 |
| Min, Max | 0.304, 1.000 | 0.057, 1.000 |
| Cycle 31 Day 1 | | |
| n | 3 | 4 |
| Mean (SD) | 0.8513 (0.13543) | 0.9290 (0.08376) |
| Median | 0.8190 | 0.9395 |
| Q1, Q3 | 0.7350, 1.0000 | 0.8580, 1.0000 |
| Min, Max | 0.735, 1.000 | 0.837, 1.000 |
| End of Treatment | | |
| n | 4 | 7 |
| Mean (SD) | 0.4630 (0.32516) | 0.7170 (0.18429) |
| Median | 0.4025 | 0.7110 |
| Q1, Q3 | 0.2365, 0.6895 | 0.5550, 0.8770 |
| Min, Max | 0.141, 0.906 | 0.451, 1.000 |

Abbreviations: Q = quartile SD = standard deviation

*****One patient was not treated but had complete baseline measurement at screening. If the patient is excluded from summary, the mean (SE) would be 0.7335 (0.20686)

Of note, EQ-5D-5L data were collected up to the EOT visit. In Table 69, n at each measurement timepoint (except for EOT) were for patients on treatment, with no patients off treatment (except for the one not treated in the ibrutinib group at screening). At EOT, n is for patients off treatment.

The extent of missing data at each measurement timepoint is shown in Table 70. In ASPEN Cohort 1 (ITT set; 102+99=201 patients), 8 patients did not have any

response assessment or at least one complete EQ-5D-5L measurement and were excluded from the utility analysis set. In addition, one patient who had a complete measurement at baseline but was not treated was excluded from the utility analysis set. Thus, 192 patients were included in the utility analysis set. The missing rate at each timepoint (excluding EOT) ranged from 0 to 38.4%.

**Table 70. EQ-5D-5L compliance rates**

| | **Zanubrutinib (N = 99)** | **Ibrutinib (N = 93)** |
|---|---|---|
| Screening | | |
| Completion | 61 (59.8) | 61 (61.6) |
| Compliance | 61/ 99 (61.6) | 61/ 93 (65.6) |
| Cycle 4 | | |
| Completion | 73 (71.6) | 63 (63.6) |
| Compliance | 73/ 96 (76.0) | 63/ 87 (72.4) |
| Cycle 7 | | |
| Completion | 81 (79.4) | 72 (72.7) |
| Compliance | 81/ 94 (86.2) | 72/ 86 (83.7) |
| Cycle 10 | | |
| Completion | 82 (80.4) | 74 (74.7) |
| Compliance | 82/ 91 (90.1) | 74/ 85 (87.1) |
| Cycle 13 | | |
| Completion | 84 (82.4) | 79 (79.8) |
| Compliance | 84/ 88 (95.5) | 79/ 84 (94.0) |
| Cycle 19 | | |
| Completion | 63 (61.8) | 60 (60.6) |
| Compliance | 63/ 66 (95.5) | 60/ 62 (96.8) |
| Cycle 25 | | |
| Completion | 22 (21.6) | 17 (17.2) |
| Compliance | 22/ 24 (91.7) | 17/ 20 (85.0) |
| Cycle 31 | | |
| Completion | 3 (2.9) | 4 (4.0) |
| Compliance | 3/ 4 (75.0) | 4/ 4 (100.0) |
| End Of Treatment | | |
| Completion | 4 (3.9) | 7 (7.1) |
| Compliance | 4/ 21 (19.0) | 7/ 22 (31.8) |

Abbreviations: EQ-5D-5L = EuroQol 5-Dimensions 5-Level

For the handling of missing data, no missing data imputation was applied during the analysis of utility data. The MMRM was used to handle the missing data with assumption of missing at random (MAR).

There are 9 patients excluded from the utility analysis. The characteristics of the included and excluded patients are summarised by treatment groups and pooled groups in Table 71.

**Table 71. Characteristics of included (column "Utility Analysis") and excluded (column "Non-Utility Analysis") patients**

| | Ibrutinib | | Zanubrutinib | | Total | |
|---|---|---|---|---|---|---|
| | Utility Analysis (N = 93) | Non-Utility Analysis (N = 6) | Utility Analysis (N = 99) | Non-Utility Analysis (N = 3) | Utility Analysis (N = 192) | Non-Utility Analysis (N = 9) |
| Age (years) | | | | | | |
| n | 93 | 6 | 99 | 3 | 192 | 9 |
| Mean (SD) | 70.0 (8.66) | 68.8 (8.04) | 68.9 (10.24) | 77.3 (8.74) | 69.4 (9.50) | 71.7 (8.80) |
| Median | 71.0 | 69.5 | 70.0 | 75.0 | 70.0 | 70.0 |
| Q1, Q3 | 65.0, 74.0 | 66.0, 75.0 | 62.0, 77.0 | 70.0, 87.0 | 63.0, 76.5 | 69.0, 75.0 |
| Min, Max | 38, 90 | 55, 78 | 45, 87 | 70, 87 | 38, 90 | 55, 87 |
| Age Group, n (%) | | | | | | |
| ≤ 75 years | 72 (77.4) | 5 (83.3) | 66 (66.7) | 2 (66.7) | 138 (71.9) | 7 (77.8) |
| > 75 years | 21 (22.6) | 1 (16.7) | 33 (33.3) | 1 (33.3) | 54 (28.1) | 2 (22.2) |
| Gender, n (%) | | | | | | |
| Male | 62 (66.7) | 3 (50.0) | 67 (67.7) | 2 (66.7) | 129 (67.2) | 5 (55.6) |
| Female | 31 (33.3) | 3 (50.0) | 32 (32.3) | 1 (33.3) | 63 (32.8) | 4 (44.4) |
| Race, n (%) | | | | | | |
| Asian | 0 (0.0) | 0 (0.0) | 4 (4.0) | 0 (0.0) | 4 (2.1) | 0 (0.0) |
| White | 90 (96.8) | 5 (83.3) | 86 (86.9) | 2 (66.7) | 176 (91.7) | 7 (77.8) |
| Black or African American | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) |
| American Indian or Alaska Native | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) |
| Aboriginal/Torres Strait Islander | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) |
| Native Hawaiian or Other Pacific Islander | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) |
| Other | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) |
| Not Reported/Unknown | 3 (3.2) | 1 (16.7) | 9 (9.1) | 1 (33.3) | 12 (6.3) | 2 (22.2) |
| Ethnicity, n (%) | | | | | | |
| Not Hispanic or Latino | 86 (92.5) | 5 (83.3) | 80 (80.8) | 2 (66.7) | 166 (86.5) | 7 (77.8) |
| Hispanic or Latino | 4 (4.3) | 1 (16.7) | 4 (4.0) | 0 (0.0) | 8 (4.2) | 1 (11.1) |
| Not Reported/Unknown | 3 (3.2) | 0 (0.0) | 15 (15.2) | 1 (33.3) | 18 (9.4) | 1 (11.1) |
| Geographic Region, n (%) | | | | | | |
| Asia | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) |
| Australia/New Zealand | 29 (31.2) | 1 (16.7) | 31 (31.3) | 1 (33.3) | 60 (31.3) | 2 (22.2) |
| Europe | 55 (59.1) | 4 (66.7) | 60 (60.6) | 1 (33.3) | 115 (59.9) | 5 (55.6) |
| North America | 9 (9.7) | 1 (16.7) | 8 (8.1) | 1 (33.3) | 17 (8.9) | 2 (22.2) |

| | Ibrutinib | | Zanubrutinib | | Total | |
|---|---|---|---|---|---|---|
| | **Utility Analysis (N = 93)** | **Non-Utility Analysis (N = 6)** | **Utility Analysis (N = 99)** | **Non-Utility Analysis (N = 3)** | **Utility Analysis (N = 192)** | **Non-Utility Analysis (N = 9)** |
| ECOG Performance Status, n (%) | | | | | | |
| 0 | 39 (41.9) | 3 (50.0) | 46 (46.5) | 0 (0.0) | 85 (44.3) | 3 (33.3) |
| 1 | 48 (51.6) | 2 (33.3) | 48 (48.5) | 2 (66.7) | 96 (50.0) | 4 (44.4) |
| 2 | 6 (6.5) | 1 (16.7) | 5 (5.1) | 1 (33.3) | 11 (5.7) | 2 (22.2) |
| Number of Prior Lines of Therapy, n (%) | | | | | | |
| 0 | 17 (18.3) | 1 (16.7) | 19 (19.2) | 0 (0.0) | 36 (18.8) | 1 (11.1) |
| 1-3 | 69 (74.2) | 5 (83.3) | 74 (74.7) | 2 (66.7) | 143 (74.5) | 7 (77.8) |
| >3 | 7 (7.5) | 0 (0.0) | 6 (6.1) | 1 (33.3) | 13 (6.8) | 1 (11.1) |
| Baseline IgM (Central Lab), n (%) | | | | | | |
| ≥ 40 g/L | 36 (38.7) | 2 (33.3) | 35 (35.4) | 1 (33.3) | 71 (37.0) | 3 (33.3) |
| < 40 g/L | 56 (60.2) | 4 (66.7) | 64 (64.6) | 2 (66.7) | 120 (62.5) | 6 (66.7) |
| Baseline β2-Microglobulin (Central Lab), n(%) | | | | | | |
| > 3 mg/L | 69 (74.2) | 5 (83.3) | 72 (72.7) | 3 (100.0) | 141 (73.4) | 8 (88.9) |
| ≤ 3 mg/L | 24 (25.8) | 1 (16.7) | 27 (27.3) | 0 (0.0) | 51 (26.6) | 1 (11.1) |
| Extramedullary Disease per IRC [b], n (%) | | | | | | |
| Yes | 68 (73.1) | 5 (83.3) | 79 (79.8) | 2 (66.7) | 147 (76.6) | 7 (77.8) |
| Lymphadenopathy | 63 (67.7) | 4 (66.7) | 77 (77.8) | 2 (66.7) | 140 (72.9) | 6 (66.7) |
| Splenomegaly | 12 (12.9) | 1 (16.7) | 17 (17.2) | 0 (0.0) | 29 (15.1) | 1 (11.1) |
| Other | 3 (3.2) | 0 (0.0) | 1 (1.0) | 0 (0.0) | 4 (2.1) | 0 (0.0) |
| No | 25 (26.9) | 1 (16.7) | 20 (20.2) | 1 (33.3) | 45 (23.4) | 2 (22.2) |
| WM IPSS per SPEP (Derived) [c], n (%) | | | | | | |
| Low | 13 (14.0) | 0 (0.0) | 17 (17.2) | 0 (0.0) | 30 (15.6) | 0 (0.0) |
| Intermediate | 40 (43.0) | 2 (33.3) | 37 (37.4) | 1 (33.3) | 77 (40.1) | 3 (33.3) |
| High | 40 (43.0) | 4 (66.7) | 45 (45.5) | 2 (66.7) | 85 (44.3) | 6 (66.7) |
| Baseline Hemoglobin, n (%) | | | | | | |
| ≤ 110 g/L | 50 (53.8) | 3 (50.0) | 65 (65.7) | 2 (66.7) | 115 (59.9) | 5 (55.6) |
| > 110 g/L | 43 (46.2) | 3 (50.0) | 34 (34.3) | 1 (33.3) | 77 (40.1) | 4 (44.4) |
| Baseline Platelet, (%) | | | | | | |
| ≤ 100 x $10^9$/L | 9 (9.7) | 3 (50.0) | 12 (12.1) | 0 (0.0) | 21 (10.9) | 3 (33.3) |
| > 100 x $10^9$/L | 84 (90.3) | 3 (50.0) | 87 (87.9) | 3 (100.0) | 171 (89.1) | 6 (66.7) |

| | Ibrutinib | | Zanubrutinib | | Total | |
|---|---|---|---|---|---|---|
| | Utility Analysis (N = 93) | Non-Utility Analysis (N = 6) | Utility Analysis (N = 99) | Non-Utility Analysis (N = 3) | Utility Analysis (N = 192) | Non-Utility Analysis (N = 9) |

Abbreviations: ECOG = Eastern Cooperative Oncology Group; IPSS = International Prognostic Scoring System; IRC = independent review committee; Q = quartile SD = standard deviation; SPEP = serum protein electrophoresis; WM = Waldenström's macroglobulinaemia

As discussed above, one patient in ITT set was not treated and was excluded from analysis although complete baseline measurement was collected. Another patient among the 8 excluded patients was not treated and no complete measurement was collected. The remaining 7 excluded patients did not have at least one complete measurement, which only accounts for 3.5% of 199 treated patients. The completely missing data of these 7 patients would not affect the estimate of the utility score.

The reasons for EQ5D-5L data not being collected at scheduled measurement timepoints were not recorded. However, the possible causes of missing data at scheduled measurement timepoints include non-attendance at scheduled visits, or non-completion of all the measurements on EQ5D-5L scale. Although some data were collected at unscheduled visits, they were not included in the analysis as the number of observations at scheduled visits was considered sufficient. If one of the five dimensions of the EQ5D-5L scale was not completed, the utility value would not be derived based on the UK value set, thus the utility value would be missing. Considering the extent of missing data (see Table 70), the partially missing data was not expected to affect the estimate of the utility score.

**B12. Priority Question: According to a recent publication on the UK Outcomes Framework (Office for National Statistics, 2019)(https://www.ethnicity-facts-figures.service.gov.uk/health/physical-health/health-related-quality-of-life-for-people-aged-65-and-over/latest#data-sources), health-related quality of life in people aged 65 and over was 0.735 for the general population.**

a. **Please justify why the progression-free health-related quality of life used in the model was higher than the average of the general population.**

b. **It is unclear whether the health-related quality of life is assumed to remain stable over the treatment period. Please clarify whether this is the case.**

c. **If assumed to remain stable, given the health-related quality of life could be assumed to decrease significantly over the treatment period due to age, please justify the assumption that it remains stable.**

d. **Please provide an updated model adjusting health state utilities for population norms and provide an updated model file. The adjustment could be done for example with the method described in this article: Ara and Brazier (2010) Populating an Economic Model with Health State Utility Values: Moving Towards Better Practice. Value in Health Vol 13, number 5, 509-518.**

There are several potential reasons for the progression-free HRQoL estimate based on ASPEN data being higher than the average of the general population, including but not limited to (1) natural differences in the clinical trial settings where patients were closely monitored compared with real-world settings, (2) differences in the geographic locations between ASPEN (Europe [60%) and Australia or New Zealand [31%]) and the UK.

HRQoL was assumed to remain stable over the treatment period, which was based on the observed utility estimates from ASPEN (shown in response to clarification question B11c, with a relatively stable trend over time throughout cycles 4 to 25). The post cycle 25 estimates appeared to be unstable, primarily driven by the low number of observations. However, it is acknowledged that the unadjusted utility estimates within the trial period were not sufficient to justify that the HRQoL would remain stable over a lifetime horizon, and as shown in the response to clarification question B11b, age was a potential predictor of utilities. Given the above and due to the relatively immature nature of ASPEN data, the model has been updated to accommodate age-related utility decrease relying on the equation from Ara and Brazier (2010) (see 'Life Table' tab of the Excel model). Specifically, the utility of the general population was estimated by age and then used to derive utility multipliers over time.

*General population utility value = 0.9508566 + 0.0212126*male – 0.0002587*age in years – 0.0000332*age$^2$*

As shown in Table 72, the ICER increased slightly after this update.

**Table 72. Analysis applying age-related utility decrease**

| | ICER (pairwise comparison with ibrutinib) | ICER (pairwise comparison with BR) | ICER (pairwise comparison with DRC) |
|---|---|---|---|
| No age-related utilities/disutilities (CS base-case analysis) | [redacted] | [redacted] | [redacted] |
| Applying utility multipliers per Ara and Brazier 2010 | [redacted] | [redacted] | [redacted] |

Abbreviations: BR = rituximab and bendamustine ;CS = company submission; DRC = dexamethasone, rituximab and cyclophosphamide; ICER = incremental cost-effectiveness ratio
Notes: For transparency and simplicity, the results above were based on the base-case analysis in the CS, revised for this specific clarification question, without accounting for the revisions of model inputs/programming in response to other clarification questions. To generate the results according to different combinations of revisions, please refer to the updated Excel model

**B13.** Section B.3.4.5 describes how the utility for the post-progression population was calculated. Due to a lack of data from the ASPEN trial, the company uses the progression utility decrements from TA491 and TA502 to calculate the quality of life of patients in the post-progression health state.

a. In reference to the company submissions of these appraisals the current model therefore applies a utility decrement of -0.1 to the pre-progression health state to calculate the post-progression health state. The final appraisal documents of both TA491 and TA502 note that the utility decrement for the post-progression state applied in the respective company submissions was too small. The post-progression state was therefore decreased to 0.6 in both previous technical appraisals. In TA502 this corresponds to a utility decrement of .18 from the pre-progression health state to the post-progression health state. Please implement this change in current model.

b. In the economic model the standard error for the post-progression utility is the same as the standard error for progression free utility. It may be argued that the uncertainty around post-progression utility is larger due to the uncertainty around its estimation. Please comment on the appropriateness of using the same standard-error for post-progression utility as for progression-free utility. If it is found to be necessary, please explore a larger standard error for post-progression utility in a scenario analysis.

It is acknowledged that the utility decrement applied in the base case analysis of company submission was based on assumptions with uncertainties. Therefore, scenario analyses were conducted by varying the health state utility for post-progression survival (to 0.60 and 0.65 separately), as presented in Section B.3.8.3. The results of applying a utility decrement of 0.18 are summarised in Table 73; this assumption had limited impact on the ICERs.

**Table 73. Scenario analysis: utility decrement of 0.18 for post-progression survival**

| | ICER (pairwise comparison with ibrutinib) | ICER (pairwise comparison with BR) | ICER (pairwise comparison with DRC) |
|---|---|---|---|
| Utility decrement of 0.10 (CS base-case analysis) | [redacted] | [redacted] | [redacted] |
| Utility decrement of 0.18 | [redacted] | [redacted] | [redacted] |

Abbreviations: BR = rituximab and bendamustine ;CS = company submission; DRC = dexamethasone, rituximab and cyclophosphamide; ICER = incremental cost-effectiveness ratio
Notes: For transparency and simplicity, the results above were based on the base-case analysis in the CS, revised for this specific clarification question, without accounting for the revisions of model inputs/programming in response to other clarification questions. To generate the results according to different combinations of revisions, please refer to the updated Excel model

Due to a lack of available data to inform the post-progression utility, using any other standard error (i.e., not assuming the same standard error as the pre-progression utility) is considered to be arbitrary. Instead of varying the standard error for post-progression utility in a scenario analysis, a more straightforward approach would be to vary the absolute value of post-progression utility in scenario analyses based on the post-progression utility applied in other appraisals for similar indications, as was done above.

**B14.** In the base-case analysis, utility values are set to be treatment-independent. However, when set to treatment-specific utility values in the economic model, there is no difference in utility values in 3 out of 4 comparators.

a. Please justify the assumption of treatment independent utility values in the base-case.

b. Please explain why the differences in utility values between intervention and comparators are minimal when treatment-specific utility values are assumed.

The company assumed that the question was intended to be “there is no difference in utility values in 3 out of 4 *treatments*”, including zanubrutinib (the intervention), ibrutinib, BR and DRC.

In the base case analysis of company submission, the estimation of total QALYs accounted for two components:

1. Treatment-independent health state utilities (estimated based on the health-related quality-of-life data collected in ASPEN for patients treated with zanubrutinib and ibrutinib) and

2. Treatment-specific AE-related utility decrements (estimated based on treatment-specific AE incidence and AE-specific utility decrements)

The settings of the base-case analysis above were considered to be the least prone to bias against any treatments, based on the data availabilities for each treatment, as summarised below:

- Treatment-specific utilities were available for zanubrutinib and ibrutinib, based on ASPEN trial data (see Section B.3.4 of the company submission).

- Treatment-specific utilities were not available for BR and DRC that were not included in ASPEN but were based on clinical studies identified from the SLR, in which no health-related quality-of-date was (publicly) available or collected.

- There was a lack of published utility/disutility studies in general for the WM population identified by the SLR (see Appendix H of the company submission).

Despite the use of treatment-independent health state utilities in the base-case analysis (driven by the data availabilities), the model was developed to allow for treatment-specific utilities if further health-related quality-of-life data become available in the future. Until then, the utility estimate for zanubrutinib (0.791) was used for BR and DRC in the model, with a note left in the reference cell (right next to the input value cell) to highlight that it was an assumption rather than a robust utility value for BR or DRC.

## *Resources and costs*

**B15.** Table 3.28 describes costs for the treatment of AEs. These costs are broken down into costs for infections and costs for AEs which are not infections.

a. Please justify the assumption that all infections/ non-infection AEs accumulate the same cost.

b. Please break down the costs for AEs further.

Costs of AEs (presented int Table B.3.28 of the company submission) were collected upfront during the review of prior economic models, including NICE TA491, in which all AEs were assigned the same cost, except for non-pneumonia infection where a separate cost was applied. However, when the type of AEs included in the cost-effectiveness analysis was later restricted to those of Grade ≥3 that occurred in ≥5% of patients in any treatment arm, the cost of non-pneumonia infection AEs was no longer relevant. As a result, all AEs presented in Table B.3.19 were considered non-infection AEs in terms of costing in the cost-effectiveness model.

In summary, although the cost for infection AEs was presented in Table B.3.28, it was not actually applied in the cost-effectiveness analysis. Only the cost for non-infection AEs was applied in the cost-effectiveness analysis and was applied to all the AEs presented in Table B.3.19.

Exploratory scenario analyses have been conducted by using AE-specific costs that from NICE TA502 (ibrutinib for treating relapsed or refractory MCL), with results presented in Table 74.

**Table 74. Scenario analyses: AE-specific costs from NICE TA502**

| Event | CS base-case analysis | | | Scenario analyses | | |
|---|---|---|---|---|---|---|
| | Cost (2020 £) | Source | ICER | Cost (2020 £) | Source | ICER |
| Anaemia | £179.94 | NHS reference cost (based on the HRG codes used in TA491); post inflation | Vs ibrutinib: [redacted]; Vs BR: [redacted] Vs DRC: [redacted] | £175.79 | TA502; post inflation | Vs ibrutinib: [redacted]; Vs BR: [redacted] Vs DRC: [redacted] |
| Hypertension | | | | £175.79 | Assumption | |
| Neutropenia | | | | £175.79 | TA502; post inflation | |
| Pneumonia | | | | £2,720 | TA502; post inflation | |
| Thrombocytopenia | | | | £175.79 | Assumption | |

Abbreviations: BR = rituximab and bendamustine ;CS = company submission; DRC = dexamethasone, rituximab and cyclophosphamide; ICER = incremental cost-effectiveness ratio
Notes: For transparency and simplicity, the results above were based on the base-case analysis in the CS, revised for this specific clarification question, without accounting for the revisions of model inputs/programming in response to other clarification questions. To generate the results according to different combinations of revisions, please refer to the updated Excel model

## *Base case and sensitivity analysis*

**B16.** Section B.3.8 describes the sensitivity analyses conducted by the company. Uncertainty around the partitioned survival model was not expressed in the scenario analyses by quantifying the impact of the use of different PFS, OS and TTD curves. Due to the immature evidence considerable uncertainty could exist around the survival curves. Exploring this uncertainty by implementing different survival curves may be valuable. Please conduct scenario analyses to express the uncertainty around the survival curves.

As partially discussed in the response to clarification question B5b, scenario analyses were conducted in the company submission by exploring alternative parametric models that were considered to be clinically plausible based on clinical expert opinion regarding the mean OS and hazard patterns. The exact parametric distributions examined in scenario analyses and the corresponding rationales were summarised in Table B.3.5 and detailed further in Section B.3.3.2. In addition, the uncertainties of survival parameters were also examined in the probabilistic sensitivity analyses through Cholesky decomposition.

**B17.** The model is programmed to allow a maximum number of 1,000 iterations for the probabilistic sensitivity analysis. The ERG is concerned that this may not be

sufficient. Please provide convergence plots for the probabilistic sensitivity analysis for incremental costs and effects separately using at least 5,000 simulations.

The model has been updated to allow for a maximum of 5,000 iterations. The company re-ran the PSA using 5,000 simulations, the results of which were consistent with the results using 1,000 simulations that the probabilistic mean costs and QALYs were close to the deterministic estimates. Such results further demonstrated that the cost-effectiveness analysis was structurally stable.

Given that the PSA is subject to change whenever any model input/setting is changed and the relatively longer time required to run 5,000 simulations (relative to 1,000 simulations), no convergence plot is provided here, but may be generated whenever needed in the 'PSA' tab in the updated Excel model.

**B18.** In Section B.3.7.1 the base-case results are presented. No fully incremental analysis (as per NICE reference case) presenting the calculation of incremental QALY gains and costs along treatment options ranked by ascending cost was done. Please provide fully incremental analyses of treatments included in the NICE final scope.

Please refer to the response to clarification question B1.

## *Validation*

**B19. Priority Question: Based on the tornado diagrams resulting from the deterministic sensitivity analysis, age and gender seem to be influential parameters. Please explain why this is the case.**

Age and gender affect the results through their impact on the age- and gender-adjusted background mortality. However, despite that age and gender appear to be influential parameters, it is important to note that those are *relatively* influential compared with other parameters and are associated with very narrow ranges of

outcomes. When considering age and gender alone, both parameters had limited impact on the results.

**B20.** Please provide any detail on internal validation exercises performed, for example by completing the TECH-VER checklist (Büyükkaramikli et al, 2019 TECH-VER A Verification Checklist to Reduce Errors in Models and Improve Their Credibility. Pharmacoeconomics 2019 Nov;37 (11):1391-1408)

A checklist and results of validation conducted are presented in Table 75.

**Table 75. Internal validation**

| Test description | Test results and documentation |
|---|---|
| *Pre-analysis calculations* | |
| Does the technology (drug/device, etc.) acquisition cost increase with higher prices? | Yes, when increasing the price of zanubrutinib, the acquisition cost increases. |
| Does the drug acquisition cost increase for higher weight or body surface area? | Yes, when increasing the mean body surface area, the acquisition cost increases for chemo-immunotherapies. |
| In a partitioned survival model, does the progression-free survival curve or the time on treatment curve cross the overall survival curve? | No, the curves did not cross in all three comparisons based on the graphical demonstration in 'Survival' sheet, as the PFS and TTD curves were adjusted such that neither curve would cross the OS curve. |
| If survival parametric distributions are used in the extrapolations or time-to-event calculations, can the formulae used for the Weibull (generalized gamma) distribution generate the values obtained from the exponential (Weibull or Gamma) distribution(s) after replacing/transforming some of the parameters? | Yes |
| Is the HR calculated from Cox proportional hazards model applied on top of the parametric distribution extrapolation found from the survival regression? | Cox proportion hazards models were not applied on top of the existing parametric models (including exponential, Weibull, log-normal, etc.) |
| For the treatment effect inputs, if the model uses outputs from WINBUGS, are the OR, HR, and RR values all within plausible ranges? (Should all be non-negative and the average of these WINBUGS outputs should give the mean treatment effect) | Not applicable. No NMA was conducted. |
| *Event-state calculations* | |
| Calculate the sum of the number of patients at each health state | Yes |
| Check if all probabilities and number of patients in a state are greater than or equal to 0 | Yes |
| Check if all probabilities are smaller than or equal to 1 | Yes |
| Compare the number of dead (or any absorbing state) patients in a period with the number of dead (or any absorbing state) patients in the previous periods? | Yes |
| In case of lifetime horizon, check if all patients are dead at the end of the time horizon | Yes, for the majority (>97%) of the patients. |

| Test description | Test results and documentation |
|---|---|
| *Discrete event simulation specific:* Sample one of the 'time to event' types used in the simulation from the specified distribution. Plot the samples and compare the mean and the variance from the sample | Not applicable |
| Set all utilities to 1 | After setting all utilities to1, QALYs=LYs |
| Set all utilities to 0 | No QALYs accumulated over the time horizon |
| Decrease all state utilities simultaneously (but keep event-based utility decrements constant) | Lower utilities were accumulated over the model horizon |
| Set all costs to 0 | No costs were accumulated at any time over the model horizon |
| Put mortality rates to 0 | Patients never die over the model horizon |
| Put mortality rate at extremely high | Patients die in the first few cycles. In this model, when mortality was set to be 1, all patients die at first cycle |
| Set the effectiveness-, utility-, and safety-related model inputs for all treatment options equal | This test generated the same total life-years and QALYs for all treatment options |
| In addition to the inputs above, set cost-related model inputs for all treatment options equal | This test generated the same total costs for all treatment options |
| Change around the effectiveness-, utility- and safety-related model inputs between two treatment options | The total life-years and QALYs were then reversed between two treatment options. |
| Check if the number of alive patients estimated at any cycle is in line with general population life-table statistics | Yes, driven by the fact that background mortality per general population life table was accounted for during the model development. |
| Check if the QALY estimate at any cycle is in line with general population utility estimates | This was discussed in the response to clarification question B12. |
| Set the inflation rate for the previous year higher | The inflation rates were based on Curtis, Lesley A. and Burns, Amanda (2019) Unit Costs of Health and Social Care. The inflation rate in any specific were not necessarily always higher than that in the previous year. |
| Calculate the sum of all ingoing and outgoing transition probabilities of a state in a given cycle | Not applicable, as this is a PSM rather than a STM. |
| Calculate the number of patients entering and leaving a tunnel state throughout the time horizon | Not applicable |
| Check if the time conversions for probabilities were conducted correctly. | Not applicable |
| *Decision tree specific:* Calculate the sum of the expected probabilities of the terminal nodes | Not applicable |
| *Patient-level model specific:* Check if common random numbers are maintained for sampling for the treatment arms | Not applicable |
| *Patient-level model specific:* Check if correlation in patient characteristics is taken into account when determining starting population | Not applicable |
| Increase the treatment acquisition cost | Costs accumulated at a given time increased during the period when the treatment is administered |
| *Population model specific:* Set the mortality and incidence rates to 0 | Not applicable |
| *Result calculations* | |
| Check the incremental life-years and QALYs gained results. Are they in line with the comparative clinical effectiveness evidence of the treatments involved? | Zanubrutinib generated positive life-years and QALYs when compared with ibrutinib, BR and DRC, which was in line with the comparative clinical effectiveness. |

| Test description | Test results and documentation |
|---|---|
| Check the incremental cost results. Are they in line with the treatment costs? | Since zanubrutinib was associated with higher acquisition cost, the incremental costs compared with ibrutinib/BR/DRC was positive. |
| Total life years greater than the total QALYs | Yes |
| Undiscounted results greater than the discounted results | Yes |
| Divide undiscounted total QALYs by undiscounted life years | This value were within the outer ranges (maximum and minimum) of all the utility value inputs |
| Subgroup analysis results: How do the outcomes change if the characteristics of the baseline change? | The baseline patient characteristics had limited impact on the results in this case. |
| Could you generate all the results in the report from the model (including the uncertainty analysis results)? | Yes |
| Do the total life-years, QALYs, and costs decrease if a shorter time horizon is selected? | Yes |
| Is the reporting and contextualization of the incremental results correct? | Yes |
| Are the reported ICERs in the fully incremental analysis non-decreasing? | Not applicable, as the model relied on pairwise comparison per MAIC results. Additional discussions surrounding the possibility of a full incremental analysis are provided in the response to clarification question B1. |
| If disentangled results are presented, do they sum up to the total results (e.g. different cost types sum up to the total costs estimate)? | Yes, all cost/LY/QALY category sum up to the total estimates. |
| Check if half-cycle correction is implemented correctly (total life-years with half-cycle correction should be lower than without) | The half-cycle correction implementation was correct. |
| Check the discounted value of costs/QALYs after 2 years | Discounted value = undiscounted/$(1 + r)^2$ |
| Set discount rates to 0 | The discounted and undiscounted results were equal. |
| Set mortality rate to 0 | The undiscounted total life-years per patient were equal to the length of the time horizon |
| Put the consequence of adverse event/discontinuation to 0 (0 costs and 0 mortality/utility decrements) | The AE-related total costs and QALY losses became 0 then. |
| Divide total undiscounted treatment acquisition costs by the average duration on treatment | This result was aligned with the per-month/cycle drug acquisition costs. |
| Set discount rates to a higher value | Total discounted results decreased |
| Set discount rates of costs/effects to an extremely high value | Total discounted results were approximately the same as the discounted results accrued in the first few cycles |
| Put adverse event/discontinuation rates to 0 and then to an extremely high level | The total costs will be lower and QALYs/LYs will be higher when adverse event rates were 0; when AE rates were extremely high, there were higher costs and lower QALYs/LYs. |
| Double the difference in efficacy and safety between the new intervention and comparator, and report the incremental results | This is not applicable (at least not in a straightforward way) for the efficacy which replied on parametric survival models.<br><br>For safety, when the difference in AE incidence was doubled, the difference in total AE-related costs and QALY losses doubled. |

| Test description | Test results and documentation |
|---|---|
| Do the same for a scenario in which the difference in efficacy and safety is halved | Similar to the above |
| *Uncertainty analysis calculations* | |
| Are all necessary parameters subject to uncertainty included in the OWSA? | Yes |
| Check if the OWSA includes any parameters associated with joint uncertainty (e.g. parts of a utility regression equation, survival curves with multiple parameters) | No |
| Are the upper and lower bounds used in the one-way sensitivity analysis using confidence intervals based on the statistical distribution assumed for that parameter? | Yes; for those parameters without confidence interval reported, 20% standard error was assumed for sensitivity analysis. |
| Are the resulting ICER, incremental costs/QALYs with upper and lower bound of a parameter plausible and in line with a priori expectations? | Yes |
| Check that all parameters used in the sensitivity analysis have appropriate associated distributions – upper and lower bounds should surround the deterministic value (i.e. upper bound ≥ mean ≥ lower bound) | Yes |
| Standard error and not standard deviation used in sampling | Yes. |
| Lognormal/gamma distribution for HRs and costs/resource use | Yes, gamma distribution was used for resource use. Lognormal distribution was not applicable as HRs were not explicitly applied as inputs in the model. |
| Beta for utilities and proportions/probabilities | Yes |
| Dirichlet for multinomial | Not applicable. |
| Multivariate normal for correlated inputs (e.g. survival curve or regression parameters) | Yes |
| Normal for other variables as long as samples do not violate the requirement to remain positive when appropriate | Yes |
| Check PSA output mean costs, QALYs, and ICER compared with the deterministic results. Is there a large discrepancy? | The PSA output and deterministic results were generally consistent. |
| If you take new PSA runs from the Microsoft Excel model do you get similar results? | Yes |
| Is(are) the CEAC line(s) in line with the CE scatter plots and the efficient frontier? | Yes |
| Does the PSA cloud demonstrate an unexpected behavior or have an unusual shape? | No |
| Is the sum of all CEAC lines equal to 1 for all WTP values? | Yes |
| Do the explored scenario analyses provide a balanced view on the structural uncertainty (i.e. not always looking at more optimistic scenarios)? | Yes |
| Are the scenario analysis results plausible and in line with a priori expectations? | Yes |
| Check the correlation between two PSA results (i.e. costs/QALYs under the SoC and costs/QALYs under the comparator) | The correlation between costs and QALYs were reasonable for any treatment arms, judging by the scatter plots for total costs and QALYs |
| If a certain seed is used for random number generation (or previously generated random | Not applicable |

| Test description | Test results and documentation |
|---|---|
| numbers are used), check if they are scattered evenly between 0 and 1 when they are plotted | |
| Compare the mean of the parameter samples generated by the model against the point estimate for that parameter; use graphical methods to examine distributions, functions | Because the scatter plots for total costs and QALYs of the PSA were aligned with the deterministic estimates, no graphic methods were further done for parameter estimates. |
| Check if sensitivity analyses include any parameters associated with methodological/structural uncertainty (e.g. annual discount rates, time horizon) | No, they were included in the scenario analyses. |
| Value of information analysis if applicable: Was this implemented correctly? | Not applicable |
| Which types of analysis? Were aggregated parameters used? Which parameters are grouped together? Does it match the write-up's suggestions? | Yes |
| Is EVPI larger than all individual EVPPIs? | Not applicable |
| Is EVPPI for a (group of) parameters larger than the EVSI of that (group) of parameter(s)? | Not applicable |
| Are the results from EVPPI in line with OWSA or other parameter importance analysis (e.g. ANCOVA)? | Not applicable |
| Did the electronic model pass the black-box tests of the previous verification stages in all PSA iterations and in all scenario analysis settings? (Additional macro can be embedded to the PSA code, which stops the PSA when an error such as negative transition probability is detected) | Yes |
| Check if all sampled input parameters in the PSA are correctly linked to the corresponding event/state calculations | Yes |

AE = adverse event; ANCOVA = analysis of covariance; BR = rituximab and bendamustine; CE = cost-effectiveness; CEAC = cost-effectiveness acceptability curve; DRC = dexamethasone, rituximab and cyclophosphamide; EVPI = expected value of perfect information; EVPPI = expected value of partially perfect information; HR = hazard ratio; ICER = incremental cost-effectiveness ratio; LY = life year; MAIC = matching adjusted indirect comparison; NMA = network meta-analysis; OR = odds ratio; OS = overall survival; OWSA = one-way sensitivity analysis; PFS = progression-free survival; PSA = probabilistic sensitivity analysis; PSM = partitioned survival model; QALY = quality-adjusted life year; RR = relative risk; SA = scenario analysis; SoC = standard of care; STM = state transition model; TTD = time to discontinuation; WTP = willingness-to-pay

**B21.** Please provide cross validations, i.e. comparisons with other relevant NICE technology appraisals focussed on similar, potentially relevant, diseases (e.g. TA491 and TA502) and elaborate on the identified differences regarding:

a. Model structure and assumptions
b. Input parameters related to:
    I. Clinical effectiveness
    II. Health state utility values
    III. Resource use and costs
c. Estimated outcomes per comparator/ intervention

I. Life years
II. QALYs
III. Costs

A summary of the comparisons with TA491 (ibrutinib for treating WM) and TA502 (ibrutinib for treating MCL) is provided in Table 76

**Table 76. Comparison of the current appraisal versus TA491 and TA502**

| | TA491 | TA502 | ID1427 | Comments |
|---|---|---|---|---|
| Model structure | Markov model | Markov model | Partitioned survival model | |
| Clinical effectiveness | Study 1118E European chart review study | One RCT comparing ibrutinib versus temsirolimus - MCL3001<br>Two single-arm studies (PCYC1104 and SPARK) | ASPEN trial<br>Tedeschi et al. 2015<br>Dimopoulos et al. 2007/Kastritis et al. 2015 | |
| Utility | 2/3/4-line: 0.799, based on the RESONATE for R/R CLL<br>BSC: This was calculated by applying a utility decrement of 12.8% to the baseline utility of 0.763 generated from the RESONATE EQ-5D-5L data for R/R CLL. This percentage utility decrement was derived from Beusterien et al. (2010), a time trade-off QoL study carried out to ascertain CLL utilities in the UK | PFS: 0.78<br>PPS: 0.68<br>R-chemo decrement: 0.2<br>Pooled data was based on RAY (MCL3001) and SPARK (MCL2001)<br>The decrement upon progression predicted using these data (0.1) is considered to be reasonable in light of "upon progression" decrements in other haematological cancers that have been used in previous NICE submissions. | PFS: 0.791, based on ASPEN trial<br>PPS: 0.691. State utility was estimated under the assumption of 0.1 decrement relative to PFS utility.<br>0.1 was based on the utility decrements for progression applied in NICE TA502 (0.10) for ibrutinib in MCL and TA429 (0.098) for ibrutinib in CLL. | The utility estimates before disease progression were generally consistent across three TA.<br>0.1 is a commonly used utility decrement in TA502 and ID1427. |
| Resource use and costs | Frequency of use of resources over time (questionnaire survey conducted by company)<br>Unit cost for resource use (NHS reference costs 2014/2015)<br>Unplanned Event Related Medical Resource Utilisation | Drug acquisition cost (MIMS Online, MIMS Online)<br>Drug administration cost (NHS reference costs 2014/2015)<br>Total annual resource use by health state and response status (clinicians' feedback, NHS reference costs | Drug acquisition cost (BNF)<br>Drug administration cost (NHS reference costs 2018/2019)<br>Frequencies and unit costs of resource use for routine care (NICE TA491 for ibrutinib in WM for frequency and NHS reference costs | The cost category and unit cost data source were generally consistent across these submissions. With NHS reference cost serving as the source for unit cost for medical services, PSSRU |

| | TA491 | TA502 | ID1427 | Comments |
|---|---|---|---|---|
| | (clinical experts' opinion)<br>Intervention and comparator costs (British national Formulary and NHS reference cost 2014/2015)<br>Health state e unit costs and resource use (BNF, NHS reference cost 2014/2015)<br>Adverse event cost (NHS reference cost 2014/2015)<br>Terminal care cost (PSSRU and Round et al 2015) | (2014/2015) and the PSSRU 2015)<br>Health state cost (Model calculations)<br>Adverse event cost (NHS reference cost 2014/2015)<br>Terminal care cost (Nuffield et al. 2014 and PSSRU 2015) | 2018/2019 for unit costs)<br>Adverse event cost (NHS reference cost 2018/2019)<br>Subsequent treatment use and distribution (the proportion of patients receiving subsequent treatment upon progression was obtained from NICE TA491 for ibrutinib in WM; distribution of subsequent treatments was based on WM Rory Morrison Registry report)<br>Terminal care cost (PSSRU 2019 and Round et al.2015) | for cost inflation, and BNF for drug acquisition cost |
| Life years | NR | 1.23 (incremental) | Zanu vs Ibru: [redacted]<br>Zanu vs BR: [redacted]<br>Zanu vs DRC: [redacted] | |
| QALYs | NR | 0.94 (incremental) | Zanu vs Ibru: [redacted]<br>Zanu vs BR: [redacted]<br>Zanu vs DRC: [redacted] | |
| Costs | NR | 69,528 (incremental) | Zanu vs Ibru: [redacted]<br>Zanu vs BR: [redacted]<br>Zanu vs DRC: [redacted] | |

Abbreviations: BNF = British National Formulary; BR = rituximab and bendamustine; BSC = best supportive care; CLL = chronic lymphocytic leukaemia; DRC = dexamethasone, rituximab and cyclophosphamide; EQ-5D-5L = EuroQol 5-Dimensions 5-Level; MCL = mantle cell lymphoma; MIMS = Monthly Index of Medical Specialities; NR = not reported; PPS = post-progression survival; PSSRU = Personal Social Services Research Unit; QALY = quality-adjusted life year; QoL = quality of life; R/R = relapsed/refractory; RCT = randomised controlled trial; TA = technology appraisal; UK = United Kingdom; WM = Waldenström's macroglobulinaemia

# References


1. © National Institute for Health and Care Excellence. Developing NICE guidelines: the manual [PMG20]. 2020. https://www.nice.org.uk/process/pmg20/chapter/identifying-the-evidence-literature-searching-and-evidence-submission#developing-search-strategies (accessed 2021 April).
2. Higgins J, Thomas J, Chandler J, et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.2. 2021. www.training.cochrane.org/handbook (accessed April 2021).
3. WMUK. First UK Waldenström's Macroglobulinaemia Registry Report. 2018. https://www.wmuk.org.uk/sites/default/files/2020-01/The%20Rory%20Morrison%20WM%20Registry%20Report%20%281%29%20-%20March%202018.pdf (accessed August 2020).
4. BeiGene. Clinical Study Report: ASPEN BGB-3111-302, June 2020.
5. Treon SP, Xu L, Guerrera ML, et al. Genomic Landscape of Waldenstrom Macroglobulinemia and Its Impact on Treatment Strategies. *J Clin Oncol* 2020; **38**(11): 1198-208.
6. Odutayo A, Wong CX, Hsiao AJ, Hopewell S, Altman DG, Emdin CA. Atrial fibrillation and risks of cardiovascular disease, renal disease, and death: systematic review and meta-analysis. *BMJ* 2016; **354**: i4482.
7. BeiGene. Zanubrutinib Summary of Product Characteristics. 2020.
8. Paydas S. Management of adverse effects/toxicity of ibrutinib. *Crit Rev Oncol Hematol* 2019; **136**: 56-63.
9. Janssen-Cilag. Ibrutinib Summary of Product Characteristics. 2021. https://www.ema.europa.eu/en/documents/product-information/imbruvica-epar-product-information_en.pdf (accessed February 2021).
10. Dimopoulos MA, Anagnostopoulos A, Kyrtsonis MC, et al. Primary treatment of Waldenström macroglobulinemia with dexamethasone, rituximab, and cyclophosphamide. *J Clin Oncol* 2007; **25**(22): 3344-9.
11. Kastritis E, Gavriatopoulou M, Kyrtsonis MC, et al. Dexamethasone, rituximab, and cyclophosphamide as primary treatment of Waldenström macroglobulinemia: final analysis of a phase 2 study. *Blood* 2015; **126**(11): 1392-4.
12. Tedeschi A, Picardi P, Ferrero S, et al. Bendamustine and rituximab combination is safe and effective as salvage regimen in Waldenstrom macroglobulinemia. *Leuk Lymphoma* 2015; **56**(9): 2637-42.
13. Paludo J, Abeykoon JP, Shreders A, et al. Bendamustine and rituximab (BR) versus dexamethasone, rituximab, and cyclophosphamide (DRC) in patients with Waldenstrom macroglobulinemia. *Ann Hematol* 2018; **97**(8): 1417-25.
14. Castillo JJ, Gustine JN, Meid K, et al. Response and survival for primary therapy combination regimens and maintenance rituximab in Waldenstrom macroglobulinaemia. *Br J Haematol* 2018; **181**(1): 77-85.
15. Paludo J, Abeykoon JP, Kumar S, et al. Dexamethasone, rituximab and cyclophosphamide for relapsed and/or refractory and treatment-naive patients with Waldenstrom macroglobulinemia. *Br J Haematol* 2017; **179**(1): 98-105.
16. Garcia-Sanz R, Dimopoulos MA, Lee H-P, et al. Updated results of the ASPEN trial from a cohort of patients with MYD88 wild-type (MYD88WT) Waldenström macroglobulinemia (WM). *Journal of Clinical Oncology* 2020; **38**(15_suppl): e20056-e.

17. Treon SP, Branagan AR, Ioakimidis L, et al. Long-term outcomes to fludarabine and rituximab in Waldenstrom macroglobulinemia. *Blood* 2009; **113**(16): 3673-8.
18. Tedeschi A, Benevolo G, Varettoni M, et al. Fludarabine plus cyclophosphamide and rituximab in Waldenstrom macroglobulinemia: an effective but myelosuppressive regimen to be offered to patients with advanced disease. *Cancer* 2012; **118**(2): 434-43.
19. Souchet L, Levy V, Ouzegdouh M, et al. Efficacy and long-term toxicity of the rituximab-fludarabine-cyclophosphamide combination therapy in Waldenstrom's macroglobulinemia. *Am J Hematol* 2016; **91**(8): 782-6.
20. Kyle RA, Greipp PR, Gertz MA, et al. Waldenstrom's macroglobulinaemia: a prospective study comparing daily with intermittent oral chlorambucil. *Br J Haematol* 2000; **108**(4): 737-42.
21. Gertz MA, Rue M, Blood E, Kaminer LS, Vesole DH, Greipp PR. Multicenter phase 2 trial of rituximab for Waldenstrom macroglobulinemia (WM): an Eastern Cooperative Oncology Group Study (E3A98). *Leuk Lymphoma* 2004; **45**(10): 2047-55.
22. Gertz MA, Abonour R, Heffner LT, Greipp PR, Uno H, Rajkumar SV. Clinical value of minor responses after 4 doses of rituximab in Waldenstrom macroglobulinaemia: a follow-up of the Eastern Cooperative Oncology Group E3A98 trial. *Br J Haematol* 2009; **147**(5): 677-80.
23. WMUK. Second UK Waldenström's Macroglobulinaemia Registry Report (personal communication prior to publication in 2021). 2021.
24. © National Institute for Health and Care Excellence. Ibrutinib for treating Waldenstrom's macroglobulinaemia [TA491]. November 2017. www.nice.org.uk/guidance/ta491 (accessed October 2020).
25. Buske C, Leblond V, Dimopoulos M, et al. Waldenstrom's macroglobulinaemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2013; **24 Suppl 6**: vi155-9.
26. BeiGene. Clinical Study Report: BGB-3111-AU-003 2020.

NICE National Institute for Health and Care Excellence

# Professional organisation submission

## Zanubrutinib for treating Waldenström's macroglobulinaemia [ID1427]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

**Information on completing this submission**

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

| **About you** | |
|---|---|
| 1. Your name | [redacted] |
| 2. Name of organisation | **British Society Haematology/ Royal College Pathologists** |

| | |
|---|---|
| 3. Job title or position | **Haematology Consultant, [redacted]** |
| 4. Are you (please tick all that apply): | ☒ an employee or representative of a healthcare professional organisation that represents clinicians?<br>☒ a specialist in the treatment of people with this condition?<br>☐ a specialist in the clinical evidence base for this condition or technology?<br>☐ other (please specify): |
| 5a. Brief description of the organisation (including who funds it). | **Registered charities, with members paying an annual subscription.** |
| 4b. Has the organisation received any funding from the manufacturer(s) of the technology and/or comparator products in the last 12 months? [Relevant manufacturers are listed in the appraisal matrix.] | **Not to the best of my knowledge** |

| | |
|---|---|
| If so, please state the name of manufacturer, amount, and purpose of funding. | |
| 5c. Do you have any direct or indirect links with, or funding from, the tobacco industry? | **no** |
| **The aim of treatment for this condition** | |
| 6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.) | - To improve the length of and quality of life of patients with symptomatic Waldenstrom macroglobulinaemia (WM) with minimal toxicity.<br>Symptoms can be related to the bone marrow failure that occurs due to marrow involvement by WM, due to lymphoma nodal involvement and then related to the paraprotein produced by the WM lymphoma cells. |
| 7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by | - Improvement in symptoms that led to treatment being required.<br>- Improvement in blood counts, transfusion independence<br>- Improvement in paraprotein level to reduce risk of complications, e.g. due to hyperviscosity<br>- Time to next treatment > 3 years in the R/R setting. |

| | |
|---|---|
| x cm, or a reduction in disease activity by a certain amount.) | |
| 8. In your view, is there an unmet need for patients and healthcare professionals in this condition? | Yes, there are no targeted therapies that are licensed and funded in the NHS for this disease, despite the increasing body of evidence that they are effective. This is in contrast to CLL which is targetable by many of the same agents, many of which are approved and funded for by the NHS. I believe as WM is a rarer disease, with few large international trials, it makes it difficult for drugs to be approved for this condition.<br><br>The risk of a number of the complications that can occur secondary to WM, e.g. increased risk of infections/ secondary malignancies can actually increase with chemoimmunotherapy, and given that many patients may have other health problems, chemoimmunotherapy is not always suitable. |
| **What is the expected place of the technology in current practice?** | |
| 9. How is the condition currently treated in the NHS? | Chemoimmunotherapy- most often Rituximab in combination with bendamustine or cyclophosphamide/ dexamethasone.<br><br>At relapse alternative chemotherapy regimens are used, although frequently ibrutinib is currently accessed via the CDF.<br><br>Clinical Trials |
| • Are any clinical guidelines used in the treatment of the condition, and if so, which? | BCSH guidelines (currently being updated)<br>Regional guidelines e.g. London Cancer Partners guidelines<br>International guidelines e.g. ESMO/NCCN |

| | |
|---|---|
| • 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.) | There is international consensus on what treatments have activity but availability of treatments differs from country to country and there is no well defined pathway as to optimal sequence of therapy.<br><br>Within this country there is variation in chemotherapy regimens used. |
| • What impact would the technology have on the current pathway of care? | It would allow patients who have relapsed or are refractory to one chemoimmunotherapy regimen to have a different technology to treat their cancer. Furthermore, it would allow patients who wouldn't be able to tolerate chemotherapy to have an effective treatment for their WM. |
| 10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice? | Yes, haematologists treating WM as well as other B cell malignancies are familiar with the use and toxicity profile of BTK inhibitors including zanubrutinib. |
| • How does healthcare resource use differ between the technology and current care? | This is a stepwise improvement in treatment options for patients with WM, as whilst BTK inhibitors may be available through other avenues e.g. clinical trials/CDF/ privately and thus many clinicians have experience of their use both for WM and other conditions, this would be the first BTK inhibitor to be approved and funded specifically for WM. |
| • In what clinical setting should the technology be used? (For example, | Secondary care |

| | |
|---|---|
| primary or secondary care, specialist clinics.) | |
| • What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.) | No other investment required for reasons as above. |
| 11. Do you expect the technology to provide clinically meaningful benefits compared with current care? | Yes, it gives patients a new class of drugs that we know are very active in B cell malignancies and in WM in particular that they would not otherwise have access to. This will help improve both progression free survival and overall survival from an efficacy perspective. Furthermore, there may be some patients in whom chemoimmunotherapy would not be suitable either if they had early relapse following initial treatment or due to toxicity concerns, and thus this would provide a new option with a different toxicity profile. |
| • Do you expect the technology to increase length of life more than current care? | Yes, there is data from the 3 year follow up of the phase 1 trial NCT02343120 which shows clear good PFS which is likely to translate to improved OS as it is a different treatment option for patients. |
| • Do you expect the technology to increase health-related quality of life more than current care? | Yes, due to this being an option that more patients will be able to tolerate than chemoimmunotherapy. Also as seen in the trial data patients QoL improved with effective treatment indicating that there was improvement in symptoms as well as response rate. |

| | |
|---|---|
| 12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population? | No- there has been suggestions from trial data with other BTK inhibitors that perhaps they are less effective in certain genomic subgroups of patients with WM, the trials with zanubrutinib do not show differences in outcomes between these genomic subgroups suggesting activity in all of them (although the trials were not powered to show differences) |
| **The use of the technology** | |
| 13. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.) | Easier- as this is oral therapy, the need to come into hospital becomes less and so patients can be more easily monitored remotely. Patients with WM are more susceptible to infections and by definition fall into the clinically extremely vulnerable group with regards to COVID and so having an oral option for patients many of whom will also have other risk factors for COVID makes it easier for them to access necessary treatment for their cancer without increasing their risk of exposure to COVID. By this time, many of these patients will have been vaccinated but we know that the vaccination is likely to be less effective in patients with WM. |

| | |
|---|---|
| 14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing? | Treatment will start as per standard guidelines, i.e. when there is a clinical indication.<br><br>Treatment will stop when the patient is no longer responding or getting intolerable toxicity. Standard monitoring will be ongoing whilst patient is on treatment to determine these factors. |
| 15. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation? | Avoiding the need to come in for chemotherapy.<br><br>Some patients may not be suitable for chemoimmunotherapy and it opens up an option for treatment for these patients. |
| 16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it | Yes, it gives a different treatment options for patients who would have just had chemoimmunotherapy options. This provides a different way of treating their disease that is well tolerated, leads to improvement in quality of life and from initial trial data would indicate that it is likely to lead to improved length of life. |

| | |
|---|---|
| improve the way that current need is met? | |
| • Is the technology a 'step-change' in the management of the condition? | yes |
| • Does the use of the technology address any particular unmet need of the patient population? | Yes, it gives patients access to a class of drugs that we know is active for WM. |
| 17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life? | The trial data for zanubrutinib shows it is quite well tolerated with an acceptable toxicity profile that can be managed. Cytopenias, diarrhoea and infections are seen usually grade 1-2 in trials and less frequently at grade 3, but we are able to manage these effectively for the majority of patients to continue on treatment. |
| **Sources of evidence** | |
| 18. Do the clinical trials on the technology reflect current UK clinical practice? | No head to head comparison with chemoimmunotherapy.<br><br>Aspen trial reports outcomes compared to another BTK inhibitor that is currently available via the CDF for treatment of R/R WM. |

| | |
|---|---|
| | However I think the trial population investigated in the single arm phase 1/2 study and also the follow on ASPEN study do reflect the population in the UK. Whilst the median age of the patient population was perhaps younger as is often seen, the maximum age was 87 and 90 with a range of performance scores and also some were heavily pretreated. |
| • If not, how could the results be extrapolated to the UK setting? | Modelling. Comparison to real world data with alternative regimens |
| • What, in your view, are the most important outcomes, and were they measured in the trials? | Time to next treatment (no, duration of response and progression free survival was measured, but sometimes patients can progress but still not require therapy immediately).<br><br>Improvement in symptoms (yes)<br><br>Safety analysis (yes) |
| • If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes? | Primary outcome measured was response rate (CR/VGPR) which can be used as a surrogate endpoint for progression free survival with chemoimmunotherapy but perhaps is less predictive for this class of drugs as seen in patients with CLL, and also in a recent retrospective study relating response rate for patients taking ibrutinib with PFS (Castillo et al BJHaem 2020). Thus this may actually underestimate the impact of BTK inhibitors on more meaningful outcomes such as PFS, time to next treatment, in that even those who achieve a partial response as opposed to VGPR or complete response will have a meaningful response. |

| | |
|---|---|
| • Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently? | Not to the best of my knowledge. Zanubrutinib is being trialled in other B cell malignancies as well as WM, and the toxicity profile is similar across trials and real world data in all these disease groups. |
| 19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence? | no |
| 20. How do data on real-world experience compare with the trial data? | No real world data with zanubrutinib specifically, but there is a lot of experience with this class of drugs in WM, and the use of them also in other disease areas. The AEs reported in the trials reflect that which was already seen and known with this class of drugs. Indeed, the ASPEN trial indicated that zanubrutinib had a more favourable safety profile than an alternative BTK inhibitor which anecdotally is our experience too. |
| **Equality** | |
| 21a. Are there any potential equality issues that should be taken into account when considering this treatment? | no |

| | |
|---|---|
| 21b. Consider whether these issues are different from issues with current care and why. | n/a |

**Key messages**

22. In up to 5 bullet points, please summarise the key messages of your submission.

- Novel therapeutic option for patients who require alternative treatment regimens to chemoimmunotherapy
- Well tolerated drug that leads to durable periods of time before progression
- Associated with improvement in quality of life for patients
- Oral therapy that allows remote monitoring
- Trial outcomes are very favourable and that just looking at response rate and depth of response may underestimate the activity and durability of this treatment option.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

……

**Your privacy**

The information that you provide on this form will be used to contact you about the topic above.

☒ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

NICE National Institute for Health and Care Excellence

**Patient organisation submission**

**Zanubrutinib for treating Waldenström's macroglobulinaemia [ID1427]**

| **About you** | |
|---|---|
| 1.Your name | **Will Franks** |
| 2. Name of organisation | Joint submission on behalf of WMUK and Lymphoma Action |
| 3. Job title or position | Chair of Trustees for WMUK |
| 4a. Brief description of the organisation (including who funds it). How many members does it have? | **WMUK** (https://www.wmuk.org.uk/), a registered Charity in England and Wales, is a patient orientated organisation focused solely on those impacted by Waldenström's macroglobulinaemia (WM). The charity currently has 1045 members. The goals of the charity are optimising access to accurate diagnosis & high-quality care, access to personalised information & support, access to new treatment, and research that matters to patients.<br><br>WMUK is primarily funded by charitable fundraising events and donation from patients, carers, family and friends, and other members of the general public. Some donations are received from pharmaceutical companies primarily to support events such as the charity's annual Patient - Doctor Summit.<br><br>**Lymphoma Action** (https://lymphoma-action.org.uk/) is a national charity, established in 1986, registered in England and Wales and in Scotland.<br><br>We provide high quality information, advice and support to people affected by lymphoma – the 5th most common cancer in the UK.<br><br>We also provide education, training and support to healthcare practitioners caring for lymphoma patients. In addition, we engage in policy and lobbying work at government level and within the National Health Service with the aim of improving the patient journey and experience of people affected by lymphoma. We are the only charity in the UK dedicated to lymphoma. Our mission is to make sure no one faces lymphoma alone. |

| | |
|---|---|
| | Our work is made possible by the generosity, commitment, passion and enthusiasm of all those who support us. We have a policy for working with healthcare and pharmaceutical companies – those that provide products, drugs or services to patients on a commercial or profit-making basis. This includes that no more than 20% of our income can come from these companies and there is a cap of £50k per company. Acceptance of donations does not mean that we endorse their products and under no circumstances can these companies influence our strategic direction, activities or the content of the information and support we provide to people affected by lymphoma. |
| 4b. Has the organisation received any funding from the manufacturer(s) of the technology and/or comparator products in the last 12 months? [Relevant manufacturers are listed in the appraisal matrix.]<br><br>If so, please state the name of manufacturer, amount, and purpose of funding. | **WMUK**<br>£2,967 from BeiGene as sponsorship towards the costs of the WMUK 2019 Patient - Doctor Summit.<br>£15,000 from Janssen in support of education and information activities.<br>Dr D'Sa (Trustee) has received grant funding for a research fellow £147k (2019-21) and is on the Medical Advisory Board of BeiGene UK Ltd.<br><br>**Lymphoma Action**<br>Janssen - £15,000 (support for education and information activities).<br>Roche Products - £20,000 (support for education and information activities). |
| 4c. Do you have any direct or indirect links with, or funding from, the tobacco industry? | None. |
| 5. How did you gather information about the | 1. We sought feedback directly from patients receiving zanubrutinib as part of the ongoing ASPEN Study of zanubrutinib vs ibrutinib in WM. We received feedback from eight respondents taking part in this trial.<br><br>2. We sought input from Dr Shirley D'Sa, who shared the experiences of some of the patients she treats for WM. |

| | |
|---|---|
| experiences of patients and carers to include in your submission? | a. Dr D'Sa is UK Chief Investigator for the ASPEN Study of zanubrutinib vs ibrutinib in WM, so has first-hand experience of treating patients with zanubrutinib.<br><br>b. Dr D'Sa is PI for the Rory Morrison Registry Project, which collects real world data on UK patients with WM (2016 ongoing), including patients treated with BTK inhibitors including zanubrutinib, and incorporates patient-related outcome measures.<br><br>3. As a charity with a wide reach for a rare disease, we have a continuous dialogue with patients and families affected by WM, which provides invaluable insights into real world experiences of those receiving BTK inhibitors as well as those on other therapies. This occurs via our advice line and portal as well as a moderated Facebook support group page. |
| **Living with the condition** | |
| 6. What is it like to live with the condition? What do carers experience when caring for someone with the condition? | WM develops over many months or years. It is associated with major disease-related symptoms that have a significant impact on the day-to-day lives of people with it. These include infections, weakness, extreme fatigue, breathlessness, and severe bone, joint and eye pain. People with WM have also reported fevers, night sweats, weight loss, and significant reduction in their mobility. Complications arising in WM patients include:<br><br>• Cryoglobulinaemia- when the IgM paraprotein has the property of precipitating in vivo and causing organ and tissue damage. Patients commonly suffer from kidneys problems, joint pain, cold feet or hands, skin ulceration, and nerve damage.<br><br>• Cold agglutinin disease - a condition in which the presence of a WM disease clone can promote a cold-mediated haemolytic anaemia due to red cell agglutination and complement fixation. The size of the clone may not be high, but the immunological consequences can be life-changing for patients. Symptoms include feeling weak and tired, dizziness and headaches, sore back, legs, or joints, irritability or changes in behaviour, pale or yellow skin, vomiting or diarrhea, cold feet or hands, and chest pains or an irregular heartbeat. Recurrent transfusions to replace the broken down red cells and the iron overload that ensues, as well as an increased rate of venous thromboembolism, can all be acutely life-threatening.<br><br>WM was traditionally viewed as a disease that affects people over the age of 65. However, this type of blood cancer is now increasingly seen in people of working age who are economically and socially active, often with young families. It can have a significant physical, psychological, social and financial impact. |

| | |
|---|---|
| | A WM patient in the UK (56) said: "Living with WM makes normal life very difficult. I've been living with constant leg and foot pain resulting from my first course of chemotherapy in 2013. Now that the WM disease burden is returning, along with cryoglobulinaemia, the debilitating symptoms have meant that I have had to stop work as the CEO of a tech company. This has had a big impact on my family's financial short- and long-term security, and continuing emotional impact on both me and my partner."<br><br>There is a significant psychological burden associated with WM, and patients frequently report emotional distress and poor mental health. One patient told us how her diagnosis has had a significant psychological impact and a "sense of loss for the future. This was not how I imagined life to be like as a young married couple.''<br><br>Daily symptoms such as fatigue, which can be intense and disabling, can also have a negative impact on quality of life.<br><br>In addition, 'watch and wait' is described as particularly stressful by patients and their carers who have to live with a high level of uncertainty, not knowing if or when they will need treatment. Having a diagnosis of cancer is life-changing and emotionally challenging. This burden is made heavier when patients are watching and waiting for symptoms to get worse and for treatment to start. Once patients do start treatment, they live with the constant threat of relapse and short or partial duration of response, as well as a high level of worry about what treatments will be available beyond first-line therapies. |
| **Current treatment of the condition in the NHS** | |
| 7. What do patients or carers think of current treatments and care available on the NHS? | **The following quotations have been taken word for word from Patient responses:**<br><br>**Patient R on the current zanubrutinib trial said:** "During and after chemotherapy treatment it is worrying: if the treatment is/has worked, was it the right choice, will it cause further damage, how long will it take to fully recover (about a year), how long will the treatment be effective? As a haematology patient you emerge unsupported and feel dumped by the NHS to live with the aftermath of treatment."<br><br>**Patient F currently on watch and wait said:** "As patients we have been living the best we can between cycles of chemotherapy based treatments and the resulting increasingly debilitating effects of both the disease and these more conventional treatments. Both patients and carers affected by this condition are acutely aware of the finite number of therapies available to us and, as treatment cycles take place, this narrows our choices as intolerance increases, or effectiveness diminishes. As patient and carers we live with this anxiety, and constantly monitor the new treatments |

| | |
|---|---|
| | emerging from the US, and lobby for the NHS to adopt the effective treatments for our quality of life and survival."<br><br>As we work towards more informed clinical practices across the UK through deployment of information (guidelines, seminars to patients and clinicians alike), certain chemotherapies (such as purine analogues) have fallen out of favour due to concerns about short- (immunosuppression) and long-term toxicities (secondary cancers).<br><br>Other effective agents such as bortezomib are not available on the NHS for patients with WM. Approaches such as high dose therapy and autologous stem cell transplantation are only recommended as possible salvage therapy in fitter patients. These are intensive treatments that typically require prolonged hospital stays and significant time off work. Many patients with WM are not fit enough to tolerate these intensive treatments and have more limited options.<br><br>Prior to the introduction of BTKi, the mainstay of therapies at front line and relapse available via the NHS was rituximab in combination with agents such as cyclophosphamide or bendamustine. Whilst these therapies are effective at the outset, patients are aware of the diminishing returns from repetitive use of such therapies, and actively seek the reassurance of further lines of therapy to keep them alive.<br><br>Current treatment choices available via the NHS do lag those licensed and available in other countries, or available privately. Zanubrutinib is one of these treatments and the ongoing ASPEN trial has literally been a life saver for some patients. We are very supportive and appreciative of the NICE assessment of zanubrutinib for the treatment of WM. We are also very appreciative of the NHS recommending ibrutinib through the Cancer Drugs Fund; this drug has been a game changer for an increasing number of our patients. |
| 8. Is there an unmet need for patients with this condition? | 1. Yes – as noted under Section 7, there is a limited range of therapies for WM, which inevitably cease to control the disease after sequential use. This results in an inexorable decline in health and well-being, increased health needs and untimely death.<br><br>2. There are a range of WM-related conditions that are not best served by chemoimmunotherapy, with a mismatch of intensity (see Section 11).<br><br>There is a clear, unmet need for an effective, well tolerated treatment that provides long-term disease control. |

| **Advantages of the technology** | |
|---|---|
| 9. What do patients or carers think are the advantages of the technology? | **The following quotations have been taken word for word (and anonymised ) from the written responses to the questionnaire sent to patients enrolled on the ASPEN Study of zanubrutinib vs ibrutinib in WM**<br><br>**Patient S:** "Zanubrutinib has been a complete game changer for me. It has controlled my WM extremely well, I have not felt this well in many years and I am so happy to have an oral daily drug to take in place of chemotherapy infusions. It has allowed me to carry on my daily activities in a way I have not been able to for many, many years. It has enabled me to participate fully with family, friends and social activities to the full as previously chemotherapy neutropenia placed severe limitations on my daily activities. My family have been very relieved to see the improvement in my ability to live a normal life and also to see an improvement in my mental well-being. Chemotherapy has led to much depression, impacts my psychological well-being."<br><br>**Patient M:** "Saved my life. I am chemo intolerant. My bone marrow was 80%, highest IGM 58.5, highest plasma viscosity 7.2, lowest Hb 65. Now, I am in remission, my paraprotein protein levels are 'too small to quantify' and I am very much enjoying living a normal healthy life again. Without question, taking zanubrutinib has saved my life and has enabled myself to be in remission from WM."<br><br>**Patient R:** "I am currently on the BGB-3111-302 clinical trial at UCLH and started cycle 1 on 29 December 2017 and I am now on cycle 37. For me this is a wonder drug and from the outset I was able to have an immediate benefit which has been ongoing. I now functioning as normal, viz. from being unable to climb the stairs at home in December 2014 without having to sit down on the bed I am now walking 100 miles per month with ease. Taking zanubrutinib orally twice a day bears no comparison whatsoever with any of the chemotherapy treatments I have received and which were very debilitating without my receiving any significant benefit."<br><br>**Patient B:** "Zanubrutinib is not as toxic as chemotherapy is. Zanubrutinib is very easy to take at home or when away visiting family. I have not needed to set timers/alarms to maintain taking the drug regularly. My personal confidence has been restored and I am able to use the skills and experience I have accumulated over my life. I have been able to support other patients and become a respected patient leader advising on improving cancer treatment and care outside WM, supporting work on national NHS priority projects, local public health messaging and helping the NHS recover from the first wave of COVID-19." |

| | |
|---|---|
| | **Patient A**: **"**Taking zanubrutinib has been nothing short of incredible. I have had no significant side effects and a complete response. It's simply a world away from chemo, let alone a stem cell transplant, with all the resultant side effects and hospitalisations. Aside from taking pills a couple of times a day, my life is normal."<br>"I can't stress enough how thankful I am to have been able to take part in the zanubrutinib trial. Before it started I couldn't have hoped for it to have gone any better. I very much hope that many others will be able to benefit from taking it as much as I have."<br><br>The availability of zanubrutinib has enabled patients to get back on track in their lives, both economically and socially, and, for most, returning to living their life as they would like to – with a sense of normality – is game-changing.<br><br>Patients we have surveyed consider Zanubrutinib an effective treatment, well tolerated, with rapid response, associated with an excellent QOL with limited side effects. This is more so with zanubrutinib than ibrutinib (*Blood.2020 Oct 29;136(18):2038-2050*).<br><br>With the ongoing COVID-19 pandemic, having home-based therapy that limits hospital visits is a great advantage.<br><br>It is no exaggeration to state that BTK inhibitors have literally been a lifeline for patients with WM who have received other therapies and progressed. |
| **Disadvantages of the technology** | |
| 10. What do patients or carers think are the disadvantages of the technology? | Any concerns expressed about the possible toxicity of BTKi are tempered by patient experience of generally excellent tolerability and physician experience of how to manage adverse effects by dose adjustment when necessary. Taking continuous indefinite therapy is a concern to a few but this becomes a limiting factor in practice.<br><br>**The following quotations have been taken word for word (and anonymised ) from the written responses to the questionnaire sent to patients enrolled on the ASPEN Study of zanubrutinib vs ibrutinib in WM**<br><br>**Patient S:** "I had an initial problem with zanubrutinib, it led to severe diarrhoea. But this was solved when Secondary ImmunoDeficiency was diagnosed and I started on a programme of IVIG. I now have no further problems taking zanubrutinib."<br><br>**Patient M:** "No disadvantages whatsoever."<br><br>**Patient R:** "I have had no challenging issues and the only adverse ongoing side effect is mild constipation and |

| | |
|---|---|
| | occasional infections."<br><br>**Patient B:** "Although there have been some clinical events that were challenging at the time, I have learnt about my body from the investigations carried out. I am not sure how many of these were related to taking zanubrutinib. There are some minor side effects but these are scarcely noticeable (e.g. petechiae). I continue on treatment."<br><br>**Patient A:** "There have been no disadvantages. I've had to travel to London every few weeks for the trial appointments, which is no real inconvenience. Other than that it has just been the usual needles for blood tests and CT scans, one excruciating biopsy at the start and a few unpleasant bone marrow biopsies. It's a tiny price to pay compared to the extreme unpleasantness of the other treatments I've had over the years." |
| **Patient population** | |
| 11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why. | We would expect zanabrutinib to benefit all patients with WM. However, there are certain patient groups who might be expected to derive particular benefit:<br><br>1. **Patients in need of rapid disease response**, such as hyperviscosity, which can be exacerbated by rituximab use, which is known to induce IgM flares and may prompt the need for invasive procedures such as plasma exchange. Whilst rituximab can be deferred in the chemoimmunotherapy setting, this removes a crucial ingredient when it is most keenly needed (early in the course of therapy).<br>2. **Patients with chemoresistant disease**: diminishing returns and accrual of toxicity from chemotherapy increases health needs and reduces QOL.<br>3. **Patients who are too frail for chemoimmunotherapy** are likely to benefit, as zanubrutinib works rapidly with a meaningful increase in blood counts and effective reduction in IgM levels/risk of hyperviscosity and less toxicity than ibrutinib (*Blood.2020 Oct 29;136(18):2038-2050*).<br>4. **Patients with wild-type MYD88 Waldenström's macroglobulinaemia** *Blood Advances 2020. In press*<br>5. **CNS involvement by WM– Bing-Neel syndrome** that requires and responds to a more targeted approach. The fact that BTKi, including zanubrutinib, cross the blood-brain barrier (*Hemasphere. 2018 Nov 30;2(6):e155*) is a huge advance as an alternative to targeting such disease is high dose methotrexate which is toxic and unrealistic for most WM patients. |

| | |
|---|---|
| | 6. WM-related immunologically driven conditions including but not restricted to those described below are likely to be especially responsive to zanubrutinib due to its mechanism of action:<br>• **Cryoglobulinaemia:** arises when the IgM paraprotein has the property of precipitating in vivo and causing organ and tissue damage (kidneys, joints, skin ulceration, nerve damage) via small vessel blockade and vasculitis.<br>• **Cold agglutinin disease:** a condition in which the presence of a WM disease clone can promote a cold-mediated haemolytic anaemia due to red cell agglutination and complement fixation. The size of the clone may not be high, but the immunological consequences can be life-changing for patients- Recurrent transfusions to replace the broken down red cells and the iron overload that ensues, as well as an increased rate of venous thromboembolism, can all be acutely life-threatening. In many cases, this disease's manifestation shows a limited response to chemoimmunotherapy directed at the underlying clonal disease. Frequently, patients remain transfusion-dependent and become progressively immunosuppressed by such treatment. There are preliminary data to suggest that BTKi are effective in this setting – Zanubrutinib would be a valuable asset in this context and could serve to obviate the need for transfusions and ongoing chemotherapy where chemoresistance has been noted.<br>• **WM mediated neuropathies** are a group of potentially disabling neuropathies due to direct infiltration of peripheral nerves or nerve roots or IgM-mediated activity against neural targets such as myelin associated glycoprotein that is found in nerve sheaths. This leads to a range of sensory and motor nerve damage with resultant progressive disability due to weakness and poor balance and increased risk of falls. There is no standard of care for these patients – chemoimmunotherapy is generally attempted with little success due to the front-loaded approach this offers as opposed to the effect that continuous targeted therapies offer, to enable stabilisation/return to functionality. |
| **Equality** | |
| 12. Are there any potential equality issues that should be taken into account when | It should be available for all patients irrespective of age and fitness, although it is expected to add especial benefit for those who are too frail for chemotherapy-based treatment. |

| | |
|---|---|
| considering this condition and the technology? | |
| **Other issues** | |
| 13. Are there any other issues that you would like the committee to consider? | BTK inhibitors demonstrate anti-inflammatory actions in a range of settings, including **severe COVID-19 infection** (*M. Roschewski et al., Sci. Immunol.10.1126/sciimmunol.abd0110 (2020*). This likely underpins the putative effectiveness of zanubrutinib in IgM-driven inflammatory disorders. In addition, it could be used in preference to chemoimmunotherapy in the era of COVID-19, which is likely to take several years to become endemic.<br>Given the current coronavirus pandemic, it is more important than ever to consider the potential benefits of well tolerated treatments that can be administered orally at home. |
| **Key messages** | |

14. In up to 5 bullet points, please summarise the key messages of your submission:

- Zanubrutinib offers a lifeline for WM patients at all stages of disease due to its targeted approach, and its inclusion in the treatment armamentarium for WM would be a game-changer for patients, extending life, improving QOL and reducing health needs.
- The feedback received from patients enrolled on the Phase 3 ASPEN study demonstrates very significant quality of life improvements for patients using zanubrutinib compared to existing chemotherapy based treatments available on the NHS. The resulting positive impact on carers is also significant.
- BTK inhibitors are highly active in WM and enable a far higher quality of life and contribution to society than existing chemotherapy-based treatments. This higher quality of life and the significantly fewer complications experienced by patients on BTK inhibitors reduces the burden on NHS services during and between treatment cycles.
- In the Phase 3 ASPEN study, zanubrutinib treatment was associated with a trend toward better response quality and less toxicity, particularly cardiovascular toxicity compared to Ibrutinib
- Continuous oral therapy, taken at home, is deliverable and acceptable to patients with WM and may offer greater opportunities for safer treatment in the current and post COVID-19 era.a

NICE National Institute for Health and Care Excellence

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

……

**Your privacy**

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……


in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

# Zanubrutinib for Waldenström's macroglobulinaemia [ID1427]

**Produced by** Kleijnen Systematic Reviews Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University

**Authors**
Rob Riemsma, Reviews Manager, Kleijnen Systematic Reviews Ltd, UK
Ben Wijnen, Health Economist, Maastricht UMC, the Netherlands
Sabine Grimm, Health Economist, Maastricht UMC
Sean Harrison, Systematic Reviewer and Statistician, KSR Ltd
Jeremy Howick, Systematic Reviewer, KSR Ltd
Nigel Armstrong, Health Economist, KSR Ltd
Willem Witlox, Health Economist, Maastricht UMC
Thomas Otten, Health Economist, Maastricht UMC
Mohammed Islam, Health Economist, KSR Ltd
Kate Misso, Information Specialist, KSR Ltd
Manuela Joore, Health Economist, Maastricht UMC
Jos Kleijnen, Director, KSR Ltd, Professor of Systematic Reviews in Health Care, Maastricht University

**Correspondence to**
Rob Riemsma, Kleijnen Systematic Reviews
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, United Kingdom
YO19 6FD

**Date completed** 02/06/2021

**Source of funding**: This report was commissioned by the NIHR Evidence Synthesis Programme as project number STA 13/35/29.

**Declared competing interests of the authors**

None.

**Acknowledgements**

None.

Commercial in confidence (CiC) data are highlighted in blue throughout the report.

Academic in confidence (AiC) data are highlighted in yellow throughout the report.

Confidential comparator prices are highlighted in green throughout the report.

Any de-personalised data are highlighted in pink throughout the report.



**Rider on responsibility for report**

The views 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:**

Riemsma R, Wijnen B, Grimm S, Harrison S, Howick J, Armstrong N, Witlox W, Otten T, Islam M, Misso K, Joore MA, Kleijnen J. Zanubrutinib for Waldenström's macroglobulinaemia [ID1427]: a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2021.

**Contributions of authors**

Rob Riemsma acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Ben Wijnen acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Sabine Grimm, Willem Witlox, Thomas Otten, Mohammed Islam and Nigel Armstrong acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Jeremy Howick acted as systematic reviewer, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Sean Harrison acted as systematic reviewers and statistician, critiqued the analyses in the company's submission and contributed to the writing of the report. Kate Misso critiqued the search methods in the submission and contributed to the writing of the report. Manuela Joore acted as health economist on this assessment, critiqued the company's economic evaluation, contributed to the writing of the report and provided general guidance. Jos Kleijnen critiqued the company's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

**Abbreviations**

| | |
|---|---|
| AE | Adverse events |
| AESI | Adverse event of special interest |
| AIC | Akaike Information Criterion |
| ALT | Alanine transaminase |
| ASCT | Autologous stem cell transplantation |
| AST | Aspartate aminotransferase |
| BCR | B-cell antigen receptor |
| BCSH | British Committee for Standards in Haematology |
| BDR | Bortezomib, rituximab and dexamethasone |
| BI | Budget impact |
| BIC | Bayesian information criteria |
| BID | Twice daily |
| BLNK | B cell linker |
| BNF | British National Formulary |
| BR | Rituximab and bendamustine |
| BSA | Body surface area |
| BSC | Best supportive care |
| BTK | Bruton's tyrosine kinase |
| Ca2+ | Calcium |
| CD19 | Cluster of differentiation 19 |
| CDF | Cancer Drug Fund |
| CE | Cost effectiveness |
| CE | Conformité Européenne (European Conformity) |
| CEA | Cost effectiveness analysis |
| CEAC | Cost effectiveness acceptability curve |
| CHMP | Committee for Medicinal Products for Human Use |
| CHOP | Cyclophosphamide, doxorubicin, vincristine and prednisone |
| CI | Confidence interval |
| Clad-R | Cladribine and rituximab |
| CLL | Chronic lymphocytic leukaemia |
| CMH | Cochran-Mantel-Haenszel |
| CR | Complete response |
| CRD | Centre for Reviews and Dissemination |
| CS | Company's submission |
| CSR | Clinical study report |
| CT | Computerised tomography |
| CVP | Cyclophosphamide, vincristine and prednisolone |
| CXCR4 | C-X-C motif chemokine receptor 4 |
| CYP | Cytochrome P450 |
| CYP3A | Cytochrome P4503A |
| DAG | 1,2 di-acyl glycerol |
| DAPS | Directly accessed pathology services |
| DCR | Dexamethasone, rituximab and cyclophosphamide |
| DOR | Duration of response |
| DSA | Deterministic sensitivity analysis |
| DSU | Decision Support Unit |
| ECOG PS | Eastern Cooperative Oncology Group Performance Status |
| EMA | European Medicines Agency |
| EORTC QLQ-C30 | European Organisation for Research and Treatment of Cancer |
| EPAR | European Public Assessment Report |
| EQ-5D-3L | European Quality of Life-5 Dimensions 3-Level |
| EQ-5D-5L | European Quality of Life-5 Dimensions 5-Level |
| ERG | Evidence Review Group |

| | |
|---|---|
| ESHAP | Etoposide, solu-medrone, cytarabine, cisplatin |
| ESMO | European Society for Medical Oncology |
| EUR | Erasmus University Rotterdam |
| FAD | Final appraisal document |
| FAS | Full analysis set |
| FCR | Fludarabine, cyclophosphamide and rituximab |
| FDA | Food and Drug Administration |
| FR | Fludarabine and rituximab |
| HR | Hazard ratio |
| HRQoL | Health-related quality of life |
| HRU | Healthcare resource use |
| HTA | Health technology assessment |
| IBR | Ibrutinib |
| IC | Indirect comparison |
| ICER | Incremental cost effectiveness ratios |
| IDARAM | Idarubicin, methotrexate, cytarabine and dexamethasone |
| Ig | Immunoglobulin |
| IKK | I kappa B kinase |
| Inv | Investigator |
| IPD | Individual patient-level data |
| IPSSWM | International Prognostic Scoring System for Waldenström's Macroglobulinaemia |
| IQR | Interquartile range |
| IRC | Independent review committee |
| ITC | Indirect treatment comparison |
| ITT | Intention-to-treat |
| IV | Intravenous |
| IWWM | International Workshop on Waldenström's Macroglobulinemia |
| IWWM-6 | Sixth International Workshop on Waldenström's Macroglobulinemia |
| IWWM-7 | Seventh International Workshop on Waldenström's Macroglobulinemia |
| KM | Kaplan-Meier |
| KSR | Kleijnen Systematic Reviews |
| LMM | Linear mixed effects model |
| LPL | Lymphoplasmacytic lymphoma |
| LS | Least square |
| LY | Life year |
| LYG | Life years gained |
| LYN | LYN proto-oncogene |
| MAA | Marketing authorisation application |
| Max | Maximum |
| MAIC | Matching adjusted indirect comparison |
| MCL | Mantle cell lymphoma |
| MeSH | Medical subject headings |
| MHRA | Medicines and Healthcare Products Regulatory Agency |
| Min | Minimum |
| MRR | Major response rate |
| MTA | Multiple technology appraisal |
| MTC | Mixed treatment comparison |
| MYD88 | Myeloid differentiation primary response gene 88 |
| *MYD88*$^{MUT}$ | Myeloid differentiation primary response gene 88 mutant |
| *MYD88*$^{WT}$ | Wild-type myeloid differentiation primary response gene 88 |
| n | Number of patients in the category |
| N | Number of patients evaluable |
| N/A | Not applicable |
| NE | Not evaluable |
| neff | Effective sample size |

| | |
|---|---|
| NFκB | Nuclear factor kappa B |
| NFAT | Nuclear factor of activated T cells |
| NHL | Non-Hodgkin's lymphoma |
| NHS | National Health Service |
| NICE | National Institute of Health and Care Excellence |
| NIHR | National Institute for Health Research |
| NMA | Network meta-analysis |
| No. | Number |
| NR | Not reported |
| NYHA | New York Heart Association |
| OD | Once daily |
| OR | Overall response |
| ORR | Objective response rate |
| OS | Overall survival |
| PAS | Patient access scheme |
| PD | Progressive disease |
| PFS | Progression-free survival |
| PH | Proportional hazard |
| Pl3K | Phosphatidylinositol-4,5-bisphosphate 3-kinase |
| PIP3 | Phosphatidylinositol (3,4,5)-trisphosphate |
| PKC | Protein kinase C |
| PLC | Phospholipase C |
| PN | Peripheral neuropathy |
| PP | Per-protocol |
| PPS | Post-progression survival |
| PR | Partial response |
| PRESS | Peer Review of Electronic Search Strategies |
| PRISMA | Preferred reporting items for systematic reviews and meta-analyses |
| PRO | Patient-reported outcome |
| PSM | Partitioned survival model |
| PSA | Probabilistic sensitivity analysis |
| PSS | Personal Social Services |
| PSSRU | Personal Social Services Research Unit |
| Pt | Patient |
| PT | Preferred Term |
| QALY | Quality adjusted life year |
| QLQ-C30 | Quality of Life Questionnaire core-30 |
| QoL | Quality of life |
| QTcF | T interval corrected for heart rate using Fridericia's formula |
| R | Rituximab or Randomised |
| R-CHOP | Rituximab, cyclophosphamide, doxorubicin, vincristine and prednisone |
| R-ESHAP | Rituximab, etoposide, solu-medrone, cytarabine and cisplatin |
| R-IDARAM | Rituximab, idarubicin, methotrexate, cytarabine and dexamethasone |
| R/R | Relapsed/refractory |
| RAP | RapGTP-binding protein |
| RCT | Randomised controlled trial |
| RR | Relative risk; Risk ratio |
| SAE | Serious adverse events |
| ScHARR | School of Health and Related Research |
| SD | Standard deviation |
| SE | Standard error |
| SLR | Systematic literature review |
| SMC | Scottish Medicines Consortium |
| SmPC | Summary of product characteristics |
| SMQ | Standardized Medical Dictionary for Regulatory Activities Query |

| | |
|---|---|
| SOC | System Organ Class |
| SPEP | Serum protein electrophoresis |
| STA | Single technology appraisal |
| STM | Sate-transition model |
| SYK | Spleen tyrosine kinase |
| TA | Technology assessment/Technology appraisal |
| TN | Treatment naïve |
| TRAE | Treatment-related adverse event |
| TTD | Time to discontinuation |
| TTO | Time trade-off |
| UK | United Kingdom |
| ULN | Upper limit of normal |
| UMC | University Medical Centre |
| Unk | Unknown |
| USA | United States of America |
| VGPR | Very good partial response rate |
| VR | Bortezomib and rituximab |
| vs | Versus |
| WHIM | Warts, hypogammaglobulinemia, infections, myelokathexis |
| WHO | World Health Organization |
| WM | Waldenström's macroglobulinaemia |
| WMUK | Waldenström's Macroglobulinaemia United Kingdom |
| WT | Wild type |
| WTP | Willingness-to-pay |
| ZANU | Zanubrutinib |

# Table of Contents

**Table of Tables**

**Table of Figures**

# 1. EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the evidence review group (ERG) as being potentially important for decision making. If possible, it also includes the ERG's preferred assumptions and the resulting incremental cost effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 presents the key model outcomes. Section 1.3 discusses the decision problem, Section 1.4 relates to the clinical effectiveness, and Section 1.5 relates to the cost effectiveness. Other key issues are discussed in Section 1.6 while a summary in presented in Section 1.7.

Background information on the condition, technology and evidence and information on key as well as non-key issues are in the main ERG report, see Sections 2 (decision problem), 3 (clinical effectiveness) and 4 to 6 (cost effectiveness) for more details.

All issues identified represent the ERG's view, not the opinion of the National Institute for Health and Care Excellence (NICE).

## *1.1 Overview of the ERG's key issues*

**Table 1.1: Summary of key issues**

| ID1427 | Summary of issue | Report sections |
|---|---|---|
| 1 | The comparators are not in line with the NICE scope | Sections 2.3, 3.3 and 3.4 |
| 2 | Patients with cardiovascular disease and those taking warfarin were excluded from the ASPEN trial | Section 2.1 |
| 3 | The evidence for treatment naïve patients is based on small numbers of patients and has limited generalisability | Section 3.2.5 |
| 4 | Survival data for zanubrutinib are immature | Section 3.2.5 |
| 5 | The indirect comparisons with rituximab and bendamustine (BR) and dexamethasone, rituximab and cyclophosphamide (DCR) are unreliable | Sections 3.3 and 3.4 |
| 6 | The choice of a partitioned survival model and its underlying assumptions | Section 4.2.2 |
| 7 | The model does not include all comparators mentioned in the NICE scope | Section 4.2.4 |
| 8 | Ibrutinib should not be included as direct comparator and/or subsequent treatment option in the economic model | Section 4.2.4 |
| 9 | The partitioned survival analysis chosen by the company relies on estimates for progression-free survival (PFS) and overall survival (OS), secondary and exploratory endpoints respectively. Only a small number of PFS and OS events had occurred at the time of this appraisal. | Section 4.2.6 |
| 10 | Plausibility of OS hazards falling below background mortality hazards. | Section 4.2.6 |
| 11 | The use of data from patients with *MYD88*$^{MUT}$ only. | Section 4.2.6 |
| 12 | Assumption of lifelong treatment effectiveness. | Section 4.2.6 |

| ID1427 | Summary of issue | Report sections |
|---|---|---|
| 13 | PFS utility higher than general UK population values. | Section 4.2.8 |
| 14 | The value and standard error implemented for post-progression utility is not evidence-based. | Section 4.2.8 |
| 15 | Large discrepancy between the deterministic incremental cost effectiveness ratio (ICER) and the probabilistic ICER. | Section 5.3.4 |

The key differences between the company's preferred assumptions and the ERG's preferred assumptions are 1) the exclusion of ibrutinib as direct comparator and as subsequent treatment and 2) treatment waning at a five-year cut-off as adopted in prior appraisals. In addition, the ERG preformed exploratory scenario analyses to explore the impact of alternative survival curves.

### *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.

Overall, the technology is modelled to affect QALYs by:

- Increasing overall survival (OS).
- Increasing progression-free survival.

Overall, the technology is modelled to affect costs by:

- Higher unit price than current treatments.

The modelling assumptions that have the greatest effect on the ICER are:

- Assumptions regarding subsequent treatments to BR and DRC in the economic model.
- Because of background mortality over-riding the OS distributions, it is likely that the driving factor in the model is short-to-medium term OS and the timepoint background mortality takes over in the zanubrutinib arm, rather than long-term extrapolation.
- Assumption of the timepoint at which treatment waning is assumed to start.

### *1.3 The decision problem: summary of the ERG's key issues*

The decision problem addressed in the company submission (CS) is partially in line with the final scope issued by NICE. However, some comparators mentioned in the NICE scope have not been included by the company (Table 1.2).

**Table 1.2: Key issue 1: The comparators are not in line with the NICE scope.**

| Report section | Sections 2.3, 3.3 and 3.4 |
|---|---|
| **Description of issue and why the ERG has identified it as important** | The comparators are not in line with the NICE scope. Fludarabine and rituximab (FR), fludarabine, cyclophosphamide and rituximab (FCR), and cladribine and rituximab (Clad-R) have not been included as comparators due to lack of data according to the company. Autologous stem cell transplantation (ASCT) has not been included in any of the literature searches reported in the company submission (CS). Ibrutinib has been |

| Report section | Sections 2.3, 3.3 and 3.4 |
|---|---|
| | included as a comparator. However, NICE explicitly excluded ibrutinib as a comparator. |
| **What alternative approach has the ERG suggested?** | In the response to clarification the company produced matching adjusted indirect comparison (MAICs) comparing zanubrutinib with FR and FCR. The ERG believes that the indirect comparisons between zanubrutinib and FR, or FCR, are just as valid as the comparisons with BR and DCR and should therefore have been included in the CS. The ERG agrees with the company that a comparison of zanubrutinib with Clad-R is not feasible. ASCT should also have been included as a comparator. |
| **What is the expected effect on the cost effectiveness estimates?** | The effect on the cost effectiveness estimates is unclear as these comparators (FC, FCR and ASCT) have not been included in the company's economic analyses. |
| **What additional evidence or analyses might help to resolve this key issue?** | The company should include comparisons between zanubrutinib and FC, FCR and ASCT in the economic model. The relevance of these comparisons could be informed by clinical expert opinion. |

### *1.4 The clinical effectiveness evidence: summary of the ERG's key issues*

The ERG identified three major concerns with the evidence presented on the clinical effectiveness, namely that patients with cardiovascular disease and those taking warfarin were excluded from the ASPEN trial (Table 1.3); survival data for zanubrutinib are immature (Table 1.4); and the indirect comparisons with BR and DCR are unreliable (Table 1.5).

**Table 1.3: Key issue 2: Patients with cardiovascular disease and those taking warfarin were excluded from the ASPEN trial**

| Report section | Section 2.1 |
|---|---|
| **Description of issue and why the ERG has identified it as important** | Patients with cardiovascular disease and those taking warfarin were excluded from the ASPEN trial.<br>The company did not want to expose patients with underlying comorbidities to known or unknown side-effects; therefore, patients with cardiovascular disease were excluded from the ASPEN trial. This means possible cardiac serious adverse events (AEs) may not have been observed due to the inclusion criteria of the ASPEN trial. |
| **What alternative approach has the ERG suggested?** | The ERG has no suggestions for an alternative approach. |
| **What is the expected effect on the cost effectiveness estimates?** | The effect on the cost effectiveness estimates is unclear. |
| **What additional evidence or analyses might help to resolve this key issue?** | The ERG has no suggestions for additional evidence or analyses. Therefore, this might imply a constraint on the population in the scope i.e. to exclude those with cardiovascular disease and those taking warfarin. This issue might be resolved if clinical expert opinion indicated that such patients were not eligible for treatment. |

**Table 1.4: Key issue 3: The evidence for treatment naïve patients is based on small numbers of patients and has limited generalisability**

| Report section | Section 3.2.5 |
|---|---|
| **Description of issue and why the ERG has identified it as important** | In the ASPEN trial, a total of 201 patients were randomised to zanubrutinib or ibrutinib; 164 patients had relapsed/refractory disease (zanubrutinib, n=83 versus ibrutinib, n=81) and 37 were treatment naïve and unsuitable for chemotherapy (zanubrutinib, n=19 versus ibrutinib, n=18). Therefore, the evidence for treatment naïve patients is based on small numbers of patients and is limited to patients who were unsuitable for chemotherapy. |
| **What alternative approach has the ERG suggested?** | The ERG has no suggestions for an alternative approach. |
| **What is the expected effect on the cost effectiveness estimates?** | The effect on the cost effectiveness estimates is unclear. However, the results of the economic analyses will be less reliable for this population. |
| **What additional evidence or analyses might help to resolve this key issue?** | The ERG has no suggestions for additional analyses. |

**Table 1.5: Key issue 4: Survival data for zanubrutinib are immature**

| Report section | Section 3.2.5 |
|---|---|
| **Description of issue and why the ERG has identified it as important** | Survival data for zanubrutinib are immature. Also, only PFS and OS were considered as outcomes in the MAIC. |
| **What alternative approach has the ERG suggested?** | Longer term follow-up from the ASPEN trial is necessary to resolve this issue. The company did provide updated results as part of the response to clarification (cut-off date of 31 August 2020). However, PFS in the updated results was based on investigator assessment rather than independent review committee (IRC) assessed as in the CS; and OS and PFS were still immature in the updated results. |
| **What is the expected effect on the cost effectiveness estimates?** | The effect on the cost effectiveness estimates is unclear. |
| **What additional evidence or analyses might help to resolve this key issue?** | Longer term follow-up from the ASPEN trial is necessary to resolve this issue. |

**Table 1.6: Key issue 5: The indirect comparisons with BR and DCR are unreliable**

| Report section | Sections 3.3 and 3.4 |
|---|---|
| **Description of issue and why the ERG has identified it as important** | The indirect comparisons (MAICs) with BR and DCR are unreliable for the following reasons:<br>- Only PFS, OS, and AEs were considered as outcomes in the MAIC<br>- These survival data for zanubrutinib are immature.<br>- There is a substantial risk of bias. The company submission listed a range of baseline patient variables considered to be potential prognostic factors or effect modifiers and would therefore likely cause bias in a MAIC if the included studies had differences in these variables. As no study presented the requisite |

| Report section | Sections 3.3 and 3.4 |
|---|---|
| | summary data to match on all variables, no MAIC matched on all these variables.<br>- In addition to the potential prognostic factors or effect modifiers listed in the company submission, other variables are also to cause bias and were not matched for in the MAICs, including socio-economic status, year of study, location of study, general health of patients.<br>- Additionally, the definitions of outcomes were not always consistent between studies, and the interventions were administered differently in each study.<br>- Finally, it is unclear to what extent the MAICs are relevant to a contemporary National Health Service (NHS) population, given differences in baseline variables between the studies in the MAICs (to which the patients in ASPEN were matched) and the patients with Waldenström's macroglobulinaemia (WM) in the UK clinical practice. |
| **What alternative approach has the ERG suggested?** | The ERG has no suggestions for an alternative approach.<br>Given the lack of direct evidence, indirect methods were used, and given the lack of randomised controlled trial (RCT) evidence to inform a network meta-analysis, MAICs using single-arm studies represent the only available evidence to adjust for confounding comparing zanubrutinib with comparator treatments. Even so, these analyses still present a substantial risk of bias. |
| **What is the expected effect on the cost effectiveness estimates?** | The effect on the cost effectiveness estimates is unclear. However, the results of the economic analyses will be less reliable due to these potential biases. |
| **What additional evidence or analyses might help to resolve this key issue?** | The ERG has no suggestions for additional analyses. |

### *1.5 The cost effectiveness evidence: summary of the ERG's key issues*

The main sources of evidence on treatment effectiveness (i.e. OS and PFS) used for intervention and comparators used in the model are two matched MAICs comparing zanubrutinib with BR and DRC based on the ASPEN trial and Tedeschi et al 2015 and Dimopoulos et al. 2007/Kastritis et al. 2015. Main issue in this submission is the immaturity of the available data, as was acknowledged by the company. The partitioned survival analysis chosen by the company relies on estimates for PFS and OS, secondary and exploratory endpoints respectively. Only a small number of PFS and OS events had occurred at the time of this appraisal and many patients were censored. The long-term predictions are therefore extremely uncertain. In relation to this, it should also be highlighted that extreme differences in survival curves do not necessary lead to an extreme impact on the median and mean survival estimates, and on the ICER, as hazards of all survival models fall below background mortality hazards after a certain period of time and background mortality is then assumed to apply in the model. Because of background mortality over-riding the OS distributions for all comparators, it is likely that the driving factor in the model is short-to-medium term OS gain, and the timepoint background mortality takes over in the zanubrutinib arm, rather than long-term OS extrapolation. These timepoints differ for each distribution in each comparison but range between seven to 10 years and two to seven years with jointly-fitted models and independently fitted models respectively (for zanubrutinib) for the DRC comparison; and five to 12 years and baseline-19 years with jointly-fitted and independent models respectively for the BR

comparison. The company could not provide any evidence to support mortality hazards for all modelled treatments dropping below general population mortality hazard. Another issue is that there appears to be a structural and large discrepancy between the deterministic and probabilistic analysis which questions the internal validity of the modelled results (i.e. difference between the deterministic ICER and the probabilistic ICER for BR of [redacted] and a difference of [redacted] for DRC). These differences were even larger in the ERG base-case analysis/scenarios. Consequently, the cost effectiveness results, calculated using the economic model submitted by the company, are subject to a large degree of uncertainty.

In the company base-case (probabilistic), the ICER amounted to [redacted] per QALY when compared to BR and [redacted] per QALY when compared to DRC. The individual ERG adjustments had large impact on the ICER, ranging from [redacted] per QALY gained to [redacted] per QALY gained compared to BR and from [redacted] per QALY gained to [redacted] per QALY gained compared to DRC. The estimated ERG base-case ICER (probabilistic) was [redacted] per QALY gained for zanubrutinib compared to BR and [redacted] per QALY gained for zanubrutinib compared to DRC.

A full summary of the cost effectiveness evidence review conclusions can be found in Section 7.4 of this report. The company's cost effectiveness results are presented in Section 6, the ERG's summary and detailed critique in Section 5, and the ERG's amendments to the company's model and results are presented in Section 7. The main ERG results are reproduced using confidential patient access schemes (i.e. for cannabidiol) in a confidential appendix. The key issues in the cost effectiveness evidence are discussed in Tables 1.7 to 1.16.

**Table 1.7: Issue 6: The choice of a partitioned survival model and its underlying assumptions.**

| **Report section** | **Section 4.2.2** |
|---|---|
| **Description of issue and why the ERG has identified it as important** | The choice of a partitioned survival model (PSM) and its underlying assumptions.<br>PSMs are often used in oncology. However, the progression-free survival and overall survival of patients are relatively long and as a result health-related quality of life (HRQoL) (except for age-related utility decrease) and cost and resource use are stable over a relatively long period. Next to that, the model's health state occupancy was based secondary (OS) and exploratory (PFS) outcomes in the ASPEN trial. |
| **What alternative approach has the ERG suggested?** | The ERG suggested to explore a state-transition model (STM) to validate outcomes of the current model but acknowledges that the applicability of a STM is questionable based on the data availability of the comparators specified in NICE's final scope. |
| **What is the expected effect on the cost effectiveness estimates?** | Unknown. |
| **What additional evidence or analyses might help to resolve this key issue?** | Further justification from the company regarding the plausibility of the HRQoL and cost and resource use being stable over a relatively long period and the potential implications of health state occupancy being based on secondary and exploratory endpoints. |

**Table 1.8: Issue 7: The model does not include all comparators mentioned in the NICE scope**

| **Report section** | **Section 4.2.4** |
|---|---|
| **Description of issue and why the ERG has identified it as important** | The model does not include all comparators mentioned in the NICE scope.<br>In contrast to the NICE scope, the model does not include FR, FCR, Clad-R and ASCT (for patients who have had at least one prior |

| Report section | Section 4.2.4 |
|---|---|
| | therapy), chlorambucil, rituximab monotherapy and best supportive care (BSC) (for patients for whom chemo-immunotherapy is unsuitable) as comparators. |
| **What alternative approach has the ERG suggested?** | Based on the company's response to question A26, the ERG argues that the company could have done exploratory analyses for the comparisons of zanubrutinib with FR/FCR and rituximab monotherapy. |
| **What is the expected effect on the cost effectiveness estimates?** | Unknown. |
| **What additional evidence or analyses might help to resolve this key issue?** | Exploratory analyses mentioned in the company's response to clarification question A26 could have been included in the model. |

**Table 1.9: Issue 8: Ibrutinib should not be included as direct comparator and/or subsequent treatment option in the economic model**

| Report section | Section 4.2.4 |
|---|---|
| **Description of issue and why the ERG has identified it as important** | According to NICE's position statement on treatments currently in Cancer Drug Fund (CDF), ibrutinib should not be included as direct comparator and/or subsequent treatment option in the economic model. |
| **What alternative approach has the ERG suggested?** | The ERG ignored the evidence for the comparison of zanubrutinib with ibrutinib and excluded the possibility of ibrutinib as a subsequent treatment in the ERG base-case analysis. Instead, the ERG assumed in its base-case that patients initially treated with BR would receive DRC as subsequent treatment and patient initially treated with DRC would receive BR. Patients initially receiving zanubrutinib received subsequent treatment according to the CS base-case (BR for 60.4% of the patients and DRC for 39.6% of the patients). |
| **What is the expected effect on the cost effectiveness estimates?** | Unknown. |
| **What additional evidence or analyses might help to resolve this key issue?** | The company should provide additional evidence (e.g. expert opinion or clinical trials) that gives insight into possible subsequent treatments in absence of ibrutinib in the UK. |

**Table 1.10: Issue 9: Only a small number of PFS and OS events had occurred at the time of this appraisal.**

| Report section | Section 4.2.6 |
|---|---|
| **Description of issue and why the ERG has identified it as important** | Only a small number of PFS and OS events had occurred at the time of this appraisal.<br>The partitioned survival analysis chosen by the company relies on estimates for PFS and OS, secondary and exploratory endpoints respectively. Only a small number of PFS and OS events had occurred at the time of this appraisal. |
| **What alternative approach has the ERG suggested?** | The ERG was not able to resolve the uncertainty caused by data immaturity. |

| | |
|---|---|
| **What is the expected effect on the cost effectiveness estimates?** | Unknown. |
| **What additional evidence or analyses might help to resolve this key issue?** | Collection of long-term follow-up data. |

**Table 1.11: Issue 10: Plausibility of OS hazards falling below background mortality hazards.**

| **Report section** | **Section 4.2.6** |
|---|---|
| **Description of issue and why the ERG has identified it as important** | Plausibility of OS hazards falling below background mortality hazards.<br>Hazards of all survival models fall below background mortality hazards and background mortality is then assumed to apply. Because of background mortality overriding the OS distributions, it is likely that the driving factor in the model is short-to-medium term OS and the timepoint background mortality takes over in the zanubrutinib arm, rather than long-term extrapolation. |
| **What alternative approach has the ERG suggested?** | The company was asked to provide any evidence to support mortality hazards dropping below general population mortality hazards. No expert opinion was provided on this in particular (only that experts expected monotonically increasing hazards). Upon request, the company assessed the impact of summing up model hazards and background mortality hazards in scenario analysis: this increased the ICER substantially in both BR and DRC and illustrates that uncertainty about long-term hazards could be a model driver. |
| **What is the expected effect on the cost effectiveness estimates?** | Unknown, but a scenario by the company suggests that alternative assumptions about long-term OS could be impactful |
| **What additional evidence or analyses might help to resolve this key issue?** | An explanation as to whether this is simply an artifact of data immaturity, or whether low mortality hazards in the long run indicate that there is a subgroup of patients with WM that are at particular risk of dying in the first years into the modelled disease trajectory, whilst the average patient has closer to normal life expectancy, or alternative explanations. |

**Table 1.12: Issue 11: The use of data from patients with *MYD88*$^{MUT}$ only.**

| **Report section** | **Section 4.2.6** |
|---|---|
| **Description of issue and why the ERG has identified it as important** | The use of data from patients with *MYD88*$^{MUT}$ only.<br>The company used the ITT population for their analyses, which only contained patients with *MYD88*$^{MUT)}$ (L265P point mutation in myeloid differentiation primary response gene 88). In addition, the final scope issued by NICE does not specify any genetic marker and refers to people with the *MYD88*$^{MUT}$ as a relevant a subgroup of the population only. |
| **What alternative approach has the ERG suggested?** | The ERG was concerned that results of cost effectiveness analyses based on cohort 1 only might not be generalisable and requested a scenario with a pooled analysis of both cohorts. The company provided this in response to question B6 by performing new MAICs, updating HRQoL inputs and performing cost effectiveness analyses with these inputs. Cost effectiveness results of pooled cohort 1 and 2 were relatively close to results of cohort 1, if slightly lower for the pooled analysis. |

| Report section | Section 4.2.6 |
|---|---|
| | The company did not state whether the analysis was weighted to reflect the mix of patients in clinical practice (i.e. 90% of *MYD88*$^{MUT}$ and 5-10% of *MYD88*$^{WT}$) and hence the ERG assumes that this weighting did not occur and the weight was instead determined by patient numbers in the cohorts. |
| **What is the expected effect on the cost effectiveness estimates?** | Unknown. |
| **What additional evidence or analyses might help to resolve this key issue?** | A weighted analysis to reflect the mix of patients in clinical practice (i.e. 90% of *MYD88*$^{MUT}$ and 5-10% of *MYD88*$^{WT}$). In addition, this would require information regarding the mix of mutations in the comparator arm. |

**Table 1.13: Issue 12: Assumption of lifelong treatment effectiveness.**

| Report section | Section 4.2.6 |
|---|---|
| **Description of issue and why the ERG has identified it as important** | Assumption of lifelong treatment effectiveness.<br>The ERG considered that the assumption of lifelong treatment effectiveness may not be justified and requested that the company implement treatment effectiveness waning in the model. The company implemented treatment waning for both PFS and OS at different time points (i.e. 30 months, five years, seven years, and 10 years), using hazard ratios of 1 from the chosen time point onwards. Results of these scenarios showed that this was influential. |
| **What alternative approach has the ERG suggested?** | The ERG, in line with previous related appraisals (e.g. TA627), adopted treatment waning at five years in its base-case, but acknowledges that this is uncertain. |
| **What is the expected effect on the cost effectiveness estimates?** | The inclusion of treatment waning increases the ICER. In the CS scenario analysis, when assuming treatment waning after five years, the ICERs increased to [redacted] per QALY gained for zanubrutinib vs BR and to [redacted] per QALY gained compared to DRC. |
| **What additional evidence or analyses might help to resolve this key issue?** | Long-term follow-up data regarding treatment waning over time. |

**Table 1.14: Issue 13: PFS utility higher than general UK population values.**

| Report section | Section 4.2.8 |
|---|---|
| **Description of issue and why the ERG has identified it as important** | PFS utility higher than general UK population values. According to a recent publication the health-related quality of life in people aged 65 or over was valued lower than the health-related quality of the utility value which was implemented in this model. Upon a clarification request the company explained that this could be due to 1) natural differences due to differences in trial and real-world setting and 2) differences between the geographic location of participants of the ASPEN trial and UK citizens. This potential lack of transferability calls into question the validity of using the utility estimates from the ASPEN trial. |
| **What alternative approach has the ERG suggested?** | Provide evidence on the justification of the PFS utility values in the model. An age-adjustment was implemented by the company however it remains uncertain whether this is a sufficient adjustment. |

| Report section | Section 4.2.8 |
|---|---|
| **What is the expected effect on the cost effectiveness estimates?** | A decrease in the utility difference is likely to increase the ICER. |
| **What additional evidence or analyses might help to resolve this key issue?** | Additional evidence about the health-related quality of life of WM patients in the UK may help to resolve this issue. |

**Table 1.15 Issue 14: The value and standard error implemented for post-progression utility is not evidence-based.**

| Report section | Section 4.2.8 |
|---|---|
| **Description of issue and why the ERG has identified it as important** | The value and standard error implemented for post-progression utility is not evidence-based.<br>Due to a lack of data about post-progression utility, the company implemented a utility decrement of -.1 to attain post-progression quality of life in reference to the company submissions in TA491 and TA502. However, in both appraisals the utility decrement was increased to around -.18. To keep this submission in line with previous appraisals this decrement of -.18 was implemented in the ERG base-case. However considerable uncertainty remains as to the real post-progression health-related quality of life. |
| **What alternative approach has the ERG suggested?** | The ERG implemented a utility decrement of -.18 to stay in line with previous technical appraisals. Furthermore, an increased standard error around the post-progression utility could be implemented to reflect the uncertainty around the post-progression utility in the probabilistic analysis (note: the ERG did not adjust the standard error). |
| **What is the expected effect on the cost effectiveness estimates?** | The increase in the decrement, decreased the ICER of zanubrutinib compared to BR and DRC in the ERG base-case. A increase in standard error around the post-progression utility is likely to slightly increase the overall uncertainty in the probabilistic analysis. |
| **What additional evidence or analyses might help to resolve this key issue?** | Additional evidence about the post-progression health-related quality of life of WM patients may help to resolve this issue. |

**Table 1.16: Issue 15: Large discrepancy between the deterministic ICER and the probabilistic ICER.**

| Report section | Section 5.3.4 |
|---|---|
| **Description of issue and why the ERG has identified it as important** | The large discrepancy between the deterministic ICER and the probabilistic ICER.<br>The ERG noted a large difference between the deterministic ICER and the probabilistic ICER for BR (difference of [redacted]) and DRC (difference of [redacted]). There appears to be a structural difference between the deterministic and probabilistic analysis with the latter have structurally higher ICERs. |
| **What alternative approach has the ERG suggested?** | The company should provide an explanation why both results differ. It could be either distributions are mis-specified or correlations ignored. |
| **What is the expected effect on the cost effectiveness estimates?** | Unknown. |

| Report section | Section 5.3.4 |
|---|---|
| **What additional evidence or analyses might help to resolve this key issue?** | The company should make sure all relevant input parameters are adequately modelled in the probabilistic analysis and demonstrate a minimal difference between both the deterministic and probabilistic results or provide an explanation as to why these differences arise. |

**Table 1.17: Issue 16: Treatment effectiveness being analysed for the different comparisons separately**

| Report section | Sections 4.2.6 and 5.1 |
|---|---|
| **Description of issue and why the ERG has identified it as important** | Treatment effectiveness being analysed for the different comparisons separately and hence a fully incremental analysis was not performed.<br>The ERG is concerned that treatment effectiveness was analysed for the different comparisons separately. This meant that no fully incremental cost effectiveness analyses were performed, due to differences in the populations between the BR and DRC MAIC in several characteristics. Nevertheless, the use of different analyses is problematic as it does not allow for comparison of zanubrutinib, DRC and BR. This warrants the questions to what extent the two ICERs (i.e. zanubrutinib vs BR and zanubrutinib vs DRC) are applicable to the same population. |
| **What alternative approach has the ERG suggested?** | The ERG questions whether a fully incremental analysis could be performed with these comparisons. |
| **What is the expected effect on the cost effectiveness estimates?** | Unknown. |
| **What additional evidence or analyses might help to resolve this key issue?** | If the company could gain access to the individual patient-level data (IPD) of both the BR and DRC clinical study data they could be pooled with the zanubrutinib trial data in order to compare all three treatments in effectively the same population. However, given the lack of overlap in variables included in the two MAICs, such an analysis would still be subject to high risk of bias given the unobserved variables. |

### *1.6 Summary of the ERG's view*

Adjustments made by the ERG, to derive the ERG base-case (using the CS base-case as starting point) are listed below. The 'fixing error' adjustments were combined. All other ERG analyses were performed incorporating the 'fixing error' adjustments given the ERG considered that the 'fixing error' adjustments corrected unequivocally wrong issues.

**Fixing errors**

1. The model of the company submission applied DRC for five and BR for seven model cycles. BR was supposed to be given for six treatment cycles, which was implemented in the ERG base-case. Given that it was unclear to the ERG for how many cycles DRC was supposed to be given, the ERG chose to implement BR for six treatment cycles with the duration of three weeks in the ERG base-case (Section 4.2.9).

**Fixing violation**

2. Ibrutinib was excluded from the model as direct comparator and as subsequent treatment (Section 4.2.4).

**Matters of judgement**

3. Assuming similar relative dose intensity rates of 97.5% for BR, DRC, and zanubrutinib instead of 100% for both.(Section 4.2.4).
4. In the CS base-case no treatment waning was assumed. Although the ERG acknowledges the difficulties in empirically assessing treatment-waning, a five-year cut-off, as adopted in prior related appraisals (e.g. TA627) was assumed (Section 4.2.6).
5. Inclusion of all AEs of Grade ≥3 which occurred in ≥1% of the population, instead of ≥5% of the trial populations (Section 4.2.7).
6. The use of age-adjusted utility values instead of (Section 4.2.8).
7. A utility decrement of 0.18 in line with TA491 and TA502 instead of a utility decrement of 0.1 (Section 4.2.8).

### *1.7 ERG exploratory scenario analyses*

The ERG performed the following exploratory scenario analyses to assess the impact of alternative assumptions conditional on the ERG base-case.

1. OS scenarios: DRC comparison: use dependent exponential for OS (to be in line with PFS), BR comparison: dependent gamma for OS
2. PFS scenarios: DRC comparison: dependent Gompertz for PFS, BR comparison: dependent lognormal for PFS
3. A scenario assuming no subsequent treatments.

**Table 1.18: Summary of ERG's preferred assumptions and ICER**

| Technologies | Total costs (£) | Total QALYs | Incremental costs (£) | Incremental QALYs | ICER (£/QALY) |
|---|---|---|---|---|---|
| **ERG base-case (ERG_1 -ERG_7)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £53,685 | 4.60 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £50,562 | 5.40 | | | |
| **Company's corrected base-case (ERG_1 & ERG_2)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £53,842 | 4.96 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £50,695 | 5.88 | | | |
| **Matter of judgement: Similar dose-intensities (ERG_1, ERG_2 & ERG_3)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £53,685 | 4.96 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |

| Technologies | Total costs (£) | Total QALYs | Incremental costs (£) | Incremental QALYs | ICER (£/QALY) |
|---|---|---|---|---|---|
| DRC | £50,552 | 5.88 | | | |
| **Matter of judgement: Treatment waning (ERG 1, ERG_2 & ERG_4)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £53,842 | 4.96 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £50,695 | 5.88 | | | |
| **Matter of judgement: Including additional AEs (ERG_1, ERG_2 & ERG_5)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £53,842 | 4.96 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £50,705 | 5.88 | | | |
| **Matter of judgement: Age-adjusted utilities (ERG_1, ERG_2 & ERG_6)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £53,842 | 4.77 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £50,695 | 5.62 | | | |
| **Matter of judgement: Post-progression utility decrement (ERG_1, ERG_2 & ERG_7)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £53,842 | 4.78 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £50,695 | 5.64 | | | |

**Table 1.19: Probabilistic ERG base-case**

| Technologies | Total costs (£) | Total QALYs | Incremental costs (£) | Incremental QALYs | ICER (£/QALY) |
|---|---|---|---|---|---|
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £53,658 | 4.51 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £50,626 | 5.38 | | | |

**Table 1.20: Deterministic scenario analyses (conditional on ERG base-case)**

| Technologies | Total costs (£) | Total QALYs | Incremental costs (£) | Incremental QALYs | ICER (£/QALY) |
|---|---|---|---|---|---|
| **Scenario: Treatment waning 10 years (ERG_1 - ERG_7, ERG_12)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £53,685 | 4.60 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £50,562 | 5.40 | | | |
| **Scenario: OS - Dependent exponential DRC (ERG_1 - ERG_7, ERG_13)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £53,685 | 4.60 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £50,561 | 5.44 | | | |
| **Scenario: OS - Dependent gamma BR (ERG_1 - ERG_7, ERG_14)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £53,563 | 4.23 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £50,562 | 5.40 | | | |
| **Scenario: PFS - Dependent Gompertz DRC (ERG_1 - ERG_7, ERG_15)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £53,685 | 4.60 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £49,908 | 5.48 | | | |

| Technologies | Total costs (£) | Total QALYs | Incremental costs (£) | Incremental QALYs | ICER (£/QALY) |
|---|---|---|---|---|---|
| **Scenario: PFS - Dependent lognormal BR (ERG_1 - ERG_7, ERG_16)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £52,651 | 4.82 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £50,562 | 5.40 | | | |
| **Scenario: No subsequent treatment (ERG_1 - ERG_7, ERG_17)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £32,039 | 4.60 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £24,859 | 5.40 | | | |

## 2. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

**Table 2.1: Statement of the decision problem (as presented by the company)**

| | Final scope issued by NICE | Decision problem addressed in the company submission | Rationale if different from the final NICE scope | ERG comment |
|---|---|---|---|---|
| **Population** | Adults with WM:<br>• who have had at least 1 prior therapy, or<br>• whose disease is untreated, for whom chemo-immunotherapy is unsuitable | As per scope | N/A | The population is in line with the scope. |
| **Intervention** | Zanubrutinib | As per scope | N/A | The intervention is in line with the NICE scope |
| **Comparator(s)** | Treatment without zanubrutinib:<br>• For people who have had at least one prior therapy:<br>o BR<br>o DRC<br>o FR<br>o FCR<br>o Clad-R<br>o ASCT in people for whom ASCT is suitable<br>• For people for whom chemo-immunotherapy is unsuitable:<br>o chlorambucil<br>o rituximab monotherapy<br>o BSC including blood product transfusions, plasma exchange, granulocyte | Treatment without zanubrutinib:<br>• BR<br>• DRC<br>• Ibrutinib | Other than BR and DRC, it was not possible to conduct comparisons with chemotherapy regimens or BSC, due to a lack of data in the literature to enable comparison of zanubrutinib with the comparators of interest (see CS, Appendix D).<br><br>However, BR and DRC currently represent the two most common regimens for the first-line treatment of WM in patients considered fit enough to tolerate them (13.1% and 16.2%, respectively [see Section B.1.3.5.2 of the CS]). In addition, BR and DRC are the third- and second-most common second-line regimens, respectively, behind ibrutinib (18.2%).[1] | The comparators are not in line with the NICE scope.<br>FR, FCR, and Clad-R have not been included as comparators due to lack of data according to the company.<br>ASCT has not been included in any of the literature searches reported in the CS.<br>Ibrutinib has been included as a comparator. However, NICE explicitly excluded ibrutinib as a comparator. |

| | Final scope issued by NICE | Decision problem addressed in the company submission | Rationale if different from the final NICE scope | ERG comment |
|---|---|---|---|---|
| | stimulating factors and intravenous Ig infusions | | Ibrutinib is also included as a comparator, given that:<br>Registry data indicates that BTK inhibitors (currently only ibrutinib is available) are an emerging standard of care in patients who have had ≥1 prior therapy, with ibrutinib being the most frequently used treatment in clinical practice (approximately 18.2% of cases).[1]<br>Ibrutinib is the only comparator for which direct head-to-head evidence is available – the safety and efficacy of zanubrutinib versus ibrutinib were evaluated in the largest Phase 3 trial of BTK inhibitors in WM (BGB-3111-302 [ASPEN]),[2] which forms the primary source of clinical evidence for this submission<br>Although ibrutinib is currently recommended for use in the CDF, the data collection arrangement for ibrutinib was anticipated to conclude in September 2020,[3] and NICE is subsequently due to update the guidance for ibrutinib in WM | |
| **Outcomes** | • OS<br>• PFS<br>• Response rates (ORR, MRR, VGPR/CR) | • Response rates (ORR, MRR, VGPR/CR)<br>• Duration of response<br>• PFS | N/A | The outcomes reported are in line with the NICE scope |

| | Final scope issued by NICE | Decision problem addressed in the company submission | Rationale if different from the final NICE scope | ERG comment |
|---|---|---|---|---|
| | • Time to next treatment<br>• Duration of response/remission<br>• Adverse effects of treatment<br>• HRQoL | • OS<br>• Time to next treatment<br>• HRQoL<br>• Adverse effects of treatment | | |
| **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.<br>• If the technology is likely to provide similar or greater health benefits at similar or lower cost than technologies recommended in published NICE technology appraisal guidance for the same indication, a cost-comparison may be carried out.<br>• 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.<br>• Costs will be considered from an NHS and Personal Social Services perspective.<br>• The availability of any patient access schemes for the intervention or comparator | Not addressed in the CS | Not addressed in the CS | Partly in line with the NICE scope. However, the company did not perform a fully incremental analysis. |

| | Final scope issued by NICE | Decision problem addressed in the company submission | Rationale if different from the final NICE scope | ERG comment |
|---|---|---|---|---|
| | technologies will be taken into account.<br>• The economic modelling should include the costs associated with diagnostic testing for MYD88 in people with Waldenström's macroglobulinaemia who would not otherwise have been tested, if appropriate. A sensitivity analysis should be provided without the cost of the diagnostic test. See section 5.9 of the Guide to the Methods of Technology Appraisals'. | | | |
| **Subgroups to be considered** | If the evidence allows the following subgroups will be considered:<br>• people with MYD88 mutation-positive Waldenström's macroglobulinaemia<br>• people with IgM-related conditions (e.g. paraproteinaemic neuropathies, cryoglobulinaemia, secondary cold agglutinin disease and Bing-Neel syndrome). | Not addressed in the CS | Not addressed in the CS | In line with the NICE scope. |

Based on Table B.1.1, pages 12-13 of the CS.[4]

ASCT = autologous stem cell transplantation; BR = bendamustine and rituximab; BSC = best supportive care; BTK = Bruton's tyrosine kinase; CDF = Cancer Drugs Fund; CHOP = cyclophosphamide, doxorubicin, vincristine and prednisone; Clad-R = cladribine and rituximab; CR = complete response; CS = company submission; DRC = dexamethasone, rituximab and cyclophosphamide; FCR = fludarabine, cyclophosphamide and rituximab; FR = fludarabine and rituximab; HRQoL = health-related quality of life; Ig = immunoglobulin; N/A = not applicable; NICE = National Institute for Health and Care Excellence; MRR = major response rate; ORR = objective response rate; OS = overall survival; PFS = progression-free survival; VGPR = very good partial response rate; WM = Waldenström's macroglobulinaemia.

### *2.1 Population*

The population defined in the scope is: Adults with Waldenström's macroglobulinaemia (WM): who have had at least one prior therapy, or whose disease is untreated, for whom chemo-immunotherapy is unsuitable.[5] The population in the CS is in line with the population in the scope.[4]

The population considered in the CS is also in line with the clinical trial for zanubrutinib in this indication, the ASPEN trial (Study BGB-3111-302) which included patients with WM who are relapsed/refractory or treatment naïve and considered to be unsuitable for chemotherapy.[4] The ASPEN trial included 37 patients who were treatment naïve (zanubrutinib (N=19) versus ibrutinib (N=18)) and 164 relapsed/refractory patients (zanubrutinib (N=83) versus ibrutinib (N=81)). Therefore, the evidence for treatment naïve patients is based on small numbers of patients.

The proposed indication for zanubrutinib is as follows: zanubrutinib as a single agent is indicated for the treatment of adult patients with WM who have received at least one prior therapy, or in first-line treatment for patients unsuitable for chemo-immunotherapy (CS, page 15).[4] The application was submitted to the European Medicines Agency (EMA) in [redacted], with a positive opinion from the Committee for Medicinal Products for Human Use (CHMP) anticipated in [redacted] [redacted]. UK approval is anticipated in [redacted].

**ERG comment:** Patients with cardiovascular disease and those taking warfarin were excluded from the ASPEN trial. According to the company, "such exclusion criteria are common in clinical trials in order to prevent patients with severe underlying comorbidities being exposed to potential side effects".[6] The company added that "a special warning regarding cardiac risk factors is included in the draft Summary of Product Characteristics (SmPC) for zanubrutinib: 'Cases of atrial fibrillation and atrial flutter have been reported particularly in patients with cardiac risk factors, hypertension, acute infections, and a previous history of atrial fibrillation.[7]' As underscored in the SmPC, 'patients with severe cardiovascular disease were excluded from [ibrutinib] clinical studies'.[7] Consequently, BeiGene did not want to expose patients with underlying comorbidities to known or unknown side-effects, patients with cardiovascular disease were excluded from the ASPEN trial".[6]

The randomised part of the ASPEN trial (Cohort 1) only included patients with *MYD88*$^{MUT}$. The company confirmed in the response to clarification that Cohort 1 did not include any patients with *MYD88*$^{WT}$ or with undetermined MYD88 status. In addition, the company stated that "A UK WM clinical expert confirmed that testing for MYD88 mutation is the standard of care at most of the 24 WM centres in the UK, which have treated 90% of the UK WM patient population since 2016".[6] Cohort 2 of the ASPEN trial received zanubrutinib and included only patients with *MYD88*$^{WT}$; the primary outcome (rate of independent review committee-assessed complete response or very good partial response) did not differ substantially between the zanubrutinib arm of Cohort 1 and Cohort 2 (28.4% versus 26.9% respectively). There is therefore little evidence of a difference in the efficacy of zanubrutinib by MYD88 mutation status for the primary outcome from the ASPEN trial.

### *2.2 Intervention*

The intervention (zanubrutinib) is in line with the scope.

The recommended daily dose of zanubrutinib is 320 mg, taken orally either OD (four 80 mg capsules) or BID (two 80 mg capsules).[7] According to the company, no additional tests or investigations are required prior to the administration of zanubrutinib (CS, page 15).[4]

### *2.3 Comparators*

The description of the comparators in the NICE scope is as follows: For people who have had at least one prior therapy: rituximab and bendamustine (BR); dexamethasone, rituximab and cyclophosphamide (DCR); fludarabine and rituximab with or without cyclophosphamide (FR or FCR); cladribine and rituximab (Clad-R); and autologous stem cell transplantation (SCT) in people for whom autologous SCT is suitable. For people for whom chemo-immunotherapy is unsuitable: chlorambucil; rituximab monotherapy; and best supportive care (BSC) including blood product transfusions, plasma exchange, granulocyte stimulating factors and intravenous immunoglobulin infusions.[5]

The company included BR and DCR as comparators. In addition, the company included ibrutinib as a comparator because registry data indicates that BTK inhibitors (currently only ibrutinib is available) are an emerging standard of care in patients who have had ≥1 prior therapy, with ibrutinib being the most frequently used treatment in clinical practice (approximately 18.2% of cases)[1] and because ibrutinib is the only comparator for which direct head-to-head evidence is available.[4]

The company stated that "other than BR and DRC, it was not possible to conduct comparisons with chemotherapy regimens or BSC, due to a lack of data in the literature to enable comparison of zanubrutinib with the comparators of interest" (CS, page 12).[4]

**ERG comment:** The feasibility of conducting indirect comparisons of zanubrutinib with FR, FCR and Clad-R will be discussed in Section 3.3 of this report. Regarding autologous SCT in people for whom autologous SCT is suitable, we asked the company why this comparator was not included in the CS because this comparator was not included in any of the literature searches performed by the company (Clarification Letter, Question A7).[6] The company responded that the eligibility criteria for the search were based on the draft scope set by NICE, not on the final scope. In addition, data from the UK WM Rory Morrison registry showed that 3% of all WM patients were considered for SCT.[1] Hence, SCT was not considered a relevant comparator by the company.

Regarding ibrutinib, NICE clearly stated in the response from NICE to comments on the draft scope, that ibrutinib had been removed as a comparator as it is currently available through the Cancer Drugs Fund and therefore not considered established practice (see NICE Response to comments on draft scope (pages 5-6).[8] When NICE recommends a drug for use within the Cancer Drugs Fund (CDF), NICE considers that there is plausible potential for the drug to satisfy the criteria for routine commissioning, but there is significant remaining clinical uncertainty which needs more investigation, through data collection in the NHS or clinical studies.[9] This means that the cost effectiveness of drugs recommended for use within the CDF has not yet been established. Therefore, any comparisons of effectiveness or cost effectiveness with CDF-drugs are equally uncertain.

### *2.4 Outcomes*

The NICE final scope lists the following outcome measures:

- Overall survival
- Progression-free survival
- Response rate
- Time to next treatment
- Duration of response/remission
- Adverse effects of treatment
- Health-related quality of life

These were all assessed in the ASPEN trial.

### *2.5 Other relevant factors*

According to the company, zanubrutinib is innovative because "treatment options for WM are limited across all lines of treatment and patients can cycle through and exhaust all available therapies.[10] No established treatment approach for WM has curative potential,[11] and once immuno-chemotherapy (e.g. rituximab combinations such as BR and DRC) and ibrutinib have been exhausted, there are no additional treatment options for relapsed/refractory patients." (CS, Section B.2.12).[4]

This appraisal does not fulfil the end-of-life criteria as specified by NICE because the life expectancy of patients eligible for zanubrutinib is well beyond 24 months. Therefore, treatment is not indicated for patients with a short life expectancy (normally less than 24 months). As stated by the company: "WM is an incurable disease with a median OS of 18.5 years in symptomatic patients. In an analysis of UK registry data from 671 patients with WM, 118 patients (18%) died between 1978 and 2019, equating to a 5-year OS of 90.5% and 10-year OS of 79.4%." (CS, page 16).[4]

According to the company, there are no known equality issues relating to the use of zanubrutinib in patients with WM (CS, Section B.1.4).[4]

# 3. CLINICAL EFFECTIVENESS

## *3.1 Critique of the methods of review(s)*

### 3.1.1 Searches

Appendix D1 of the CS details a systematic review performed to identify published evidence for current and future treatment options for patients with WM.

Searches were conducted on 24 September 2020, without a date limit. The searches were limited to English language only, and study design filters for randomised controlled trials and observational studies were applied. A summary of the sources searched is provided in Table 3.1.

**Table 3.1: Data sources for the clinical effectiveness systematic review**

| Search strategy element | Resource | Host/ source | Reported date range | Date searched |
|---|---|---|---|---|
| Electronic databases | Embase | Proquest | Not reported | 24.9.20 |
| | Medline & In-Process | Proquest | Not reported | 24.9.20 |
| | Cochrane Central Register of Controlled Trials (CENTRAL) | Wiley | 1991-2019/12 | 24.9.20 |
| | DARE | Wiley | Not searched | Not searched |
| Source: Appendix D of the Company's submission.[12] | | | | |

**ERG comment:** A single set of searches was undertaken to identify clinical effectiveness studies. The CS Appendix D[12] provided sufficient details for the ERG to appraise the literature searches. Three databases were searched; no trials registers or grey literature was included. For the most part, searches were well documented, making them transparent and reproducible.

The clinical effectiveness searches presented in the CS Appendix D were conducted in September 2020. As the clinical effectiveness searches were run from over eight months ago, the ERG considers it possible that potentially relevant studies published since September 2020 may be missing from the systematic review. This was queried during the clarification process and the company responded to say they had conducted a separate targeted search to explore whether there were new publications. The clarification response[13] stated that two new publications which might be included were not relevant to the network meta-analysis. Details of the strategies used, date span, date of search, number of results retrieved were not provided to the ERG. The clarification response did not include an updated PRISMA flowchart. The ERG is unable to assess how the targeted searches were conducted or screened.

The company's clinical effectiveness searches were well constructed. A single simultaneous search of Embase, Medline and In-Process was carried out via Proquest. The Proquest search included both randomised controlled trial and observational study design filters, which contained a wide range of synonyms and word variants. Appropriate Emtree and MeSH indexing was incorporated into the Proquest strategy. Comprehensive use of adjacency, truncation and wildcards was noted in all strategies; for the most part, this enabled effective use of phrase searching.

The company reported inclusion of the Database of Abstracts of Reviews of Effects (DARE) in their Cochrane Library search (D.1.1.1), however no strategy was presented. DARE was removed from the

Cochrane Library in September 2018. During clarification, the company confirmed this was a reporting error and DARE was not searched for this topic.

The ERG noted that both the Embase and Cochrane Library search strategies used very limited synonyms for the condition, WM. The ERG identified several word and spelling variants (see below) which could have been included to increase recall of potentially relevant WM studies.

| | | |
|---|---|---|
| 1 | Plasmacytoid Lymphocytic Lymphoma.ti,ab. | 16 |
| 2 | (Lymphoplasmacytoid Lymphoma or Macroglobuli?emia*).ti,ab. | 4749 |
| 3 | (waldenstroem macroglobulin* or atypical macroglobulin* or waldenstrom macroglobulin*).ti,ab. | 1848 |
| 4 | (macro globuli?nemia or macrocryoglobulin?emia).ti,ab. | 13 |
| 5 | or/1-4 | 4965 |

The ERG noted that both the Proquest and Cochrane Library search strategies used very limited synonyms for the drugs included in the treatment facet. The ERG identified several word and spelling variants, and CAS Registry Numbers could have been included to increase recall of potentially relevant WM studies. Please see Appendix 1 for full details of alternative synonyms which could have improved strategy performance.

The Embase search strategy contains a typographical error (line 75, pg 17). The omission of a space between the OR search operator and the final free text term means this search missed +550 references (see below):

| | | |
|---|---|---|
| 2 | (Bendamustine or belrapzo or bendamustin or bendeka or cytostasan* or levact or ribomustin or ribovact or treanda or?imet*).ti,ab. | 4280 |
| 3 | (Bendamustine or belrapzo or bendamustin or bendeka or cytostasan* or levact or ribomustin or ribovact or treanda or ?imet*).ti,ab. | 4855 |

The ERG conducted further tests on the performance of the ? wildcard character at the start of a word. According to Ovid, the optional wild card '?' character stands for zero or one characters within a word or at the end of a word. This wildcard should not be used at the start of a word. The ERG identified three word variant synonyms for bendamustine that should have been searched as full words for the strategy to work correctly.

| | | |
|---|---|---|
| 8 | ?imet.ti,ab. | 156 |
| 9 | (cimet or imet or zimet).ti,ab. | 199 |

The impact on recall was sufficient that the ERG concluded the search phrase '?imet*' may not have performed as the searcher intended.

During the clarification process, the ERG queried the rationale for applying an English language limit to the Embase/MEDLINE clinical and cost effectiveness searches. The company responded that "The rationale for limiting the searches to English literature only was based on guidance provided by NICE; Chapter 5.4 of Developing NICE guidelines: the manual states that with regards to limits and filters, searches should be limited to studies reported in English." The company cited the NICE manual for

developing guidelines[14] as their source. Although the Guidelines manual does recommend using an English language limit, the user guide for company STA evidence submissions[15] refers to the guide to the methods of technology appraisal[16] and guidance from the Centre for Reviews and Dissemination.[17] The latter source clearly states that "limiting searches to English language papers can introduce language bias."[17] Consequently the ERG remains concerned that limiting the searches to English language only studies may have introduced language bias. Current best practice states that "Whenever possible review authors should attempt to identify and assess for eligibility all possibly relevant reports of trials irrespective of language of publication"[18] and that "research related to language bias supports the inclusion of non-English studies in systematic reviews".[19, 20]

In summary, the ERG felt that the clinical effectiveness searches would have benefited from clearer reporting, more comprehensive terminology for WM and all named drugs, a full update search, and removal of the English language limit.

### 3.1.2 Inclusion criteria

The eligibility criteria used in the search strategy for randomised controlled trials (RCTs) and non-RCTs is presented in Table 3.2.

**Table 3.2: Eligibility criteria used in search strategy for RCT and non-RCT evidence**

| Category | Inclusion criteria | Exclusion criteria |
|---|---|---|
| Population | • Adult patients with WM, with or without previous treatment (i.e. treatment naïve or relapsed/refractory) | • Patients receiving treatment for secondary malignancies (focus of treatment aims to treat another underlying malignancy)<br>• Healthy subjects<br>• Children (<18 years of age) |
| Study design | • Interventional<br>• RCT (blinded and open label)<br>• Non-RCT (including prospective, interventional observational/real-world pragmatic studies)<br>• Single-arm trials<br>• Phase 2 and 3 | • Studies which do not have as main objective to study intervention effectiveness, e.g.<br>o Biomarker studies<br>o Prognostic factor studies<br>o Non-interventional studies<br>• SLRs<br>• Meta-analyses<br>• Phase 1 studies<br>• Case studies<br>• Non-human studies<br>• Study design not included in the inclusion criteria |
| Interventions | • Zanubrutinib<br>Chemo-immunotherapy including the following treatments:<br>• BR<br>• DRC<br>• FR<br>• FCR<br>• Clad-R<br>• CHOP with or without rituximab<br>• Ibrutinib in people who have had at least 1 prior therapy. | • Rituximab combinations not listed as relevant interventions<br>• Interventions not included in the inclusion criteria |

| Category | Inclusion criteria | Exclusion criteria |
|---|---|---|
| | For people who are not eligible for chemo-immunotherapy:<br>• Chlorambucil, with or without rituximab<br>• Rituximab monotherapy<br>• Ibrutinib in people who have had at least 1 prior therapy<br>• BSC | |
| Comparators | • Relevant interventions<br>• No comparator (single-arm trials) | • Rituximab combinations not listed as relevant interventions<br>• Comparators not included in the inclusion criteria |
| Outcomes | • ORR<br>• CR<br>• VGPR<br>• PR<br>• MRR<br>• Duration of response (CR/VGPR/PR)<br>• OS<br>• PFS<br>• Safety (e.g. including AEs, discontinuation)<br>• HRQoL | • Only reporting outcomes not included in the inclusion criteria |
| Time limit | • N/A | • N/A |
| Language | • English | • All other languages |

Source: CS, Appendix D, Table B.5.5, page 19-20.
AE = adverse event; BR = rituximab and bendamustine; BSC = best supportive care; CHOP = cyclophosphamide, doxorubicin, vincristine and prednisone; Clad-R = cladribine and rituximab; CR = complete response; DRC = dexamethasone, rituximab and cyclophosphamide; FCR = fludarabine, cyclophosphamide and rituximab; FR = fludarabine and rituximab; HRQoL = health-related quality of life; MRR = major response rate; N/A = not applicable; ORR = overall response rate; OS = overall survival; PFS = progression-free survival; PR = partial response; RCT = randomised controlled trial; SLR = systematic literature reviews; VGPR = very good partial response; WM = Waldenström's Macroglobulinaemia

**ERG comment:** Given the final scope issued by NICE, the study designs, population, intervention, and outcomes were appropriate. However, the ERG notes that not all the comparators specified in the NICE scope were included in the CS.[5] Specifically, the NICE scope lists autologous stem cell transplantation (SCT), in people for whom autologous SCT is suitable as a relevant comparator for people who have had at least one prior therapy. This comparator was not included in the searches performed by the company.

### 3.1.3 Critique of data extraction

Data extraction was performed by a single reviewer using a predefined data extraction template and was quality checked by a separate reviewer against the source publication.

**ERG comment:** To minimise error during data extraction, it is usually advised that data extraction is carried out independently by two independent reviewers.[21]

### 3.1.4 Quality assessment

It is unclear whether the quality assessment was carried out by a single reviewer or two (or more) independent reviewers. They quality assessment was based on the seven-item NICE quality assessment checklist,[22] including selection bias, performance and detection bias, attrition bias, and reporting bias.

**ERG comment:** To minimise error in quality assessment, it is usually advised that quality assessment is carried out independently by two independent reviewers.[21]

### 3.1.5 Evidence synthesis

As stated by the company, "Efficacy data supporting the use of zanubrutinib for the treatment of WM are primarily provided by a single Phase 3 study (ASPEN). Therefore, a meta-analysis was not conducted." (CS, Section B.2.8, page 46).[4]

The company provided an indirect treatment comparison for the comparisons with rituximab and bendamustine (BR) and dexamethasone, rituximab and cyclophosphamide (DRC) (See also CS, Section B.2.9).

**ERG comment:** The ERG agrees that a meta-analysis of zanubrutinib studies is not warranted. The indirect comparison for the comparisons with BR and DRC is critiqued in Sections 3.3 and 3.4 of this report.

## *3.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)*

### 3.2.1 Details of the included trial: the ASPEN trial

The main evidence for the clinical effectiveness of zanubrutinib was from the ASPEN trial (the BGB-3111-302 study).[2] Supplemental long-term efficacy data is provided by the Phase 1/2 BGB-3111-AU-003 study (see Section 3.2.7 of this report), and pooled safety data are presented for all WM patients (see Section 3.2.6 of this report).

The ASPEN trial is an ongoing Phase 3, open-label, two-arm, multicentre, randomised study of zanubrutinib versus ibrutinib for the treatment of WM in patients with relapsed/refractory disease, or who are treatment naïve and ineligible for chemoimmunotherapy.[2] The study was designed with two cohorts, according to MYD88 status. Cohort 1 includes patients with *MYD88*$^{MUT}$ who were randomised to either zanubrutinib or ibrutinib. Cohort 2 included patients with *MYD88*$^{WT}$, and all patients were assigned to zanubrutinib.[2] As Cohort 2 was on single-arm observational study, we will focus on Cohort 1 in this part of our report. Cohort 2 will be discussed in Section 3.2.7 of this report (Supporting evidence). The company confirmed in the response to clarification that Cohort 1 did not include any patients with *MYD88*$^{WT}$ or with undetermined MYD88 status.

A total of 201 patients were randomised to zanubrutinib or ibrutinib; 164 patients had relapsed/refractory disease (zanubrutinib, n=83 versus ibrutinib, n=81) and 37 were treatment naïve (zanubrutinib, n=19 versus ibrutinib, n=18). In response to the clarification letter, the company added that "A total of 33 UK patients were randomised in the ASPEN study (30 randomised to Cohort 1 and 3 to Cohort 2), and 32 patients were treated. In Cohort 1, 13 patients were treated with ibrutinib, and 16 of 17 randomised were treated with zanubrutinib".[6]

**Table 3.3: The ASPEN trial – Cohort 1 (randomised patients)**

| | |
|---|---|
| **Study** | Study BGB-3111-302 (ASPEN; NCT03053440)[23, 24] |
| **Study design** | Multicentre, randomised, open-label, Phase 3 trial |
| **Locations (number of patients recruited)** | Australia (68), UK (33), Italy (27), Spain (24), US (21), Poland (19), Greece (13), Czech Republic (9), Sweden (7), Netherlands (5), Germany (2) and France (1) |
| **Study status** | Ongoing<br>First patient treated: 25 January 2017<br>Data cut-off date: 31 August 2019 |
| **Population** | Patients with WM who are relapsed/refractory (n=164) or treatment naïve and considered to be unsuitable for chemotherapy (n=37) |
| **Key eligibility criteria** | Men and women aged ≥18 years<br>Clinical and definitive histologic diagnosis of WM that is either treatment naïve and unsuitable for standard chemoimmunotherapy, or relapsed/refractory<br>Meet at least one criterion from the Seventh IWWM<br>ECOG Performance Status 0–2<br>No prior exposure to a BTK inhibitor<br>No WM central nervous system involvement |
| **Intervention(s)** | Zanubrutinib 160 mg BID to progression (n=102) |
| **Comparator(s)** | Ibrutinib 420 mg OD to progression (n=99) |
| **Reported outcomes specified in the decision problem** | OS<br>PFS<br>Response rate<br>Time to next treatment<br>DOR/remission<br>Adverse effects of treatment<br>HRQoL |
| **All other reported outcomes** | Symptom resolution<br>Serum IgM<br>TTD |

Source: CS, Table B.2.1 and B.2.2, pages 22-25.[4]
BID = twice daily; BTK = Bruton's tyrosine kinase; DOR = duration of response; ECOG = Eastern Cooperative Oncology Group; HRQoL = health-related quality of life; IWWM = International Workshop on Waldenström's Macroglobulinemia; OD = once daily; OS = overall survival; PFS = progression-free survival; TTD = time to discontinuation; WM = Waldenström's macroglobulinaemia.

### 3.2.2 Statistical analyses of the ASPEN trial

The ASPEN trial was designed to show the superiority of zanubrutinib over ibrutinib in patients with *MYD88*$^{MUT}$ WM (relapsed/refractory arm of Cohort 1). The primary efficacy analyses were the rate of complete response or very good partial response (CR/VGPR), as assessed by IRC with adaption of the response criteria updated at the Sixth International Workshop on Waldenström Macroglobulinemia (IWWM) every 28 days and every 84 days after Cycle 12. Two analysis sets were considered: 1) the relapsed/refractory analysis set, and 2) the intention to treat (ITT) analysis set. The ITT analysis set comprised all randomised patients assigned to a treatment arm, while the relapsed/refractory analysis set comprised all patients in the ITT analysis set with at least one prior line of therapy. In the CS, the primary results are for the ITT analysis set.

The analyses of the primary endpoint used a Cochran-Mantel-Haenszel test in a hierarchical fixed-sequence procedure to adjust for multiplicity to test for differences in CR/VGPR rates, stratified by the CXCR4 status (WHIM versus WT/missing), prior line of therapy (1–3 versus >3 for the relapsed/refractory analysis set; 0 versus 1–3 versus >3 in the ITT analysis set) and age group (≤65 years versus >65 years). The analysis was performed in the relapsed/refractory analysis set first, at least 15 months after 90% enrolment in this analysis set was completed. If superiority was demonstrated with statistical significance in the relapsed/refractory analysis set, superiority was further tested in the ITT analysis set.

**Table 3.4: Summary of statistical analysis: Cohort 1 (ITT analysis set)**

| | **ASPEN** |
|---|---|
| **Hypothesis objective** | Demonstrate superiority of zanubrutinib compared with ibrutinib in CR/VGPR rate, planned for 12 months after the last relapsed/refractory patient was recruited (~15 months average follow-up) |
| **Statistical analysis** | The proportion of patients achieving CR/VGPR at ~15 months was compared between zanubrutinib and ibrutinib using a Cochran-Mantel-Haenszel test in a hierarchical fixed-sequence procedure to adjust for multiplicity. The analysis was initially performed in the relapsed/refractory analysis set, then, if statistical significance was shown, in the ITT analysis set. *CXCR4* status (WHIM versus WT/missing), prior line of therapy (1–3 versus >3 for the relapsed/refractory analysis set; 0 versus 1–3 versus >3 in the ITT analysis set) and age group (≤65 years versus >65 years) were used as stratification variables. Superiority was to be declared if the 2-sided P value from the Cochran-Mantel-Haenszel test was <0.05 and the estimated difference was positive. |
| **Sample size, power calculation** | The study aimed to recruit a total of 150 relapsed/refractory patients randomised 1:1 in Cohort 1, which would provide 81.4% power to demonstrate superiority under an assumed CR/VGPR rate of 35% for zanubrutinib versus 15% for ibrutinib, using a normal approximation of a binomial test and a 2-sided alpha of 0.05. |
| **Data management, patient withdrawals** | No specific information presented in the CS or clinical study report. |
| Source: CS, Section B.2.4.[2]<br>CR = complete response, CS = company submission, MRR = major response rate, VGPR = very good partial response | |

**ERG comment:** The analysis of the ASPEN trial mainly used appropriate statistical methods, though the ERG has one concern: it is unclear how the analysis dealt with missing data.

### 3.2.3 Baseline characteristics of the ASPEN trial

The median age of all patients in the ITT analysis set (Cohort 1) was 70.0 years. The majority of patients were male (66.7%), white (91.0%), had an ECOG performance status of 0 or 1 and were enrolled in sites in Europe (59.7%), Australia/New Zealand (30.8%) or North America (9.5%). The demographics and baseline characteristics were generally similar across treatment arms, however, more patients randomised to zanubrutinib than ibrutinib were >75 years old (33.3% and 22.2%, respectively) and more were anaemic (haemoglobin ≤110 g/L in 65.7% and 53.5% of patients, respectively).[24]

A summary of baseline characteristics and demographics for Cohort 1 is shown in Table 3.5 below.

**Table 3.5: Demographics and baseline characteristics: Cohort 1 (ITT analysis set)**

| Demographic/baseline characteristic | Zanubrutinib (N=102) | Ibrutinib (N=99) | Total (N=201) |
|---|---|---|---|
| Median age (min, max), years | 70.0 (45, 87) | 70.0 (38, 90) | 70.0 |
| >75 years, n (%) | 34 (33.3) | 22 (22.2) | 56 (27.9) |
| Gender, n (%) | | | |
| Male | 69 (67.6) | 65 (65.7) | 134 (66.7) |
| Race, n (%) | | | |
| White | 88 (86.3) | 95 (96.0) | 183 (91.0) |
| Asian | 4 (3.9) | 0 | 4 (2.0) |
| Unknown | 10 (9.8) | 4 (4.0) | 14 (7.0) |
| ECOG PS | | | |
| 0 | 46 (45.1) | 42 (42.4) | 88 (43.8) |
| 1 | 50 (49.0) | 50 (50.5) | 100 (49.8) |
| 2 | 6 (5.9) | 7 (7.1) | 13 (6.5) |
| Prior lines of therapy, n (%) | | | |
| 0 | 19 (18.6) | 18 (18.2) | 37 (18.4) |
| 1-3 | 76 (74.5) | 74 (74.7) | 150 (74.6) |
| >3 | 7 (6.9) | 7 (7.1) | 14 (7.0) |
| Genotype | | | |
| *MYD88*$^{MUT}$/*CXCR4*$^{WT}$ | 91 (89.2) | 90 (90.9) | 181 (90.0) |
| *MYD88*$^{MUT}$/*CXCR4*$^{WHIM}$ | 11 (10.8) | 8 (8.1) | 19 (9.5) |
| IPSS WM, n (%) | | | |
| Low | 17 (16.7) | 13 (13.1) | 30 (14.9) |
| Intermediate | 38 (37.3) | 42 (42.4) | 80 (39.8) |
| High | 47 (46.1) | 44 (44.4) | 91 (45.3) |
| Haemoglobin ≤110 g/L, n (%) | 67 (65.7) | 53 (53.5) | 120 (59.7) |

Source: CS, Table B.2.7, page 33.[4]
*CXCR4* = C-X-C Motif Chemokine Receptor 4; ECOG PS = Eastern Cooperative Oncology Group performance status; IPSS WM = International Prognostic Scoring System for Waldenström's Macroglobulinemia; ITT = intention-to-treat; MYD88 = myeloid differentiation primary response gene 88; n = number of patients in the category; N = number of patients evaluable; WHIM = warts, hypogammaglobulinemia, infections, myelokathexis; WT = wild-type

**ERG comment:** The proportion of >75 year olds was higher in the zanubrutinib group compared with the ibrutinib group. The disease affects older people more, but this also means that people over 75 years of age may have more room for improvement. The company was asked to explain how the different proportion of people over 75 years of age in each arm of ASPEN may bias the results and how this was adjusted for in the analyses.[6] In their response to Question A8 of the clarification letter,[6] the company confirmed that "In Cohort 1 of ASPEN, more patients in the zanubrutinib arm were >75 years or ≤65 years than those in the ibrutinib arm (33.3% versus 22.2% and 40.2% versus 29.3%, respectively)." In addition, the company stated that "Patients aged >75 years had a lower VGPR/CR rate compared with those aged ≤75 years in the zanubrutinib arm (20.6% vs 32.4%), and a higher VGPR/CR rate than those aged ≤75 years in the ibrutinib arm (31.8% vs 15.6%).[24] Based on the pre-specified subgroup analysis,

zanubrutinib treatment was favoured in patients ≤75 years with a risk difference (95% CI) of 16.8% (3.0, 30.5), and ibrutinib treatment was favoured in patients >75 years with a risk difference (95% CI) of -11.2% (-35, 12.5) (Table 1 in Response to Clarification). As a result, having more >75 years patients in the zanubrutinib group could bias the risk difference towards ibrutinib".[6]

In the primary analysis, age group (>65 vs ≤65 years) was included in the stratified Cochran-Mantel-Haenszel (CMH) test in the primary and secondary efficacy analyses. However, analyses were not adjusted for age (>65 vs ≤65 years) or for age group of >75 versus ≤75 years.[6]

The proportion of anaemic patients was also higher in the zanubrutinib group compared with the ibrutinib group (haemoglobin ≤110 g/L in 65.7% and 53.5% of patients, respectively).

### 3.2.4 Risk of bias assessment of the ASPEN trial

The company assessed the quality of the ASPEN trial using the seven-item NICE quality assessment checklist,[22] including randomisation, allocation concealment, baseline characteristics, blinding, withdrawals, outcome selection and reporting, and ITT analysis. No information was provided on the number of reviewers who assessed the quality of the ASPEN trial, the company concluded that the study "was of good quality with respect to randomisation methods, baseline comparability between the randomised groups, imbalances in withdrawals between the randomised groups, selective outcomes, and reporting and statistical analysis methods".[4]

**Table 3.6: Study quality assessment using the NICE checklist (ASPEN Trial)**

| Risk of bias item | Company | ERG |
|---|---|---|
| Randomisation: was randomisation carried out appropriately? | Yes | The method of randomisation was appropriate. |
| Concealment grade: was the concealment of treatment allocation adequate? | NC | The study was not reported as having used allocation concealment. |
| Baseline comparability: were the groups similar at the outset of the study in terms of prognostic factors? | Yes, patients were well balanced between study arms for key characteristics, Randomisation was stratified by warts, hypogammaglobulinemia, immunodeficiency, and myelokathexis (WHIM) (CXCR4WHIM) syndrome-like mutation status and number of prior lines of therapy. | In spite of being stratified for some domains, there was an 11% difference in the percentage of patients older than 75 (zanubrutinib 33% versus ibrutinib 22%), and more were anaemic (haemoglobin ≤110 g/L in 66% vs 54%). |
| Blinding: were the care providers, participants and outcome assessors blind to treatment allocation? | No formal blinding was used as the study is open-label. Open-label might be mitigated by determination of the primary endpoint by independent review committee | The study was not reported as blinded. The clinical study report states that "The independent DMC was not blinded."[24] The clinical study report also reports that the lack of blinding may have introduced bias: "Due to the open-label nature of the study, access to aggregated data summaries with actual study treatment assignment of |

| Risk of bias item | Company | ERG |
|---|---|---|
| | | the randomized arms (Cohort 1 Arm A versus Arm B) whilst the study was ongoing may have introduced unwanted bias due to the possibility of inconsistent queries among patients with different treatments, or over-interpretation of immature, accruing data."[24]<br>The determination of the primary outcome by an independent review committee partly but does not fully mitigate this. |
| Follow-up: were there any unexpected imbalances in drop-outs between groups? | No, similar study treatment discontinuation between groups | Discontinuation rates were comparable (9% in ibrutinib versus 4% in zanubrutinib) |
| Selective reporting and other sources of bias: is there any evidence to suggest that the authors measured more outcomes than they reported? State any important concerns about bias not addressed in the other domains in the tool. If particular questions/entries were pre-specified in the review's protocol, responses should be provided for each question/entry. | No | The primary and secondary endpoints were consistent between the published protocol and trial reports. |
| Analysis: did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data? | Yes | An ITT analysis appears to have been done, however the AE assessment does not include one patient from each of the groups |
| Source: CS, Appendix D, Table B.5.13, page 34.[12]<br>ITT = intention-to-treat; N/A = not applicable; NC = not clear; NICE = National Institute for Health and Care Excellence | | |

**ERG comment**: The ERG examined the protocol,[25] the publication[2] and clinical study report[24] for the ASPEN trial and assessed it against the above criteria. The NICE checklist cited by the company does not have seven, but five sections with a total of 27 items. It is normally recommended that two reviewers are involved in the assessment of study quality to avoid bias and error.

Randomisation method was adequately described, and allocation concealment procedures were not reported. Therefore, the trial is at an unclear risk of selection bias. None of the study groups were reported as being blinded, raising the risk of assessment bias; this is acknowledged in the clinical study report for the ASPEN trial. The primary endpoints described in the protocol and CSR were the same (complete response or very good partial response). The clinical study report (page 54) notes that "an unstratified analysis of the primary endpoint was to be performed in the sensitivity analysis but was not

conducted".[24] Based on the above considerations, our overall assessment is that the ASPEN trial has an unclear risk of bias.

### 3.2.5 Efficacy results of the ASPEN trial

The median follow-up time as of the data cut-off date was 19.5 months for zanubrutinib-treated patients and 19.4 months for ibrutinib-treated patients. A total of 201 patients were randomised (102 in the zanubrutinib arm and 99 in the ibrutinib arm) with 164 (81.6%) patients having relapsed/refractory disease (83 in the zanubrutinib treatment arm and 81 in the ibrutinib treatment arm). Two relapsed/refractory patients were randomised but not treated, one in the zanubrutinib treatment arm due to an adverse event (AE; unrelated to screening procedures) and one in the ibrutinib treatment arm due to progressive disease (central nervous system). As of the data cut-off date (31 August 2019), a total of 158 patients (78.6%) were continuing study treatment (81 patients [79.4%] in the zanubrutinib treatment arm and 77 patients [77.8%] in the ibrutinib treatment arm). The most common reason for discontinuing study treatment was progressive disease (seven [6.9%] zanubrutinib versus five [5.1%] ibrutinib-treated patients) and AE (four [3.9%] zanubrutinib treated patients versus nine [9.1%] ibrutinib-treated patients). A total of 158 (78.6%) patients were continuing to participate in the study and 41 (20.4%) discontinued from the study.

Patient disposition for Cohort 1 (ITT analysis set) are summarised in Table 3.7 below.

**Table 3.7: Patient disposition: Cohort 1 (ITT analysis set)**

| Category | Zanubrutinib (n=102) | Ibrutinib (n=99) | Total (n=201) |
|---|---|---|---|
| Randomised, not treated, n (%) | 1 (1.0) | 1 (1.0) | 2 (1.0) |
| AE | 1 (1.0) | 0 (0.0) | 1 (0.5) |
| Progressive disease | 0 (0.0) | 1 (1.0) | 1 (0.5) |
| Treated, n (%) | 101 (99.0) | 98 (99.0) | 199 (99.0) |
| On treatment, n (%) | 81 (79.4) | 77 (77.8) | 158 (78.6) |
| Discontinued, n (%) | 20 (19.6) | 21 (21.2) | 41 (20.4) |
| AE | 4 (3.9) | 9 (9.1) | 13 (6.5) |
| Progressive disease | 7 (6.9) | 5 (5.1) | 12 (6.0) |
| Investigator's discretion | 2 (2.0) | 4 (4.0) | 6 (3.0) |
| Withdrawal by patient | 5 (4.9) | 0 (0.0) | 5 (2.5) |
| Other | 2 (2.0) | 3 (3.0) | 5 (2.5) |
| Median study follow-up (months) | 19.47 | 19.38 | 19.45 |
| Min, Max | 0.4, 31.2 | 0.5, 31.1 | 0.4, 31.2 |

Source: CS, Table B.2.6, page 32.[4]
AE = adverse event; ITT = intention-to-treat; n = number of patients in the category; N = number of patients evaluable

#### 3.2.5.1 IRC-assessed VGPR/CR rate (primary endpoint)

In Cohort 1, the rate of IRC-assessed CR and VGPR was 28.4% in all patients treated with zanubrutinib and 19.2% in patients treated with ibrutinib (95% CI, -1.5–22.0; p=0.09). The estimated difference between the two arms adjusted for the stratification factors and age group was 10.2% (Table 3.8).[2]

**Table 3.8: IRC-assessed response in Cohort 1 (ITT analysis set)**

| Assessment | Zanubrutinib (n=102) | Ibrutinib (n=99) |
|---|---|---|
| CR + VGPR, n (%) | 29 (28.4) | 19 (19.2) |
| CR + VGPR risk difference (95% CI) | 10.2 (-1.5–22.0); p=0.09 | |
| OR, n (%) | 96 (94.1) | 92 (92.9) |
| MRR, n (%) | 79 (77.5) | 77 (77.8) |
| Best overall response, n (%) | | |
| CR | 0 (0.0) | 0 (0.0) |
| VGPR | 29 (28.4) | 19 (19.2) |
| PR | 50 (49.0) | 58 (58.6) |
| Minor response | 17 (16.7) | 15 (15.2) |
| Stable disease | 3 (2.9) | 3 (3.0) |
| Progressive disease | 2 (2.0) | 2 (2.0) |
| Not evaluable | 0 (0.0) | 2 (2.0) |

Source: CS, Table B.2.9, page 36.[4]
CI = confidence interval; CR = complete response; IRC = independent review committee; ITT = intention-to-treat; MRR = major response rate; n = number of patients in the category; N = number of patients evaluable; OR = overall response; PR = partial response; VGPR = very good partial response rate.

In the relapsed/refractory population, 28.9% of patients treated with zanubrutinib and 19.8% treated with ibrutinib achieved VGPR or CR (with estimated difference of 10.7% (95% CI: 2.5 to 23.9; p=0.116)).[2]

According to the company, the testing for the primary endpoint of VGPR or CR rate superiority required testing in the relapsed/refractory analysis set prior to testing in the ITT analysis set. The primary efficacy endpoint was not significant in the relapsed/refractory analysis set (p=0.116), thus the study did not meet the primary efficacy endpoint and testing for other endpoints and resulting p-values in the following sections are descriptive.[24]

#### 3.2.5.2 IRC-assessed duration of response (secondary endpoint)

In Cohort 1, the median durations of VGPR or CR and major response according to overall combined assessment were not reached in either treatment arm who achieved a response to the study treatment, as shown in Table 3.9.

**Table 3.9: IRC-assessed duration of response in Cohort 1 (ITT analysis set)**

| Assessment | Zanubrutinib (n=102) | Ibrutinib (n=99) |
|---|---|---|
| Duration of CR or VGPR | | |
| Median follow-up, months (95% CI) | 13.6 (9.7–16.6) | 7.7 (2.8–12.9) |
| Median DOR, months (95% CI) | NE (NE-NE) | NE (8.0-NE) |
| Event-free rate at, % (95% CI) | | |
| 12 months | 100.0 (NE–NE) | 64.2 (28.8–85.4) |
| 18 months | 92.9 (59.1–99.0) | 64.2 (28.8–85.4) |
| 24 months | NE (NE–NE) | NE (NE–NE) |
| Duration of Major Response | | |
| Median follow-up, months (95% CI) | 14.8 (13.8–16.8) | 13.9 (12.3–15.7) |
| Median DOR, months (95% CI) | NE (NE-NE) | NE (NE-NE) |
| Event-free rate at, % (95% CI) | | |
| 12 months | 94.4 (85.8–97.9) | 87.9 (77.0–93.8) |
| 18 months | 85.2 (71.7–92.6) | 87.9 (77.0–93.8) |
| 24 months | 85.2 (71.7–92.6) | 81.6 (62.4–91.6) |

Source: CS, Table B.2.10, page 37.
CI = confidence interval; CR = complete response; DOR = duration of response; IRC, independent review committee; ITT = intention-to-treat; N = number of patients evaluable; NE = not evaluable; VGPR = very good partial response.

### 3.2.5.3 IRC-assessed progression-free survival (secondary endpoint)

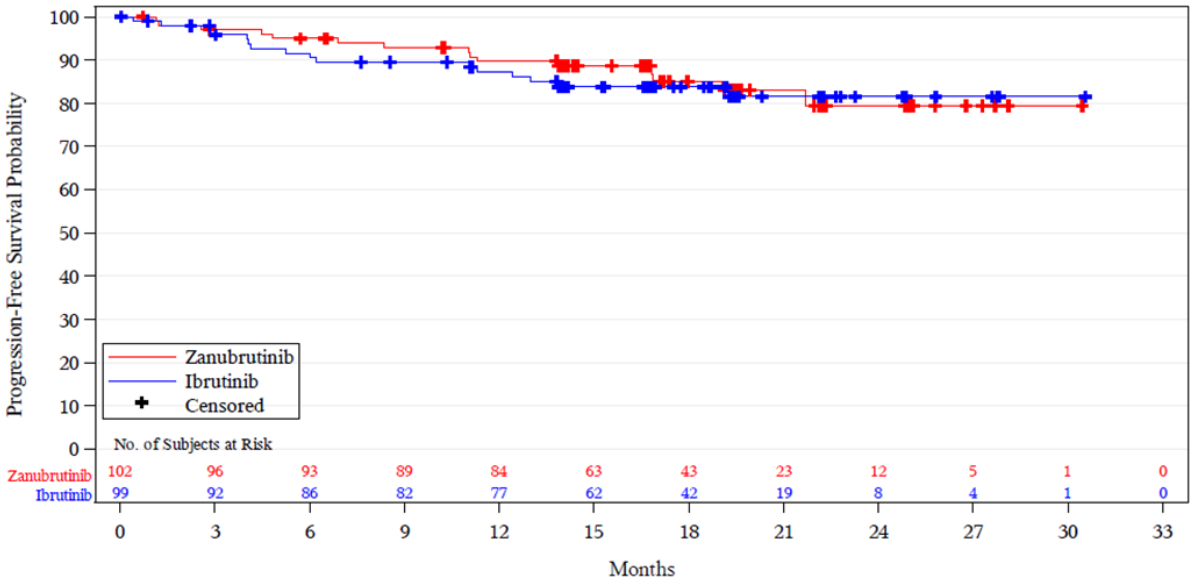
At the time of the data cut-off date, median PFS was not reached in either treatment arm of Cohort 1. The event-free rates at 12 months for patients treated with zanubrutinib or ibrutinib were 89.7% and 87.2%, respectively,[24] and 85.0% and 83.8% at 18 months[2] (Table 3.10 and Figure 3.1).

**Table 3.10: IRC-assessed progression-free survival in Cohort 1 (ITT analysis set)**

| Assessment | Zanubrutinib (n=102) | Ibrutinib (n=99) |
|---|---|---|
| Median follow-up, months (95% CI) | 18.0 (16.7–19.4) | 18.5 (16.7–19.3) |
| Median PFS, months (95% CI) | NE (NE–NE) | NE (NE–NE) |
| Events, n (%) | | |
| Progressive disease | 13 (12.7) | 10 (10.1) |
| Death | 2 (2.0) | 6 (6.1) |
| Event-free rate at, % (95% CI) | | |
| 6 months | 95.0 (88.4–97.9) | 91.6 (83.9–95.7) |
| 9 months | 92.9 (85.7–96.5) | 89.5 (81.3–94.2) |
| 12 months | 89.7 (81.7–94.3) | 87.2 (78.6–92.5) |
| 18 months | 85.0 (75.2–91.2) | 83.8 (74.5–89.9) |
| 24 months | 79.4 (66.2–88.0) | 81.5 (71.1–88.5) |

Source: CS, Table B.2.11, page 38.[4]
CI = confidence interval; IRC = independent review committee; ITT = intention-to-treat; n = number of patients in the category; N = number of patients evaluable; NE = not evaluable; PFS = progression-free survival

**Figure 3.1: IRC-assessed progression-free survival in Cohort 1 (ITT analysis set)**




Source: Response to clarification, Question A12, Figure 1.[6]
No. = number; PFS = progression free survival

### 3.2.5.4 IRC-assessed time to response (secondary endpoint)

The median time to VGPR or CR according to overall combined IRC assessment was shorter in the zanubrutinib arm than the ibrutinib arm (4.8 versus 7.4 months).[24] Time to major response (2.8 versus 2.8 months)[2] and overall response (1.0 versus 1.0 months)[24] were the same between the treatment groups.

### 3.2.5.5 Overall survival (exploratory endpoint)

At the time of the data cut-off date, OS had not been reached in either treatment arm (Table 3.11).[24]

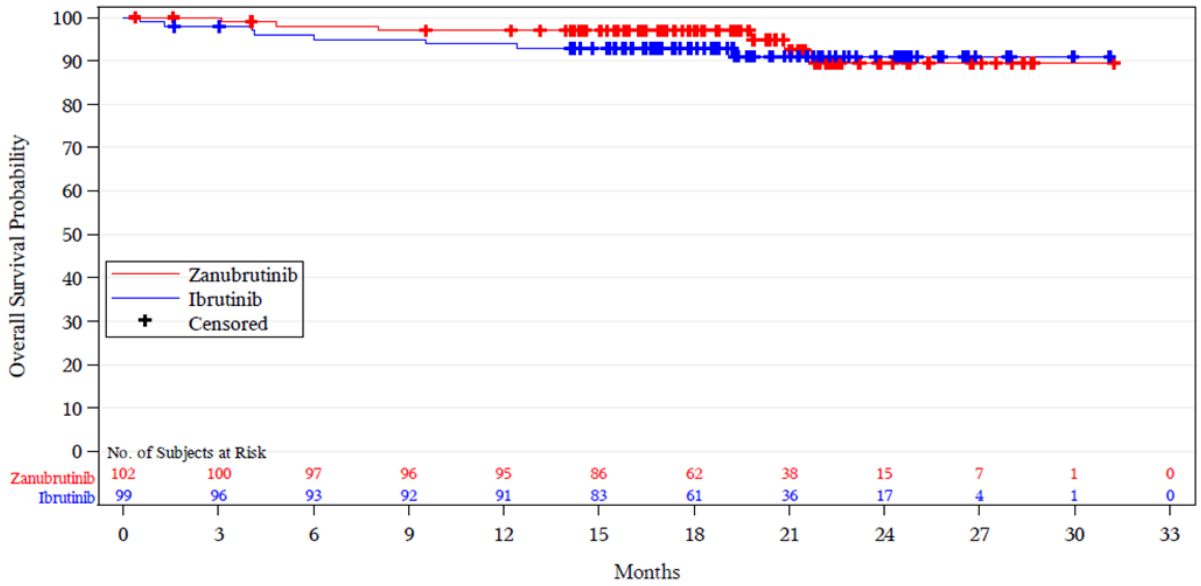
**Table 3.11: OS in Cohort 1 (ITT analysis set)**

| Assessment | Zanubrutinib (n=102) | Ibrutinib (n=99) |
|---|---|---|
| Median follow-up, months (95% CI) | 19.5 (18.1–20.8) | 19.7 (18.7–20.9) |
| Median OS, months (95% CI) | NE (NE–NE) | NE (NE–NE) |
| Event-free rate at, % (95% CI) | | |
| 12 months | 97.0 (90.9, 99.0) | 93.9 (86.8, 97.2) |
| 18 months | 97.0 (90.9, 99.0) | 92.8 (85.5, 96.5) |
| 24 months | 89.5 (76.4, 95.5) | 91.0 (82.5, 95.5) |

Source: CS, Table B.2.12, page 39.[4]
CI = confidence interval; ITT = intention-to-treat; N = number of patients evaluable; NE = not evaluable; OS = overall survival

OS at 12 months was 97.0% among patients treated with zanubrutinib and 93.9% among patients treated with ibrutinib (Figure 3.2).[24]

**Figure 3.2: OS in Cohort 1 (ITT analysis set)**




Source: Response to clarification, Question A12, Figure 2.[6]
No. = number; OS = overall survival

### 3.2.5.6 Other exploratory endpoints

Serum IgM levels decreased over time for patients in both the zanubrutinib and ibrutinib treatment arms (79%, interquartile range [IQR] 88–63 versus 72%, IQR 86–58).[2] Zanubrutinib demonstrated greater and more sustained reductions in IgM by both the repeated-measured mixed-effect model comparing the IgM reductions over time (p=0.0314) and AUC (p=0.0370) compared with ibrutinib (See also CS, Figure B.2.6, page 40).[2, 24]

Zanubrutinib and ibrutinib demonstrated similar improvements in quality of life (QoL) from baseline, with notable improvements in EQ-5D-5L and EORTC QLQ-C30 seen for loss of appetite, fatigue (mean decrease ~30%), physical (mean change from baseline >10%) and role functioning (mean increase from baseline ~20%), and dyspnoea (mean decrease >30%; See also CS, Figures B.2.7 and B.2.9, pages 40-41).[24]

The median times to initiation of non-protocol anti-cancer therapy were 6.83 months in the zanubrutinib treatment arm and 6.44 months in the ibrutinib treatment arm.[24]

## 3.2.6 Updated efficacy results of the ASPEN trial (2020 data)

As part of the response to clarification the company provided data based on a follow-up analysis of safety and efficacy conducted with the cut-off date of 31 August 2020. These results are reported below.

Overall, these data represent 12 months of additional follow-up from the initial data cut-off date (31 August 2019). The median follow-up times on study for patients in Cohort 1, treated with ibrutinib and zanubrutinib, were 31.24 months and 30.78 months, respectively.

A total of [redacted] patients were continuing study treatment: ([redacted] patients on the ibrutinib arm and [redacted] patients on the zanubrutinib arm) as of the 2020 data cut-off date. The most common reasons for treatment discontinuation for patients in Cohort 1 were adverse events ([redacted] ibrutinib-treated patients versus [redacted] zanubrutinib-treated patients) and progressive disease ([redacted] ibrutinib-treated patients versus [redacted] zanubrutinib-treated patients).

Patient disposition for Cohort 1 (ITT analysis set, 2020 data) are summarised in Table 3.12 below.

**Table 3.12: Patient disposition: Cohort 1 (ITT analysis set, 2020 data)**

| Category | Zanubrutinib (n=102) | Ibrutinib (n=99) | Total (n=201) |
|---|---|---|---|
| Randomised, not treated, n (%) | [redacted] | [redacted] | [redacted] |
| AE | [redacted] | [redacted] | [redacted] |
| Progressive disease | [redacted] | [redacted] | [redacted] |
| Treated, n (%) | [redacted] | [redacted] | [redacted] |
| On treatment, n (%) | [redacted] | [redacted] | [redacted] |
| Discontinued, n (%) | [redacted] | [redacted] | [redacted] |
| AE | [redacted] | [redacted] | [redacted] |
| Progressive disease | [redacted] | [redacted] | [redacted] |
| Investigator's discretion | [redacted] | [redacted] | [redacted] |
| Withdrawal by patient | [redacted] | [redacted] | [redacted] |
| Other | [redacted] | [redacted] | [redacted] |
| Median study follow-up (months) | [redacted] | [redacted] | [redacted] |
| Min, Max | [redacted] | [redacted] | [redacted] |

Source: Response to Clarification, Question A10.[6]
AE = adverse event; ITT = intention-to-treat; n = number of patients in the category; N = number of patients evaluable

While close to a [redacted] of patients in Cohort 1 missed at least one efficacy assessment ([redacted] in the ibrutinib group, and [redacted] in the zanubrutinib), most patients ([redacted]) missed assessments at a single visit only.

Based on a review of the study data associated with the missed visits, including safety assessments, no important protocol deviations were identified. The impact of missed central lab IgM assessments was minimized by allowing response assessments by the investigator using the local lab IgM results as available across the study, and there was no significant impact to the study efficacy assessments. Taken together, no major impact to study conduct and no impact to study conclusion related to COVID-19 were observed as of the data cut-off date of 31 August 2020 according to the company.[26]

### 3.2.6.1 IRC-assessed VGPR/CR rate (primary endpoint) and secondary endpoints

All 2020 efficacy analyses presented by the company are based on assessments by the investigator.[26] Therefore, IRC-assessed VGPR/CR rate (primary endpoint), IRC-assessed duration of response, IRC-assessed progression-free survival and IRC-assessed time to response (secondary endpoints) are not reported for the data cut-off date of 31 August 2020.

### 3.2.6.2 Overall survival (exploratory endpoint, 2020 data)

In Cohort 1, the median follow-up time was 31.7 months for patients on the ibrutinib arm and 31.5 months for patients on the zanubrutinib arm. The median overall survival had not been reached overall or in treatment naïve patients or patients with relapsed/refractory disease in either treatment arm as of the data cut-off date (Table 3.13 and Figure 3.3).

A total of [redacted] deaths occurred as of the data cut-off date: [redacted] deaths on each of the ibrutinib and zanubrutinib arms. The event-free rate for patients overall on the zanubrutinib arm versus the ibrutinib arm was [redacted] versus [redacted] at 12 months, [redacted] versus [redacted] at 18 months, [redacted] versus [redacted] at 24 months, and [redacted] versus [redacted] at 30 months. Among patients with relapsed/refractory disease, the event-free rate was higher for patients on the zanubrutinib arm versus those on the ibrutinib arm at 12 months ([redacted]) through 30 months ([redacted]).

**Table 3.13: OS in Cohort 1 (ITT analysis set)**

| Assessment | Zanubrutinib (n=102) | Ibrutinib (n=99) |
|---|---|---|
| Median follow-up, months (95% CI) | [redacted] | [redacted] |
| Median OS, months (95% CI) | [redacted] | [redacted] |
| Event-free rate at, % (95% CI) | | |
| 12 months | [redacted] | [redacted] |
| 18 months | [redacted] | [redacted] |
| 24 months | [redacted] | [redacted] |
| 30 months | [redacted] | [redacted] |

Source: Response to Clarification, Question A10.[6]
CI = confidence interval; ITT = intention-to-treat; N = number of patients evaluable; NE = not evaluable; OS = overall survival

OS at 30 months was [redacted] among patients treated with zanubrutinib and [redacted] among patients treated with ibrutinib (Figure 3.2).[24]

**Figure 3.3: OS in Cohort 1 (ITT analysis set, 2020 data)**

[redacted]

Source: Response to Clarification, Question A10.[6]
No. = number; OS = overall survival

### 3.2.6.3 Other exploratory endpoints

[redacted] were not reported for the data cut-off date of 31 August 2020.

#### 3.2.6.4 Adverse events

[redacted] were also not reported for the data cut-off date of 31 August 2020 in the document provided by the company.[26]

#### 3.2.6.5 Conclusion

All 2020 efficacy analyses presented by the company are [redacted].[26]

Comparing the 2020 results with 2019 results reported in the CSR, it looks like the 2020 results are in line with the 2019 results. However, this cannot be verified for the outcomes reported in the CS.

### 3.2.7 Adverse events

#### 3.2.7.1 Adverse events in the ASPEN trial

The CS reported adverse events (AEs) that occurred in the zanubrutinib and comparator group (ibrutinib) in the ASPEN trial (CS, page 54-68) over follow-up periods lasting from 16.4 months (zanubrutinib Cohort 2) and 18.7 months (Cohort 1) (CS, page 54 – see also Table 3.14 below). The CS reported that overall, the percentages of patients with any AEs (zanubrutinib 94.6% v ibrutinib 99%), serious AEs (zanubrutinib 39.5% v ibrutinib 40.8%), and treatment related AEs (79.1% versus 85.7%) were similar in both groups. Only the combined AEs were >10% higher in Cohort 1 compared with Cohort 2 of the zanubrutinib groups (97% versus 85.7%).

**Table 3.14: Overview of AEs (safety analysis set)**

| Event | Zanubrutinib | | | Ibrutinib (n=98) |
|---|---|---|---|---|
| | Cohort 1 (n=101) | Cohort 2 (n=28) | Total (n=129) | |
| AEs, n (%) | 98 (97.0) | 24 (85.7) | 122 (94.6) | 97 (99.0) |
| Grade ≥3 | 59 (58.4) | 18 (64.3) | 77 (59.7) | 62 (63.3) |
| SAEs | 40 (39.6) | 11 (39.3) | 51 (39.5) | 40 (40.8) |
| AEs leading to death | 1 (1.0) | 0 (0.0) | 1 (0.8) | 4 (4.1) |
| AEs leading to discontinuation | 4 (4.0) | 2 (7.1) | 6 (4.7) | 9 (9.2) |
| TRAEs, n (%) | 80 (79.2) | 22 (78.6) | 102 (79.1) | 84 (85.7) |
| Grade ≥3 | 33 (32.7) | 13 (46.4) | 46 (35.7) | 42 (42.9) |
| AESIs, n (%) | 86 (85.1) | 23 (82.1) | 109 (84.5) | 81 (82.7) |

Source: CS, Table B.2.21, page 54.[4]
AE = adverse event; AESI = adverse event of special interest; n = number of patients in the category; N = number of patients evaluable; SAE = serious adverse event; TRAE = treatment-related adverse event

AEs which were most frequently reported by system organ class (SOCs) and preferred term (PT) with a frequency of >10% in any group were described in the CS (page 55, Table B.2.22,[4] and Table 3.15 below). Most of these were comparable across groups. Some were more prevalent in the ibrutinib arm compared with the combined zanubrutinib arms, including muscle spasms (23.5% versus 10.9%), and atrial fibrillation (14.3% versus 2.3%). The AE that was more prevalent (>10% higher) in the zanubrutinib arms compared with the ibrutinib treatment arm was neutropenia (22.5% versus 12.2%).

**Table 3.15: AEs by SOC and PT reported in >10% of patients (safety analysis set)**

| Event | Zanubrutinib | | | Ibrutinib (n=98) |
|---|---|---|---|---|
| | Cohort 1 (n=101) | Cohort 2 (n=28) | Total (n=129) | |
| AEs, n (%) | 98 (97.0) | 24 (85.7) | 122 (94.6) | 97 (99.0) |
| Infections and infestations | | | | |
| Upper respiratory tract infection | 24 (23.8) | 6 (21.4) | 30 (23.3) | 28 (28.6) |
| Urinary tract infection | 10 (9.9) | 4 (14.3) | 14 (10.9) | 10 (10.2) |
| Nasopharyngitis | 11 (10.9) | 2 (7.1) | 13 (10.1) | 7 (7.1) |
| Pneumonia | 2 (2.0) | 4 (14.3) | 6 (4.7) | 12 (12.2) |
| Gastrointestinal disorders | | | | |
| Diarrhoea | 21 (20.8) | 8 (28.6) | 29 (22.5) | 31 (31.6) |
| Constipation | 16 (15.8) | 4 (14.3) | 20 (15.5) | 7 (7.1) |
| Nausea | 15 (14.9) | 1 (3.6) | 16 (12.4) | 13 (13.3) |
| Vomiting | 9 (8.9) | 2 (7.1) | 11 (8.5) | 13 (13.3) |
| Blood and lymphatic system disorders | | | | |
| Neutropenia | 25 (24.8) | 4 (14.3) | 29 (22.5) | 12 (12.2) |
| Anaemia | 12 (11.9) | 6 (21.4) | 18 (14.0) | 10 (10.2) |
| Thrombocytopenia | 10 (9.9) | 3 (10.7) | 13 (10.1) | 10 (10.2) |
| General disorders and administration site conditions | | | | |
| Fatigue | 19 (18.8) | 4 (14.3) | 23 (17.8) | 15 (15.3) |
| Pyrexia | 13 (12.9) | 6 (21.4) | 19 (14.7) | 12 (12.2) |
| Oedema peripheral | 9 (8.9) | 4 (14.3) | 13 (10.1) | 19 (19.4) |
| Injury, poisoning and procedural complications | | | | |
| Contusion | 13 (12.9) | 6 (21.4) | 19 (14.7) | 23 (23.5) |
| Musculoskeletal and connective tissue disorders | | | | |
| Back pain | 14 (13.9) | 4 (14.3) | 18 (14.0) | 6 (6.1) |
| Arthralgia | 13 (12.9) | 3 (10.7) | 16 (12.4) | 16 (16.3) |
| Pain in extremity | 11 (10.9) | 1 (3.6) | 12 (9.3) | 7 (7.1) |
| Muscle spasms | 10 (9.9) | 4 (14.3) | 14 (10.9) | 23 (23.5) |
| Respiratory, thoracic and mediastinal disorders | | | | |
| Cough | 13 (12.9) | 5 (17.9) | 18 (14.0) | 17 (17.3) |
| Dyspnoea | 14 (13.9) | 1 (3.6) | 15 (11.6) | 6 (6.1) |
| Epistaxis | 13 (12.9) | 1 (3.6) | 14 (10.9) | 19 (19.4) |
| Nervous system disorders | | | | |
| Headache | 15 (14.9) | 3 (10.7) | 18 (14.0) | 11 (11.2) |
| Dizziness | 13 (12.9) | 1 (3.6) | 14 (10.9) | 9 (9.2) |
| Skin and subcutaneous tissue disorders | | | | |
| Rash | 13 (12.9) | 3 (10.7) | 16 (12.4) | 16 (16.3) |
| Pruritus | 9 (8.9) | 4 (14.3) | 13 (10.1) | 5 (5.1) |
| Vascular disorders | | | | |

| | | | | |
|---|---|---|---|---|
| Hypertension | 11 (10.9) | 3 (10.7) | 14 (10.9) | 16 (16.3) |
| Renal and urinary disorders | | | | |
| Haematuria | 7 (6.9) | 1 (3.6) | 8 (6.2) | 10 (10.2) |
| Cardiac disorders | | | | |
| Atrial fibrillation | 2 (2.0) | 1 (3.6) | 3 (2.3) | 14 (14.3) |
| Source: CS, Table B.2.22, page 55.[4]<br>AE = adverse event; n = number of patients in the category; N = number of patients evaluable; PT = preferred term; SOC = system organ class | | | | |

The CS summarised serious AEs (Grade ≥3) by SOC and PT in Table B.2.23 (page 57).[4] The Grade ≥3 AE more prevalent (>5% higher) in the zanubrutinib arms compared with the ibrutinib treatment arm were neutropenia (14.7% versus 8.2%). Grade ≥3 AEs more prevalent in the ibrutinib arm compared with the zanubrutinib arm were pneumonia (7.1% versus 1.0%) and hypertension (11.2% versus 5.9).

The CS reported treatment related AEs (TRAEs) arising in >5% of participants (page 57-8 and Table B.2.24).[4] These occurred in 79.2% of the participants in Cohort 1, 78.6% of participants in Cohort 2, and 85.7% of participants in the ibrutinib group, and the results were comparable to all AEs, with neutropenia standing out as being higher in the zanubrutinib groups (19.4% versus 11.2%).

Serious adverse events (SAEs) were reported on page 59 and Table B.2.26.[4] At least one patient had a SAE in 39.6% of participants in Cohort 1, 39.3% of participants in Cohort 2, and 40.8% of participants in the ibrutinib group. Several of these arose in the zanubrutinib but not the ibrutinib groups: febrile neutropenia (n=3), neutropenia (n=3), anaemia (n=2), thrombocytopenia (n=2), lower respiratory tract infection (n=3), drug withdrawal syndrome (n=2), periorbital haematoma (n=2), subdural haemorrhage (n=2), basal cell carcinoma (n=2), and respiratory failure (n=2). Others arose in the ibrutinib but not zanubrutinib groups: pericarditis (n=2), cholecystitis (n=2), and loss of consciousness (n=2).

The deaths (any cause) were similar in all groups: 5.9% in Cohort 1, 10.7% in Cohort 2, and 7.1% in the ibrutinib group. Fatal AEs are summarised in Table B.2.28 of the CS (CS, page 61[4]). Four deaths due to AEs in the ibrutinib arm were due to cause unknown, acute cardiac failure, bacterial sepsis and sepsis; the single death due to an AE in the zanubrutinib arm was caused by cardiomegaly.

AEs of special interest were also reported in the CS on pages 62 to 67.[4] A higher proportion of patients treated with ibrutinib experienced haemorrhage compared with zanubrutinib (59.2% versus 48.5%). A higher rate of atrial fibrillation was detected in the ibrutinib (14.3%) than the zanubrutinib (2.3% groups. The cytopenia events also differed by >10%). Within the category of any grade cytopenias, more patients in the zanubrutinib groups had neutropenia (27.1% versus 13.3%). Within the category of any cytopenias (Grade ≥3)**,** more patients in the zanubrutinib groups had neutropenia (17.8% versus 8.2%). The other AEs of special interest did not differ >10%: major haemorrhages, hypertension, secondary primary malignancy, infections.

The CS reported AEs leading to treatment discontinuation by SOC and PT on pages 67-8 and Table B.2.36.[4] In the ibrutinib group, 9.2% of participants discontinued treatment due to AEs compared with 4.7% in the combined zanubrutinib groups. The incidences of AEs leading to discontinuation of study drug was similar between Cohort 1 and Cohort 2 (4% and 7.1%).

Possible cardiac serious AEs may not have been measured due to the inclusion criteria of the ASPEN trial. As stated in the Summary of Product Characteristics (SmPC) for zanubrutinib: “patients with severe cardiovascular disease were excluded from [ibrutinib] clinical studies”.[7] This is because atrial

fibrillation and atrial flutter were reported in some patients with cardiac risk factors,, hypertension, acute infections, and a previous history of atrial fibrillation.[7] The company stated in the response to clarification that since they did not want to expose patients with underlying comorbidities to known or unknown side-effects, patients with cardiovascular disease were excluded from the ASPEN trial.

### 3.2.7.2 Pooled adverse events

The CS reported pooled safety data for all Waldenström's macrobulinaemia (WM) patients (n=253) across four studies (the ASPEN trial, BGB-3111-210, BGB-3111-AU-003, and BGB-3111-1002) (no comparison group, CS pages 68-77,[4] and Table 3.16 below). The median duration of exposure in all WM patients was [redacted]. A summary of AEs is included in Table B.2.37 in the CS[4] (Table 3.16 below). [redacted] reported one AE, [redacted] reported at least one serious AE, [redacted] had an AE leading to discontinuation, and [redacted]. The only event that led to treatment discontinuation in >1 patient was [redacted].

**Table 3.16: Overview of AEs for all WM patients (four studies)**

| Event | All WM (n=253) |
|---|---|
| AEs, n (%) | [redacted] |
| Grade ≥3 | [redacted] |
| SAEs | [redacted] |
| AEs leading to death | [redacted] |
| AEs leading to discontinuation | [redacted] |
| TRAEs, n (%) | [redacted] |
| AESIs, n (%) | [redacted] |
| Grade ≥3 | [redacted] |
| Serious | [redacted] |

Source: CS, Table B.2.37, page 69.[4]
AE = adverse event; AESI = adverse event of special interest; n = number of patients in the category; N = number of patients evaluable; SAE = serious adverse event; TRAE = treatment-related adverse event; WM = Waldenström's macroglobulinaemia

The CS reported the most common AEs (CS, pages 70-71), with [redacted]. The most frequent (>15%) AEs in the All WM group were: [redacted] TRAEs were reported in the CS on page 71, with the most common (>10%) being [redacted]

The most frequent SAEs were [redacted]. The CS does not highlight [redacted] in their summary of the SAEs. Also, the CS does not report that neoplasms benign, malignant (including cysts and polyps) were [redacted]; the CS only includes one row below this general category 'basal cell carcinoma' which affected [redacted].

At least one adverse events of special interest (AESIs) were reported by [redacted] of zanubrutinib-treated patients. AEs within the categories of [redacted] were reported most frequently. Events that met the criteria for seriousness and/or were Grade ≥3 were reported in [redacted] and [redacted] of patients, respectively.

The CS also reports that a total of [redacted] reported at least one occurrence of treatment-emergent anaemia. [redacted] patients with treatment-emergent anaemia received red blood cell transfusion within 30 days of onset.

#### 3.2.7.3 Safety conclusions

Overall, zanubrutinib has a comparable safety and tolerability profile compared with ibrutinib. Neutropenia was consistently more prevalent in the zanubrutinib groups compared with the ibrutinib group ([redacted]). The zanubrutinib treated patients had a lower rate of several AEs compared with ibrutinib, such as [redacted]. There were also fewer AEs leading to [redacted] with zanubrutinib compared with ibrutinib.

In a pooled analysis of 253 patients with WM, the most common AEs reported by zanubrutinib treated patents were [redacted].

### 3.2.8 Included studies: Supporting evidence

Two studies providing supporting evidence were highlighted in the CS. Neither was included in the indirect comparisons or the cost effectiveness analysis.

The first study was Cohort 2 from the ASPEN study, which included 28 patients with *MYD88*WT, all of whom were assigned to zanubrutinib.[24, 27, 28] The exploratory objective of this study was to assess the anti-cancer activity and safety of zanubrutinib. The population, intervention, and outcomes were the same as for Cohort 1 (see Section 3.2.1 to 3.2.5). To be eligible patients needed to have a histologic diagnosis of WM, meet at least one criterion for treatment initiation, and if treatment naïve must be considered unsuitable for standard chemo-immunotherapy, and not have received prior BTK inhibitors. All patients in this cohort received zanubrutinib 160 mg (80 mg x 2) orally BID with at least eight hours between doses. The outcomes were the same as it was for Cohort 1 (see Section 3.2.2) and included CR or VGPR. The median age of the patients in Cohort 2 was 72. Half were male, and all but one were white. Most (71.4%) had received between one and three prior lines of therapy; 23 of the 28 patients in this arm of the ASPEN study were relapses/refractory, and the median follow-up time was 17.87 months (17.15 months for relapsed/refractory patients). As of the data cut-off date (31 August 2019), 39.3% had discontinued treatment, mostly due to progressive disease. Seven patients in this cohort (26.9%) had either CR or VGPR; 85.7% of patients in Cohort 2 experienced at least one adverse reaction, 39.5% experienced at least one serious adverse event, and there were three deaths (10.7%).

The second study was the Phase 1/2 BGB-3111-AU-003 study (see CS, Section B.2.6.2;[4] and see Table 3.17 below). This was an open-label, multiple-dose, multicentre dose escalation and expansion study in patients with B-cell lymphoid malignancies, including but not limited to WM. Adult patients with various B-cell malignancies were enrolled in this trial. In Part 1 of this study, sequentially enrolled patients received zanubrutinib doses of between 40 mg and 320 mg orally once/day or 160 mg twice/day. In Part 2 of the study, patients with relapsed/refractory WM were assigned to different groups by alternate allocation; patients with treatment-naïve WM were assigned to a separate group. Patients in Part 2 received 160 mg of zanubrutinib twice/day or 320 mg once/day orally. The primary endpoint was either CG or VGPR. Seventy-eight patients were enrolled in this study, 24 were treatment-naïve, 16% were over 75 years old, and 85% were white. The CS reports that at a median follow-up of 30.3 months, overall response rate was 95.9% and rates of VGPR/CR increased with prolonged treatment

from 20.5% at 6 months, to 32.9% at 12 months and 43.8% at 24 months.[29, 30] In total, nine patients died with five deaths due to AEs, two due to progressive disease (PD) and two due to unknown causes.

**Table 3.17: Study BGB-3111-AU-003**

| | |
|---|---|
| **Study** | Study BGB-3111-AU-003 (NCT02343120)[29] |
| **Study design** | Multicentre, Phase 1/2 trial |
| **Population** | Patients with WM who are relapsed/refractory or treatment naïve |
| **Intervention(s)** | Zanubrutinib 40 mg OD, 80 mg OD, 160 mg OD, 320 mg OD or 160 mg BID |
| **Comparator(s)** | N/A |
| **Reported outcomes specified in the decision problem** | Response rate<br>OS<br>DOR<br>PFS |
| **All other reported outcomes** | N/A |

Source: CS, B.2.1, pages 23.[4]
BID = twice daily; DOR = duration of response; OD = once daily; OS = overall survival; PFS = progression-free survival; WM = Waldenström's macroglobulinaemia.

### 3.2.9 Ongoing studies

According to the company, "there are no additional ongoing studies due to provide additional evidence in the next 12 months for relapsed/refractory or treatment-naïve WM" (CS, Section B.2.11, page 77).[4]

## *3.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison*

The company undertook a systematic literature review (SLR) to identify published evidence for current and future treatment options for patients with WM, detailed in Appendix D of the CS. Specifically, the SLR aimed to identify the efficacy of treatment options for patients with WM and published evidence for current and future treatment options for patients with WM, as well as the safety and tolerability of treatment options for patients with WM. This SLR did not include any search terms for autologous SCT (see also Section 2.3 of this report).

According to the flow diagram presented in Appendix D of the CS (CS, Appendix D, Figure B.5.1, page 20), 34 publications for the treatment of WM were identified through the SLR.

### 3.3.1 Comparator data

The SLR identified four articles for bendamustine and rituximab (BR), five articles for dexamethasone, rituximab and cyclophosphamide (DRC), six articles for fludarabine, cyclophosphamide and rituximab (FCR) or fludarabine and rituximab (FR), two articles for chlorambucil monotherapy and six articles for rituximab monotherapy. No studies were identified for cladribine and rituximab (Clad-R) or best supportive care (BSC). All identified studies were single-arm studies, except one retrospective study investigating both BR and DRC.[31] This adds up to 23 included studies: the company does not explain the difference between 34 included publications and 23 included studies.

Due to the lack of RCTs comparing zanubrutinib with any comparator other than ibrutinib, and a lack of common comparators for an anchored indirect treatment comparison (ITC), the company conducted a matching adjusted indirect comparison (MAIC). The company assessed the included studies for inclusion in the MAIC, on the definition of patient populations, availability of progression-free survival (PFS) and overall survival (OS) Kaplan-Meier (KM) curves to inform the cost effectiveness analysis, availability and comparability of baseline patient characteristics, study design, sample size, and geographical location. Only PFS, OS, and AEs were considered as outcomes in the MAIC.

In the company submission, the company included two single arm trials in the MAIC, one for BR[32] and another for DRC.[33, 34] These studies were considered by the company to be the best trials from those of BR and DRC based on the assessment above. The company stated that all identified studies for FCR and FR could not be included in the MAIC because of the relatively small sample sizes and a lack of reporting of OS KM curves or PFS KM curves. However, in the clarification response, the company performed further MAICs using data from additional studies.

The population in the study of chlorambucil monotherapy identified from the company submission SLR was adults with WM with unknown prior treatment history. The population in the study of rituximab monotherapy identified from the company submission SLR included both treatment-naïve (suitable for chemo-immunotherapy) and relapsed/refractory patients. The company considered it unfeasible to match the treatment naïve (unsuitable for chemo-immunotherapy) patients in ASPEN to these studies given the small number of such patients in the zanubrutinib arm (n=19), however, MAICs were performed for one of each treatment in the clarification response. The populations in the trials included in the clarification response exploratory MAICs were varied, including both treatment naïve and treatment refractory patients, these results should be interpreted with caution.

In the company submission, the only comparisons in the MAICs were between zanubrutinib, BR and DRC. The aim of the MAICs were therefore to compare outcomes between zanubrutinib, BR and DRC, reducing the potential for bias by adjusting for confounding variables. Given the relative scarcity of data, the MAICs were necessary for any comparisons between zanubrutinib and treatments other than ibrutinib. Individual participant data (IPD) was only available for the zanubrutinib arm of APSEN: summary data from the included trials were used for the BR and DRC trials. In the clarification response, the company performed MAICs for the following treatments: BR (three additional analyses), DRC (three additional analyses), fludarabine, cyclophosphamide and rituximab or fludarabine and rituximab (FCR/FR, four analyses), chlorambucil (one analysis), and rituximab (one analysis).

### 3.3.2 MAIC methodology in the company submission

Three MAICs were conducted, comparing the overall zanubrutinib population of ASPEN (n=102) with the populations the BR (n=71) and DRC (n=72) separately, and also for patients with relapsed/refractory disease in the zanubrutinib arm of ASPEN (n=83) to the BR population, as the BR population only included relapsed/refractory patients.

In each MAIC, the matching algorithm proposed by Signorovitch et al.[35] was used to reweight zanubrutinib patients in ASPEN so that the mean baseline characteristics of specified matching variables matched those reported in the comparator trial. The matching variables considered by the company comprised: age, number of prior therapies, Eastern Cooperative Oncology Group (ECOG) performance status, MYD88/CXCR4 mutation status, IgM concentration, β2-microglobulin concentration, platelet count, haemoglobin concentration, presence of extramedullary disease, and International Prognostic Scoring System for Waldenströms macroglobulinemia (WM IPSS). The matching variables were converted to binary or categorical variables, depending on the data availability

in the comparator trial. Table 3.18 shows an overview of MAIC studies, including data on the matching variables.

**Table 3.18: Overview of MAIC studies in the company submission**

| | **Zanubrutinib, ASPEN[2]** | **BR, Tedeschi et al. 2015[32]** | **DRC, Dimopoulos et al. 2007 / Kastritis et al. 2015[33, 34]** |
|---|---|---|---|
| IPD available | Yes | No | No |
| **Study characteristics** | | | |
| Study design | Multicentre, phase 3 | Multicentre (phase not applicable) | Multicentre, phase 2 |
| Country | Europe (59.7%); Australia or New Zealand (30.8%) | Italy | Greece |
| Intervention | Zanubrutinib<br>(160 mg twice daily until disease progression)<br>(ibrutinib is included in ASPEN but not included in MAIC) | BR<br>(Six 28-day course of bendamustine 50–90 mg/m$^2$ IV on days 1, 2) and rituximab (375 mg/m$^2$ IV on day 1) | DRC<br>(Six 21-day courses of dexamethasone 20 mg IV, followed by rituximab IV 375 mg/m$^2$ and oral cyclophosphamide 100 mg/m$^2$ twice daily [days 1 to 5]) |
| Patient population | Mixed treatment-naïve (unsuitable for chemo-immunotherapy) and relapsed/refractory WM | Relapsed/refractory WM | Treatment-naïve (suitable for chemo-immunotherapy) WM |
| Sample size, N | Treatment naïve: 19<br>Relapsed/refractory: 83 | 71 | 72 |
| Median follow-up | 19.47 months | 19 months | 23.4 months per Dimopoulos et al. 2007; 8 years per Kastritis et al. 2015 |
| **Outcomes of interest** | | | |
| PFS KM | IPD available | Reported | Reported |
| OS KM | IPD available | NR | Reported |
| AE incidence | IPD available | NR | Reported |
| **Patient baseline characteristics available in any comparator trial** | | | |
| **Age, year** | | | |
| Mean (SD) | 69.5 (9.46) | NR | NR |
| Median | 70 | 72 | 69 |
| Range | 38-90 | 49-88 | 33-89 |
| > 65, n (%) | 61 (59.8%) | NR | 63% |
| **Female proportion** | 134 (66.7%) | 25 (35.2%) | 45 (62.5%) |
| **IgM, g/L** | | | |
| Mean (SD) | 34.72 (19.62) | NR | NR |
| Median | 32.85 | 38.15 | NR |

| | Zanubrutinib, ASPEN[2] | BR, Tedeschi et al. 2015[32] | DRC, Dimopoulos et al. 2007 / Kastritis et al. 2015[33, 34] |
|---|---|---|---|
| Range | 2.4-108.0 | 2.4-96.2 | NR |
| **Platelet count, $10^9$/L** | | | |
| Mean (SD) | 238.63 (108.21) | NR | NR |
| Median | 236.00 | NR | NR |
| Range | 34.0-564.0 | NR | NR |
| ≤ 100, n (%) | 12 (11.8%) | NR | 3 (4.2%) |
| **Haemoglobin, g/L** | | | |
| Mean (SD) | 104.39 (19.24) | NR | NR |
| Median | 102.50 | NR | NR |
| Range | 53.0-152.0 | NR | NR |
| <100, n (%) | 78 (47.1%) | NR | 41 (56.9%) |
| **Prior line of treatment** | | | |
| Median | 1 | 2 | N/A |
| Range | 0-3 | 1-5 | N/A |
| 0, n (%) | 19 (18.6) | NR | N/A |
| 1-3, n (%) | 76 (74.5) | NR | N/A |
| >3, n (%) | 7 (6.9) | NR | N/A |
| **Prior treatment regimen, n (%)** | | | |
| Nucleoside analogue-containing therapies | 39 (23.8%) | 21 (29.6%) | N/A |
| Bortezomib-containing therapies | 20 (12.2%) | 7 (9.9%) | N/A |
| Cyclophosphamide-containing therapies | 139 (84.8%) | 64 (90.1%) | N/A |
| Rituximab alone or in combination therapy | 150 (91.5%) | 55 (77.5%) | N/A |
| **Extramedullary disease, n (%)** | | | |
| Adenopathy and/or splenomegaly | 63 (61.8%) | 30 (42.3%) | N/A |
| Lymphadenopathy | 61 (59.8%) | NR | 28 (38.9%) |
| Splenomegaly | 16 (15.7%) | NR | 23 (31.9%) |
| **IPSSWM score, n (%)** | | | |
| Low risk | 17 (16.7%) | 12 (21.4%[a]) | NR |
| Intermediate risk | 38 (37.3%) | 17 (30.4%[a]) | NR |
| High risk | 47 (46.1%) | 27 (48.2%[a]) | NR |

Source: CS, Appendix D, Table B.5.8.[12]

AE = adverse event; BR = bendamustine rituximab; CS = company submission; DRC = dexamethasone, rituximab, and cyclophosphamide; IgM = Immunoglobulin M; IPD = individual patient-level data; ITC = indirect treatment

| | Zanubrutinib, ASPEN[2] | BR, Tedeschi et al. 2015[32] | DRC, Dimopoulos et al. 2007 / Kastritis et al. 2015[33, 34] |
|---|---|---|---|
| comparison; IV = intravenous; KM = Kaplan-Meier; N/A = not applicable; NR = not reported; OS = overall survival; PFS = progression-free survival; SD = standard deviation<br>[a] Based on 56 patients. | | | |

Individual patient-level event and censoring times for OS and PFS were extracted from the KM curves presented for BR and DRC, which were compared with the KM curves created using both unmatched and matched data for zanubrutinib. Survival was compared by estimating hazard ratios (HRs) using Cox proportional hazard (PH) models, using the reconstructed patient data extracted for BR and DRC.

### 3.3.3 MAIC methodology in the clarification response

The company performed 12 additional MAICs in the clarification response using the same methodology as in the company submission including additional studies from Table B.5.7 in Appendix D.[6, 12] Three additional analyses were performed for BR, three additional analyses were performed for DRC, four analyses were performed for FCR/FR, one analysis was performed for chlorambucil, and one analysis was performed for rituximab. The company noted that consideration should be given to differences between ASPEN and the study populations for all MAICs in the clarification response (although this is also true for the MAICs in the company submission). Table 3.19 shows an overview of MAIC studies, including the variables for which ASPEN was matched (all matching variables were binary or categorical).

**Table 3.19: Overview of MAIC studies in the clarification response**

| Study | N | Population | Matching variables | Outcomes |
|---|---|---|---|---|
| **BR** | | | | |
| Paludo (2018) | 60 | Treatment naïve (n=17) & R/R (n=43)<br>Single centre, US | Age, platelet count, haemoglobin, IPSS, IgM, β2 microglobulin | PFS |
| Paludo (2018)* | 43 | R/R (n=43)<br>Single centre, US | Age, platelet count, haemoglobin, IPSS, IgM, β2 microglobulin | PFS |
| Castillo (2018) | 57 | Treatment naïve (n=57)<br>Single centre, US | Age, platelet count, haemoglobin, IPSS, IgM, β2 microglobulin | PFS, OS |
| **DRC** | | | | |
| Paludo (2017/18) | 100 | Treatment naïve (n=50) & R/R (n=50)<br>Single/multiple centres, US | Age, platelet count, haemoglobin, IPSS, IgM, β2 microglobulin | PFS |
| Paludo (2017/18)* | 50 | R/R (n=50)<br>Single/multiple centres, US | Age, platelet count, haemoglobin, IPSS, IgM, β2 microglobulin | PFS |
| Castillo (2018) | 38 | Treatment naïve (n=38)<br>Single centre, US | Age, platelet count, haemoglobin, IgM, β2 microglobulin | PFS, OS |
| **FCR/FR** | | | | |
| Treon (2009) | 43 | Treatment naïve (n=27) & R/R (n=16) | Age, platelet count, no prior therapy, IgM | PFS |

| Study | N | Population | Matching variables | Outcomes |
|---|---|---|---|---|
| | | Multiple centres, multi-national | | |
| Tedeschi (2012) | 43 | Treatment naïve (n=28) & R/R (n=15)<br>Multiple centres, Italy | Age, platelet count, haemoglobin, IPSS, IgM, adenopathy/splenomegaly/extranodal involvement | OS |
| Souchet (2016) | 82 | Treatment naïve (n=25) & R/R (n=57)<br>Multiple centres, France | Age, platelet count, haemoglobin, prior therapy, β2 microglobulin | PFS |
| Souchet (2016)* | 57 | R/R (n=57)<br>Multiple centres, France | Age, platelet count, haemoglobin, prior therapy, β2 microglobulin | PFS |
| **Chlorambucil** | | | | |
| Kyle (2000) | 46 | Unknown prior treatment history (n=46)<br>Single centre, US (assumed) | Age, platelet count, haemoglobin, lymphadenopathy | PFS |
| **Rituximab** | | | | |
| Gertz 2004/2009 | 69 | Treatment naïve (n=34) & R/R (n=35)<br>Multiple centres, US | Age, platelet count, haemoglobin, IgM, β2 microglobulin, ECOG | PFS, OS |

Source: company clarification response.[6]
BR = bendamustine-rituximab; DRC = dexamethasone, rituximab and cyclophosphamide; ECOG = Eastern Cooperative Oncology Group; FCR/FR = fludarabine, cyclophosphamide and rituximab or fludarabine and rituximab; IgM = immunoglobulin M; IPSS = International Prognostic Scoring System; OS = overall survival; PFS = progression-free survival; R/R = relapsed/refractory, US = United States of America
*Compared with the R/R analysis set (n=83) of ASPEN

Not all studies from Table B.5.7 in Appendix D of the company submission were included in the MAICs in the clarification response. Tam 2005[36] and Ngan 2003[37] for FCR/FR were not included due to the extremely low sample size of the WM population (3 and 5, respectively). Tedeschi 2013[38] for FCR/FR was not included due to a lack of survival outcomes for the overall population but only for treatment responders. Ngan 2003 for FCR/FR was not included due to the lack of reporting of any survival KM curves. Dimopoulos 2002a,[39] Dimopoulos 2002b,[40] Byrd 1999,[41] and Treon 2001[42] for rituximab monotherapy were also excluded due to the lack of reporting of any survival KM curves.

The MAIC results for the same treatment from different studies could not easily be combined; as such, only results from the individual MAICs are presented.

### 3.3.4 MAIC results in the company submission

Table 3.20 shows the results for survival from the MAICs in the company submission after matching, comparing zanubrutinib with BR and DRC, where a HR less than 1 indicates zanubrutinib improves survival. It is unclear whether the effect estimate for BR is from a comparison with all patients in ASPEN, or just the relapsed/refractory patients, although the survival curves were similar.

**Table 3.20: MAIC survival results after matching in the company submission**

| | BR | DRC |
|---|---|---|
| PFS, HR (95% CI) | [redacted] | [redacted] |
| OS, HR (95% CI) | [redacted] | [redacted] |

Source: CS, Section B.2.9.3.[4]
BR = bendamustine rituximab; CI = confidence interval; DRC = dexamethasone, rituximab, and cyclophosphamide; OS = overall survival; PFS = progression-free survival

### 3.3.5 MAIC results in the clarification response

Table 3.21 shows the results for survival from the MAICs in the clarification response after matching, comparing zanubrutinib with all other treatments, where a HR less than 1 indicates zanubrutinib improves survival. Studies with an asterisk were compared with the relapsed/refractory patients in ASPEN only, as opposed to all patients in ASPEN.

**Table 3.21: MAIC survival results after matching in the clarification response**

| Outcome | Study | N (N naïve, N R/R) | HR (95% CI) |
|---|---|---|---|
| PFS | **BR** | | |
| | Paludo (2018) | 60 (17, 43) | 0.200 (0.076 to 0.524) |
| | Paludo (2018)* | 43 (0, 43) | 0.103 (0.028 to 0.381) |
| | Castillo (2018) | 57 (57, 0) | 3.690 (0.797 to 17.079) |
| | **DRC** | | |
| | Paludo (2017/18) | 100 (50, 50) | 0.198 (0.077 to 0.512) |
| | Paludo (2017/18)* | 50 (0, 50) | 0.134 (0.044 to 0.413) |
| | Castillo (2018) | 38 (38, 0) | 0.328 (0.088 to 1.221) |
| | **FCR/FR** | | |
| | Treon (2009) | 43 (27, 16) | 0.588 (0.194 to 1.785) |
| | Souchet (2016) | 82 (25, 57) | 0.811 (0.225 to 2.919) |
| | Souchet (2016)* | 57 (0, 57) | 0.744 (0.135 to 4.101) |
| | **Chlorambucil** | | |
| | Kyle (2000) | 46 (NR, NR) | 0.206 (0.055 to 0.775) |
| | **Rituximab** | | |
| | Gertz (2004/2009) | 69 (34, 35) | 0.237 (0.071 to 0.793) |
| OS | **BR** | | |
| | Castillo (2018) | 71 (0, 71) | 4.429 (0.678 to 28.906) |
| | **DRC** | | |
| | Castillo (2018) | 72 (72, 0) | 0.905 (0.083 to 9.916) |
| | **FCR/FR** | | |
| | Tedeschi (2012) | 43 (28, 15) | 0.397 (0.105 to 1.504) |
| | **Rituximab** | | |
| | Gertz (2004/2009) | 69 (34, 35) | 0.232 (0.070 to 0.763) |

Source: company clarification response.[6]
BR = bendamustine-rituximab; DRC = dexamethasone, rituximab and cyclophosphamide; FCR/FR = fludarabine, cyclophosphamide and rituximab or fludarabine and rituximab; OS = overall survival; PFS = progression-free survival; R/R = relapsed/refractory

*Compared with the R/R analysis set (n=83) of ASPEN

### *3.4 Critique of the indirect comparison and/or multiple treatment comparison*

The MAICs were conducted using single-arm studies to compare zanubrutinib with comparator treatments other than ibrutinib, given the lack of direct RCT evidence. In the company submission, only BR and DRC were considered as comparators, though in the clarification response, FCR/FR, chlorambucil and rituximab were also considered. Given the lack of direct evidence, indirect methods were used, and given the lack of RCT evidence to inform a network meta-analysis, MAICs using single-arm studies represent the best available evidence comparing zanubrutinib with comparator treatments. Even so, these analyses still present a substantial risk of bias. The company updated the key assumptions of the MAIC in updated Table B.2.19 (Clarification response, Question A25[6]), which is relevant for this critique.

MAICs allow for patients in different studies to be made more comparable by matching on (or adjusting for) variables thought to cause bias when comparing study results. The degree to which bias is reduced is typically unknown, and given the observational nature of MAIC analyses, it is impossible to state definitively whether any individual analysis is free from bias. However, it is possible to judge whether any MAIC is likely to have residual bias, given the populations used in each arm and the matched variables.

For all MAICs conducted, both in the company submission and the clarification questions, there is a substantial risk of bias from confounding, reverse causation and effect modification, as variables that are matched for have often been reduced from continuous to binary or categorical variables, and no MAIC matches on all variables likely to cause bias. While reducing continuous variables to binary or categorical variables is necessary given the number of variables necessary to match on, the number of participants in each study and limitations in the presented data in studies other than ASPEN, there is likely residual confounding given the distribution of each matched variable has not been fully matched, only a binary or categorical indicator.

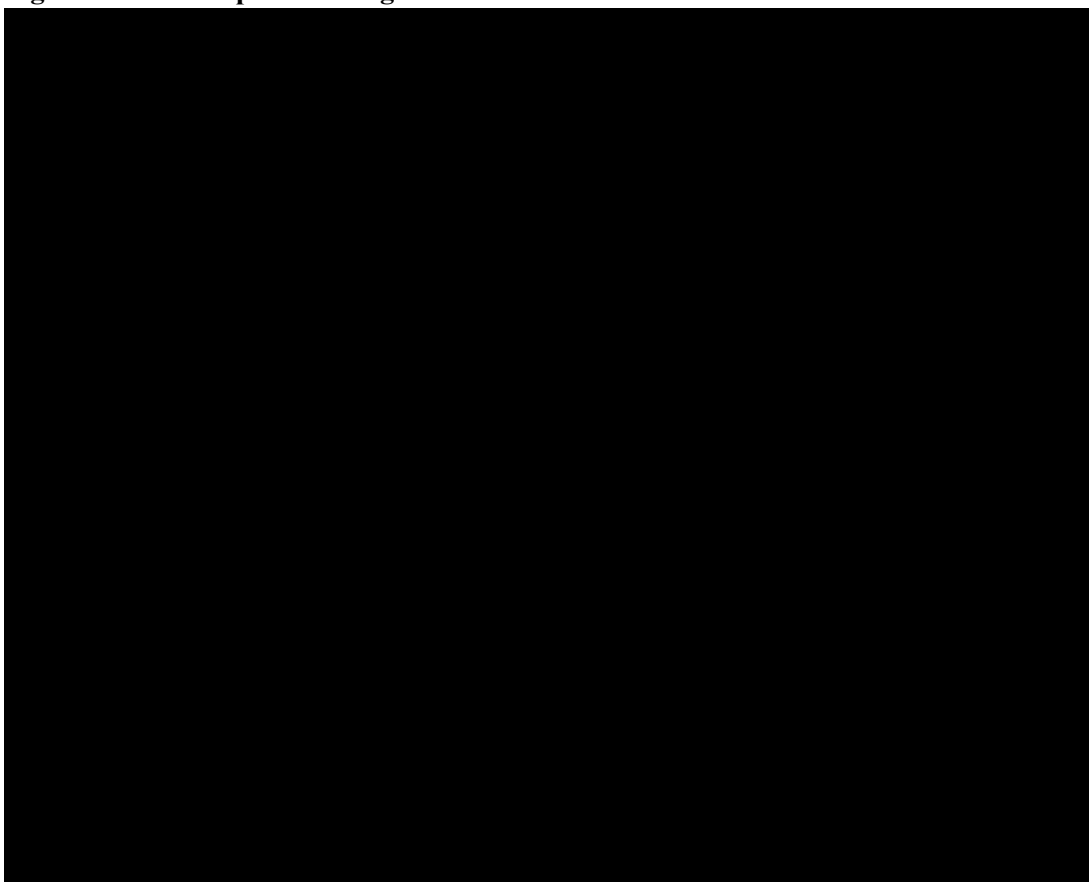
The company submission listed a range of baseline patient variables considered to be potential prognostic factors or effect modifiers and would therefore likely cause bias in a MAIC if the included studies had differences in these variables. In no MAICs were all these variables matched, as no study presented the requisite summary data to match on all variables. For example, in the company submission, the MAIC for BR did not include ECOG performance status, MYD88/CXCR4 mutation status, β2-microglobulin concentration, platelet count, and haemoglobin concentration. Additionally, the MAIC for DRC did not include the number of prior therapies, ECOG performance status, MYD88/CXCR4 mutation status, IgM concentration, β2-microglobulin concentration, or WM IPSS. Similarly, the MAICs in the clarification response did not include many variables considered to be potential prognostic factors or effect modifiers. In all cases, this was due to insufficient data being reported in included studies, but the result is that there is a substantial risk of bias in all MAICs. In the clarification response (Question A21[6]), the company discusses the comparability of the studies included in the company submission MAICs, but the ERG believes that the risk of bias remains substantial due to incomplete variable matching.

In addition to the potential prognostic factors or effect modifiers listed in the company submission, other variables are also to cause bias and were not matched for in the MAICs, including socio-economic status, year of study, location of study, general health of patients. Additionally, the definitions of outcomes were not always consistent between studies, and the interventions were administered differently in each study. Of particular note, the prior treatment status of patients in each study was a

strong prognostic indicator: treatment naïve patients likely have much better outcomes than relapsed/refractory patients. Where possible, the MAICs were conducted in the overall and relapsed/refectory-only populations to account for the differences in prognosis, but the small number of relapsed/refractory patients made this unfeasible for some studies, and the small number of treatment naïve patients in ASPEN made this unfeasible for others. As such, MAICs comparing ASPEN (mixed treatment naïve and relapsed/refractory patients) with studies of treatment naïve patients only (or a higher percentage of treatment naïve patients) are expected to show that the comparator treatment is better for progression-free and overall survival, given the patients themselves are likely to have better outcomes. However, this is indicative of the general problems of using MAIC analyses compared with RCTs: any differences between studies, including their design and populations, could potentially bias the comparison of zanubrutinib and comparator treatments in any direction with any strength. As such, the results from all MAICs should be interpreted with caution, as an unknown, but potentially substantial bias, could affect all results.

However, although the risk of bias is high, there is little evidence of substantial bias in favour of zanubrutinib. Figures 3.4 and 3.5 show the results from all MAICs conducted both in the company submission and clarification responses, split by outcome (PFS and OS) and intervention. The results have not been meta-analysed, as all MAICs are comparisons with ASPEN (and so the variances of all effect estimates are related), and the matched populations are different between each MAIC. Similarly, if the patients in ASPEN had unusually favourable or unfavourable outcomes, this would not be seen as differences on the forest plot as all effect estimates would be affected. Still, the results for the MAICs in the company submission (BR: Tedeschi (2015);[32] DRC: Dimopoukos (2007)[33]/Kastritis (2015)[34]) [redacted] [43] which included only treatment naïve patients and thus the comparison with ASPEN would be expected to favour the comparator treatment.

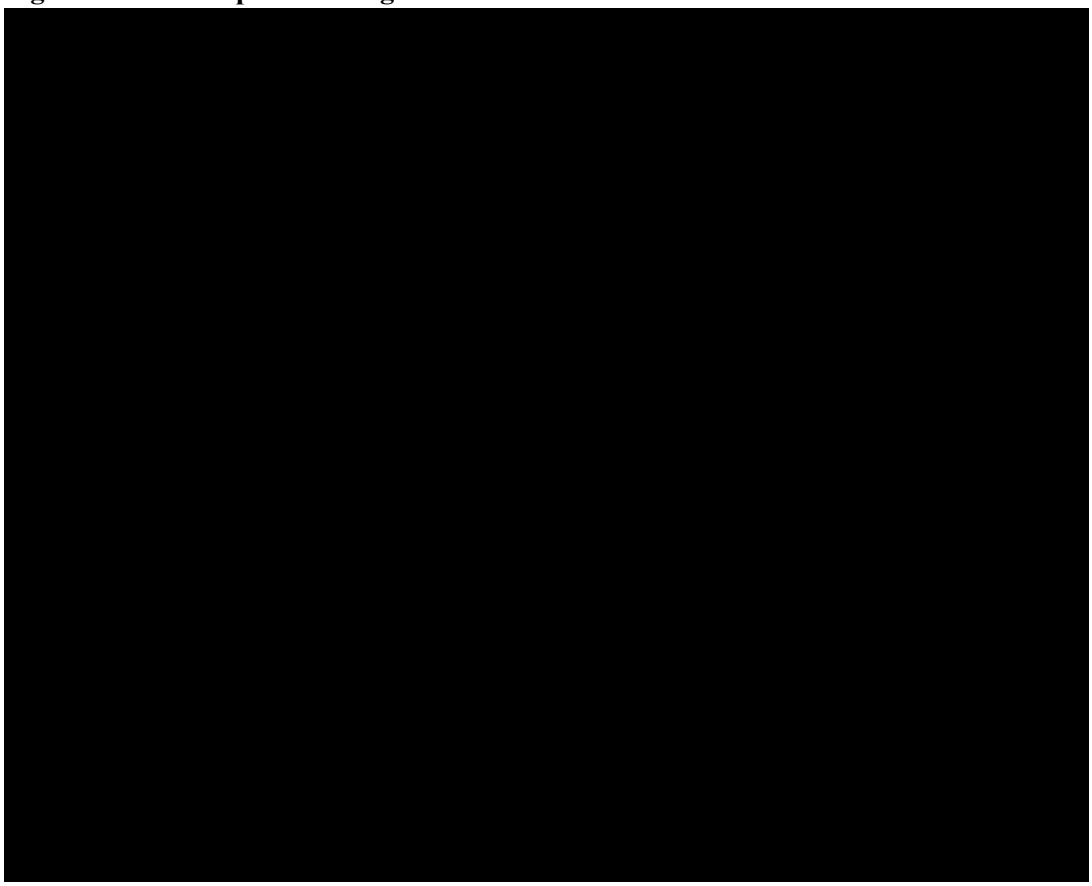
**Figure 3.4: Forest plot showing the hazard ratios of all MAICs conducted for PFS**



BR = bendamustine-rituximab; DRC = dexamethasone, rituximab and cyclophosphamide; FCR/FR = fludarabine, cyclophosphamide and rituximab or fludarabine and rituximab; PFS = progression-free survival; R/R = relapsed/refractory

*Compared with the R/R analysis set (n=83) of ASPEN

**Figure 3.5: Forest plot showing the hazard ratios of all MAICs conducted for OS**



BR = bendamustine-rituximab; DRC = dexamethasone, rituximab and cyclophosphamide; FCR/FR = fludarabine, cyclophosphamide and rituximab or fludarabine and rituximab; OS = overall survival; R/R = relapsed/refractory
*Compared with the R/R analysis set (n=83) of ASPEN

The company submission did not include MAICs for FCR/FR, chlorambucil or rituximab given the differences in study populations. While there may be more risk of bias in these MAICs compared with the BR and DRC MAICs, the ERG believes all MAICs already have a substantial risk of bias, and as such presenting the results for all MAICs seems appropriate.

Finally, it is unclear to what extent the MAICs are relevant to a contemporary NHS population, given differences in baseline variables between the studies in the MAICs (to which the patients in ASPEN were matched) and the patients with WM in the UK today. This could be considered an addition to the already substantial risk of bias from differences between the ASPEN and comparator study populations, increasing the uncertainty in the estimated effectiveness of zanubrutinib.

Overall, the MAIC analyses represent the best use of the available evidence to compare zanubrutinib with comparator treatments (other than ibrutinib). There remains a substantial risk of bias from incomplete matching of prognostic and effect modifier variables, and as such the MAIC results should be interpreted with extreme caution, with appreciation that the risk of bias should increase the uncertainty of the MAIC effect estimates.

### *3.5 Additional work on clinical effectiveness undertaken by the ERG*

No further additional work on clinical effectiveness undertaken by the ERG.

### *3.6 Conclusions of the clinical effectiveness section*

The main evidence for the clinical effectiveness of zanubrutinib was from the ASPEN trial (the BGB-3111-302 study).[2] The ASPEN trial is an ongoing Phase 3, open-label, two-arm, multicentre, randomised study of zanubrutinib versus ibrutinib for the treatment of WM in patients with relapsed/refractory disease, or who are treatment naïve and ineligible for chemoimmunotherapy.[2] The study was designed with two cohorts, according to MYD88 status. Cohort 1 includes patients with *MYD88*$^{MUT}$ who were randomised to either zanubrutinib or ibrutinib. Cohort 2 included patients with *MYD88*$^{WT}$, and all patients were assigned to zanubrutinib.[2] A total of 201 patients were randomised to zanubrutinib or ibrutinib; 164 patients had relapsed/refractory disease (zanubrutinib, n=83 versus ibrutinib, n=81) and 37 were treatment naïve (zanubrutinib, n=19 versus ibrutinib, n=18). Therefore, the evidence for treatment naïve patients is based on small numbers of patients. Patients with cardiovascular disease and those taking warfarin were excluded from the ASPEN trial.

The ASPEN trial was designed to show the superiority of zanubrutinib over ibrutinib in patients with *MYD88*$^{MUT}$ WM (relapsed/refractory arm of Cohort 1). The primary efficacy analyses were the rate of complete response or very good partial response (CR/VGPR), as assessed by IRC with adaption of the response criteria updated at the Sixth International Workshop on Waldenström Macroglobulinemia (IWWM) every 28 days and every 84 days after Cycle 12. Two analysis sets were considered: 1) the relapsed/refractory analysis set, and 2) the ITT analysis set. The ITT analysis set comprised all randomised patients assigned to a treatment arm, while the relapsed/refractory analysis set comprised all patients in the ITT analysis set with at least one prior line of therapy. In the CS, the primary results are for the ITT analysis set.

The median age of all patients in the ITT analysis set (Cohort 1) was 70.0 years. The majority of patients were male (66.7%), white (91.0%), had an ECOG performance status of 0 or 1 and were enrolled in sites in Europe (59.7%), Australia/New Zealand (30.8%) or North America (9.5%). The demographics and baseline characteristics were generally similar across treatment arms, however, more patients randomised to zanubrutinib than ibrutinib were >75 years old (33.3% and 22.2%, respectively) and more were anaemic (haemoglobin ≤110 g/L in 65.7% and 53.5% of patients, respectively).[24]

The median follow-up time as of the data cut-off date was 19.5 months for zanubrutinib-treated patients and 19.4 months for ibrutinib-treated patients. As of the data cut-off date (31 August 2019), a total of 158 patients (78.6%) were continuing study treatment (81 patients [79.4%] in the zanubrutinib treatment arm and 77 patients [77.8%] in the ibrutinib treatment arm). The most common reason for discontinuing study treatment was progressive disease (seven [6.9%] zanubrutinib versus five [5.1%] ibrutinib-treated patients) and AE (four [3.9%] zanubrutinib treated patients versus nine [9.1%] ibrutinib-treated patients). A total of 158 (78.6%) patients were continuing to participate in the study and 41 (20.4%) discontinued from the study.

In Cohort 1, the rate of IRC-assessed CR and VGPR was 28.4% in all patients treated with zanubrutinib and 19.2% in patients treated with ibrutinib (95% CI, -1.5–22.0; p=0.09). The estimated difference between the two arms adjusted for the stratification factors and age group was 10.2%.[2] In the relapsed/refractory population, 28.9% of patients treated with zanubrutinib and 19.8% treated with ibrutinib achieved VGPR or CR (with estimated difference of 10.7% (95% CI: 2.5 to 23.9; p=0.116)).[2] According to the company, the testing for the primary endpoint of VGPR or CR rate superiority required testing in the relapsed/refractory analysis set prior to testing in the ITT analysis set. The primary efficacy endpoint was not significant in the relapsed/refractory analysis set (p=0.116), thus the study did not meet the primary efficacy endpoint and testing for other endpoints and resulting p-values in the following sections are descriptive.[24]

At the time of the data cut-off date, median PFS and OS had not been reached in either treatment arm.[24] Therefore, survival data for zanubrutinib are currently immature.

Overall, zanubrutinib has a comparable safety and tolerability profile compared with ibrutinib. Neutropenia was consistently more prevalent in the zanubrutinib groups compared with the ibrutinib group (22.5% overall, 24.8% in Cohort 1 vs 12.2%). The zanubrutinib treated patients had a lower rate of several AEs compared with ibrutinib, such as atrial fibrillation (2.0% versus 15.3%), major haemorrhage (5.9% versus 9.2%) and hypertension (10.9% versus 16.3%). There were also fewer AEs leading to death (1.0 versus 4.1%), discontinuation due to AEs (4.0 versus 9.2%) and AEs leading to dose reduction (13.9 versus. 23.5%) with zanubrutinib compared with ibrutinib. In a pooled analysis of 253 patients with WM, the most common AEs reported by zanubrutinib treated patents were upper respiratory tract infection (32.4%) and diarrhoea (21.7%), vascular disorders (19.4%), confusion (17.8%), renal and urinary disorders (17%), neutropenia (16.6%), and cough (16.6%), fatigue (15%).

As part of the response to clarification the company provided data based on a follow-up analysis of safety and efficacy conducted with the cut-off date of 31 August 2020. However, all 2020 efficacy analyses are based on assessments by the investigator.[26] Therefore, none of the primary and secondary outcomes presented in the company submission have been reported for the data cut-off date of 31 August 2020 (IRC-assessed VGPR/CR rate (primary endpoint), IRC-assessed duration of response, IRC-assessed progression-free survival and IRC-assessed time to response (secondary endpoints)). Comparing the 2020 results with 2019 results reported in the CSR, it looks like the 2020 results are in line with the 2019 results. However, this cannot be verified for the outcomes reported in the CS.

The description of the comparators in the NICE scope is as follows: For people who have had at least one prior therapy: rituximab and bendamustine (BR); dexamethasone, rituximab and cyclophosphamide (DCR); fludarabine and rituximab with or without cyclophosphamide (FR or FCR); cladribine and rituximab (Clad-R); and autologous stem cell transplantation (SCT) in people for whom autologous SCT is suitable. For people for whom chemo-immunotherapy is unsuitable: chlorambucil; rituximab monotherapy; and best supportive care (BSC) including blood product transfusions, plasma exchange, granulocyte stimulating factors and intravenous immunoglobulin infusions.[5]

The company undertook a systematic literature review (SLR) to identify published evidence for current and future treatment options for patients with WM. Specifically, the SLR aimed to identify the efficacy of treatment options for previously untreated patients published evidence for current and future treatment options for patients with WM, as well as the safety and tolerability of treatment options for patients with WM. This SLR did not include any search terms for autologous SCT.

The SLR identified four articles for bendamustine and rituximab (BR), five articles for dexamethasone, rituximab and cyclophosphamide (DRC), six articles for fludarabine, cyclophosphamide and rituximab (FCR) or fludarabine and rituximab (FR), two articles for chlorambucil monotherapy and six articles for rituximab monotherapy. No studies were identified for cladribine and rituximab (Clad-R) or best supportive care (BSC). All identified studies were single-arm studies, except one retrospective study investigating both BR and DRC.[31]

The company only included BR and DCR as comparators. In addition, the company included ibrutinib as a comparator. However, ibrutinib was explicitly excluded as a comparator by NICE because it is currently available through the Cancer Drugs Fund and therefore not considered established practice.[8] The feasibility of conducting indirect comparisons of zanubrutinib with FR, FCR is discussed in Sections 3.3 and 3.4 of this report. The ERG believes that the indirect comparisons between zanubrutinib and FR, or FCR are equally as valid as the comparisons with BR and DCR and should

therefore have been included in the CS. The ERG agrees with the company that a comparison of zanubrutinib with Clad-R is not feasible.

Due to the lack of RCTs comparing zanubrutinib with any comparator other than ibrutinib, and a lack of common comparators for an anchored indirect treatment comparison (ITC), the company conducted a matching adjusted indirect comparison (MAIC). The company assessed the included studies for inclusion in the MAIC, on the definition of patient populations, availability of progression-free survival (PFS) and overall survival (OS) Kaplan-Meier (KM) curves to inform the cost effectiveness analysis, availability and comparability of baseline patient characteristics, study design, sample size, and geographical location. Only PFS and OS were considered as outcomes in the MAIC. It should be noted that survival data for zanubrutinib are immature.

In the company submission, the only comparisons in the MAICs were between zanubrutinib, BR and DRC. The aim of the MAICs were therefore to compare outcomes between zanubrutinib, BR and DRC, reducing the potential for bias by adjusting for confounding variables. Given the relative scarcity of data, the MAICs were necessary for any comparisons between zanubrutinib and treatments other than ibrutinib. Individual participant data (IPD) was only available for the zanubrutinib arm of APSEN: summary data from the included trials were used for the BR and DRC trials. In the clarification response, the company performed MAICs for the following treatments: BR (three additional analyses), DRC (three additional analyses), fludarabine, cyclophosphamide and rituximab or fludarabine and rituximab (FCR/FR, four analyses), chlorambucil (one analysis), and rituximab (one analysis).

Three MAICs were conducted in the company submission, comparing the overall zanubrutinib population of ASPEN (n=102) with the populations the BR (n=71) and DRC (n=72) separately, and also for patients with relapsed/refractory disease in the zanubrutinib arm of ASPEN (n=83) to the BR population, as the BR population only included relapsed/refractory patients. Individual patient-level event and censoring times for OS and PFS were extracted from the KM curves presented for BR and DRC, which were compared with the KM curves created using both unmatched and matched data for zanubrutinib. Survival was compared by estimating hazard ratios (HRs) using Cox proportional hazard (PH) models, using the reconstructed patient data extracted for BR and DRC.

Table 3.22 shows the results for survival from the MAICs in the company submission after matching, comparing zanubrutinib with BR and DRC, where a HR less than 1 indicates zanubrutinib improves survival. It is unclear whether the effect estimate for BR is from a comparison with all patients in ASPEN, or just the relapsed/refractory patients, although the survival curves were similar.

**Table 3.22: MAIC survival results after matching in the company submission**

| | BR | DRC |
|---|---|---|
| PFS, HR (95% CI) | [redacted] | [redacted] |
| OS, HR (95% CI) | [redacted] | [redacted] |

Source: CS, Section B.2.9.3.[4]
BR = bendamustine rituximab; CI = confidence interval; DRC = dexamethasone, rituximab, and cyclophosphamide; OS = overall survival; PFS = progression-free survival

Given the lack of direct evidence, indirect methods were used, and given the lack of RCT evidence to inform a network meta-analysis, MAICs using single-arm studies represent the only available evidence to adjust for confounding comparing zanubrutinib with comparator treatments. Even so, these analyses still present a substantial risk of bias.

For all MAICs conducted, both in the company submission and the clarification questions, there is a substantial risk of bias as variables that are matched for have often been reduced from continuous to binary or categorical variables, and no MAIC matches on all variables likely to cause bias. While reducing continuous variables to binary or categorical variables is necessary given the number of variables necessary to match on, the number of participants in each study and limitations in the presented data in studies other than ASPEN, there is likely residual confounding given the distribution of each matched variable has not been fully matched, only a binary or categorical indicator.

The company submission listed a range of baseline patient variables considered to be potential prognostic factors or effect modifiers and would therefore likely cause bias in a MAIC if the included studies had differences in these variables. In no MAICs were all these variables matched, as no study presented the requisite summary data to match on all variables. For example, in the company submission, the MAIC for BR did not include ECOG performance status, MYD88/CXCR4 mutation status, β2-microglobulin concentration, platelet count, and haemoglobin concentration. Additionally, the MAIC for DRC did not include the number of prior therapies, ECOG performance status, MYD88/CXCR4 mutation status, IgM concentration, β2-microglobulin concentration, or WM IPSS. Similarly, the MAICs in the clarification response did not include many variables considered to be potential prognostic factors or effect modifiers. In all cases, this was due to insufficient data being reported in included studies, but the result is that there is a substantial risk of bias in all MAICs.

In addition to the potential prognostic factors or effect modifiers listed in the company submission, other variables are also to cause bias and were not matched for in the MAICs, including socio-economic status, year of study, location of study, general health of patients. Additionally, the definitions of outcomes were not always consistent between studies, and the interventions were administered differently in each study. As such, the results from all MAICs should be interpreted with caution, as an unknown, but potentially substantial bias, could affect all results. However, although the risk of bias is high, there is little evidence of substantial bias in favour of zanubrutinib.

The company submission did not include MAICs for FCR/FR, chlorambucil or rituximab given the differences in study populations. While there may be more risk of bias in these MAICs compared with the BR and DRC MAICs, the ERG believes all MAICs already have a substantial risk of bias, and as such presenting the results for all MAICs seems appropriate.

Finally, it is unclear to what extent the MAICs are relevant to a contemporary NHS population, given differences in baseline variables between the studies in the MAICs (to which the patients in ASPEN were matched) and the patients with WM in the UK today. This could be considered an addition to the already substantial risk of bias from differences between the ASPEN and comparator study populations, increasing the uncertainty in the estimated effectiveness of zanubrutinib.

Overall, the MAIC analyses represent the best the only available evidence that attempts to adjust for confounding to compare zanubrutinib with comparator treatments (other than ibrutinib). There remains a substantial risk of bias from incomplete matching of prognostic and effect modifier variables, and as such the MAIC results should be interpreted with extreme caution, with appreciation that the risk of bias should increase the uncertainty of the MAIC effect estimates.

# 4. COST EFFECTIVENESS

## *4.1 ERG comment on company's review of cost effectiveness evidence*

This section pertains mainly to the review of cost effectiveness analysis studies. However, the search section (4.1.1) also contains summaries and critiques of other searches related to cost effectiveness presented in the company submission. Therefore, the following section includes searches for the cost effectiveness analysis review, measurement and evaluation of health effects as well as for cost and healthcare resource identification, measurement and valuation.

### 4.1.1 Searches performed for cost effectiveness section

The following paragraphs contain summaries and critiques of all searches related to cost effectiveness presented in the company submission.

Appendix G of the CS details systematic searches of the literature used to identify evidence on the economic outcomes of WM, including economic analyses, cost studies and observational/real world studies.

Searches were conducted on 25 September 2020. The search methods in G1.1 of the CS Appendices[12] reported that the searches were restricted to English language and did not include a date limit. Database date spans were not reported.

The sources searched are described in G.1.1.2. Although a search strategy was presented and databases were listed, the reporting was unclear and contained omissions and errors. This section reported Medline & Medline In-Process were searched via PubMed, however there was no documentation for a PubMed search strategy. It was unclear which host was used to search Embase and EconLit. A strategy presented in Table B.5.29 of the CS did not specify for which database it had been used to search. During the clarification process, the company explained that Embase, Medline and EconLit were searched simultaneously via Proquest[13] The company clarified that this combined strategy was reported in Table B.5.29 of the CS.

A single, simultaneous search of Medline, Embase and EconLit included a study design filter to capture cost effectiveness, resource cost, health utility studies. No database date spans or date limits were reported. A summary of the sources searched is provided in Table 4.1.

**Table 4.1: Data sources for the cost effectiveness systematic review**

| Search strategy element | Resource | Host/ source | Reported date range | Date searched |
|---|---|---|---|---|
| Electronic databases | Embase | Proquest | Not reported. | 25.9.20 |
| | Medline & Medline In-Process | Proquest | Not reported. | 25.9.20 |
| | EconLit | Proquest | Not reported. | 25.9.20 |
| Source: Appendix G of the Company's submission and the clarification response.[12] | | | | |

**ERG comment:** Searches was undertaken to identify published economics evaluations. Appendix G[12] provided sufficient details for the ERG to appraise the literature searches. Three databases were reported as searched. No conference proceedings, reference checking or grey literature searches were reported. For the most part, searches were well documented, making them reproducible.

The cost effectiveness searches presented in the CS Appendix G were conducted in September 2020. As the cost effectiveness searches were run from over eight months ago, the ERG considers it possible that potentially relevant studies published since September 2020 may be missing from the review. As with the clinical effectiveness search, the ERG queried search currency during the clarification process and the company responded to say they had conducted a separate targeted search to explore whether there were new publications. The clarification response[13] stated that two new publications which might be included were not relevant to the network meta-analysis. There was no information about additional results in relation to the cost effectiveness search. No details of the strategies used, date span, date of search, number of results retrieved were not provided to the ERG. The clarification response did not include an updated PRISMA flowchart. The ERG is unable to assess how the targeted searches were conducted or screened.

The company confirmed that the single cost effectiveness search strategy reported in Table B.5.29 presented a combined, simultaneous search of Medline, Medline In-Process, Embase and EconLit. Under closer examination, the ERG has some concerns about the accuracy of reporting in this table. The strategy includes Medline and Embase specific indexing. At the end of Table B.5.29 are several lines specifying which lines present Medline and Embase results, with and without duplication. At no point are the EconLit results clearly described, and there are no details regarding how the EconLit duplicate records were handled. The PRISMA flowchart (figure B.5.6, page 55 of the clarification response)[13] reports only results from Embase and Medline. There is no indication that EconLit had contributed to the search strategy or the PRISMA flowchart. The ERG was not satisfied with the clarification explanation about inclusion of EconLit in a simultaneous search. There is a marked lack of transparency in the reporting of the cost effectiveness searches. As the ERG was not able to reproduce the search, due to lack of access to the Proquest host., and it remains unclear whether or not EconLit was actually searched.

The ERG noted that the combined cost effectiveness strategy used very limited synonyms for the condition, WM. The ERG identified several word and spelling variants (see below) which could have been included to increase recall of potentially relevant WM studies.

| | | |
|---|---|---|
| 1 | Plasmacytoid Lymphocytic Lymphoma.ti,ab. | 16 |
| 2 | (Lymphoplasmacytoid Lymphoma or Macroglobuli?emia*).ti,ab. | 4749 |
| 3 | (waldenstroem macroglobulin* or atypical macroglobulin* or waldenstrom macroglobulin*).ti,ab. | 1848 |
| 4 | (macro globuli?nemia or macrocryoglobulin?emia).ti,ab. | 13 |
| 5 | or/1-4 | 4965 |

A single search was carried out to identify evidence for the cost effectiveness, health-related quality of life and cost and healthcare resource identification, measurement and valuation sections of the company submission. Three separate study design filters were combined within the single strategy, resulting in duplicated lines. Please see Appendix 1 for full details of the redundant, duplicates lines. These lines could have been streamlined to reduce repetition however strategy performance was not adversely affected.

Section H.1.2.1 reported that the combined cost effectiveness search results in Table B.5.2.29 were screened for published utility and disutility studies. On page 60 in Appendix H,[12] the number of results screened was reported as 275 records. The strategy in Table B.5.29 reported 265 results which was

repeated in the PRISMA flowchart (Figure B.5.6). The ERG considers it likely that the number reported in section H.1.2.1 may be a typographical error.

During the clarification process, the ERG queried the rationale for applying an English language limit to the Embase/MEDLINE clinical and cost effectiveness searches. The company responded that "The rationale for limiting the searches to English literature only was based on guidance provided by NICE; Chapter 5.4 of Developing NICE guidelines: the manual states that with regards to limits and filters, searches should be limited to studies reported in English." The company cited the NICE manual for developing guidelines[14] as their source. Although the Guidelines manual does recommend using an English language limit, the user guide for company STA evidence submissions[15] refers to the guide to the methods of technology appraisal[16] and guidance from the Centre for Reviews and Dissemination.[17] The latter source clearly states that "limiting searches to English language papers can introduce language bias."[17] Consequently the ERG remains concerned that limiting the searches to English language only studies may have introduced language bias. Current best practice states that "Whenever possible review authors should attempt to identify and assess for eligibility all possibly relevant reports of trials irrespective of language of publication"[18] and that "research related to language bias supports the inclusion of non-English studies in systematic reviews".[19, 20]

On the whole, the ERG felt that the cost effectiveness searches would have benefited from clearer reporting, a full update search, more comprehensive terminology for WM, and removal of the English language limit.

### 4.1.2 Inclusion/exclusion criteria

In- and exclusion criteria for the review on cost effectiveness studies, utilities and costs and resource use are presented in Table 4.2.

**Table 4.2: Eligibility criteria for the systematic literature reviews**

| | Inclusion criteria | Exclusion criteria |
|---|---|---|
| **Patient population** | Adult patients with WM, with or without previous treatment (i.e. treatment naïve or relapsed/refractory) | Patients receiving treatment for secondary malignancies (focus of treatment aims to treat another underlying malignancy)<br>Healthy subjects<br>Children (<18 years of age) |
| **Intervention** | No restrictions applied | N/A |
| **Comparator** | No restrictions applied | N/A |
| **Outcomes(s) 1 (Published economic evaluations)** | Total costs<br>QALYs<br>LYG<br>ICER/ICUR<br>Cost per progression-free year | Any outcome not specified in inclusion criteria |
| **Outcomes(s) 2 (Utility studies)** | Health State Utility values elicited using direct methods: TTO and standard gamble<br>Preference-Based methods: (e.g.EQ-5D, HUI3, SF-6D, aqol, QWB, 15D) | Publications that do not report data on relevant outcomes |

| | Inclusion criteria | Exclusion criteria |
|---|---|---|
| | VAS | |
| **Outcomes(s) 3 (Cost/resource use studies)** | Any outpatient and inpatient healthcare resource utilisation<br>Any direct costs of inpatient and outpatient services<br>Any indirect costs<br>Any costs of AEs | Any outcome not specified in inclusion criteria |
| **Study design 1 (Cost effectiveness analysis studies)** | CUA<br>CEA<br>CMA | Study designs not specified in inclusion criteria |
| **Study design 2 (Utility studies)** | Clinical trials<br>Observational studies | Study designs not specified in inclusion criteria |
| **Study design 3 (Cost/resource use studies)** | Economic evaluations<br>Patient chart reviews<br>Patient and disease registry studies<br>Claims data analyses/observational studies (excluding any studies reporting frequency of AEs) | Study designs not specified in inclusion criteria |
| Source: CS appendix G Table B.5.30 and CS appendix H Table B.5.33 | | |

**ERG comment:** The ERG agrees that the eligibility criteria are suitable to fulfil the company's objective to identify cost effectiveness studies. The rationales for excluding CE studies after full paper reviewing are considered appropriate given the defined in- and exclusion criteria.

### 4.1.3 Conclusions of the cost effectiveness review

The CS provides an overview of the included cost effectiveness, utility and resource use and costs studies, but no specific conclusion was formulated. The eligibility criteria were suitable for the SLR performed. However, the ERG felt that the cost effectiveness searches would have benefited from clearer reporting, a full update search, more comprehensive terminology for WM, and removal of the English language limit as reported above.

## *4.2 Summary and critique of company's submitted economic evaluation by the ERG*

### 4.2.1 NICE reference case checklist

**Table 4.3: NICE reference case checklist**

| Element of health technology assessment | Reference case | ERG comment on company's submission |
|---|---|---|
| **Population** | As per NICE scope | The company used the ITT of the ASPEN trial population for their (base-case) analyses, which only contained patients with MYD88 mutation. |
| **Comparators** | Therapies routinely used in the National Health Service (NHS), including | In contrast to the NICE scope, the model does not include FR, FCR, Clad-R and ASCT (for |

| Element of health technology assessment | Reference case | ERG comment on company's submission |
|---|---|---|
| | technologies regarded as current best practice | patients who have had at least one prior therapy), chlorambucil, rituximab monotherapy and BSC (for patients for whom chemo-immunotherapy is unsuitable) as comparators. |
| **Perspective on outcomes** | All direct health effects, whether for patients or, when relevant, carers | In line with NICE reference case |
| **Perspective on costs** | NHS and PSS | In line with NICE reference case |
| **Type of economic evaluation** | Cost utility analysis with fully incremental analysis | The company did not perform a fully incremental analysis. |
| **Time horizon** | Long enough to reflect all important differences in costs or outcomes between the technologies being compared | In line with NICE reference case |
| **Synthesis of evidence on health effects** | Based on systematic review | In line with NICE reference case |
| **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. | In line with NICE reference case |
| **Source of data for measurement of health-related quality of life** | Reported directly by patients and/or carers | In line with NICE reference case |
| **Source of preference data for valuation of changes in health-related quality of life** | Representative sample of the UK population | In line with NICE reference case |
| **Equity considerations** | An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit | In line with NICE reference case |
| **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 | In line with NICE reference case |
| **Discounting** | The same annual rate for both costs and health effects (currently 3.5%) | In line with NICE reference case |

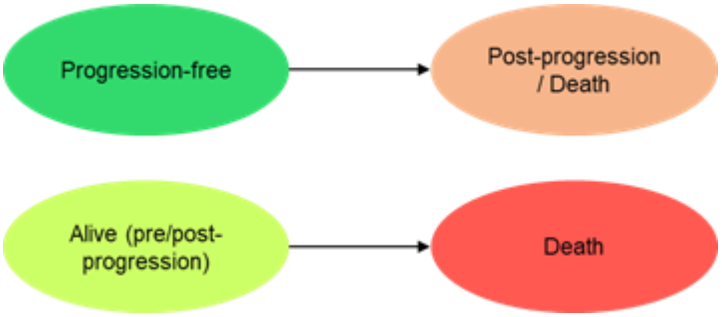
### 4.2.2 Model structure

The company developed a cohort partitioned survival model (PSM) in Excel to project the long-term clinical and economic consequences.

#### 4.2.2.1 Health states/events and transitions

The PSM consisted of three mutually exclusive health states: 1) pre-progression, 2) post-progression and 3) death (Figure 4.1). All patients started in the pre-progression health state and could either remain progression-free and stay in this health state or transition to the post-progression (if disease progression occurred) or death (if mortality event occurred) health states. OS and PFS were used to calculate the health state occupancy in each model cycle.

**Figure 4.1: Model structure**




Source: Based on Figure B.3.1 of the CS

**ERG comment:** The main concern of the ERG relates to the underlying assumptions of the choice of a partitioned survival model.

The company chose a PSM modelling approach over a STM approach and argued that this was in line with common modelling approaches and assumptions in oncology. Although the ERG agrees that PSMs are often used in oncology, these conditions often have a relatively short PFS and OS. In the current model, PFS and OS are relatively long (for example, median PFS DRC 5.66 years and median PFS BR 5.00 years as presented in the CS base-case model file) and as a result HRQoL and cost and resource use are stable over a relatively long period. Next to that, the model's health state occupancy was based on secondary (PFS) and exploratory (OS) outcomes in the ASPEN trial. In question B4 of the clarification letter,[6] the ERG asked the company to justify the above mentioned assumptions, to justify the use of a PSM given the issues highlighted in NICE DSU TSD 19 (particularly regarding the extrapolation of PFS and OS while assuming structural independence between these endpoints) and to provide a STM. In their response, the company highlighted why a STM approach was deemed less appropriate for this submission, but did not elaborate on the specific sub questions from the ERG related to using secondary/exploratory outcomes for health state occupancy and the plausibility of HRQoL (except for age-adjusted utility values in the company's revised base-case after clarification) and cost and resource use being stable over a relatively long period. The ERG therefore questions the plausibility of the assumption that cost and resource use are stable over the time period patients remain in the progression-free health state. The company justified their PSM approach stating that a STM approach was not applicable because of lack of data to inform this. The ERG acknowledges that this is likely a limitation. However, the ERG also questions the underlying assumptions (such as the assumption that the modelled survival endpoints are structurally independent, as highlighted in NICE DSU TSD 19) of the current PSM approach.

### 4.2.3 Population

In line with its anticipated marketing authorisation, the final scope issued by NICE and the population in the ASPEN phase III trial[2], zanubrutinib was considered in the cost effectiveness model for the treatment of adult patients with WM previously treated with at least one prior line of therapy, or who

are treatment naïve and unsuitable for chemo-immunotherapy. The company used the ITT of the ASPEN trial population for their analyses, which only contained patients with MYD88 mutation.

For the comparisons of zanubrutinib with BR and DRC, populations of the ASPEN trial[2] and respectively the Tedeschi et al 2015[32] and Dimopoulos et al. 2007/Kastritis et al. 2015[33, 34] were matched. In the base-case analysis, baseline patient characteristics were based on the unadjusted data of the ASPEN ITT population for all three pairwise comparisons. Scenario analyses were conducted using the baseline patient characteristics after matching adjustment (Table 4.4)

**Table 4.4: Key baseline patient characteristics used in the economic model**

| | Base-case (N=201) | Zanubrutinib (match BR; Neff = [redacted]) | Zanubrutinib (match DRC; Neff = [redacted]) | Source |
|---|---|---|---|---|
| Female proportion (%) | 33.33 | 39.05 | 39.52 | ASPEN IPD |
| Mean age (years) | 69.53 | 70.84 | 69.39 | ASPEN IPD |
| Body surface area ($m^2$) | 1.86 | 1.84 | 1.87 | ASPEN IPD |
| Source: based on CS Tables B.3.3 and B.3.4<br>Abbreviations: IPD = individual patient data, Neff = effective sample size | | | | |

**ERG comment:** The main concern of the ERG relates to the use of only patients with a *MYD88*$^{MUT}$ mutation.

As discussed in Section 2.1, the company based their analyses on patients from Cohort 1 of the ASPEN trial, which all had a *MYD88*$^{MUT}$ mutation. Although, according to the company, 90% of patients with WM have this mutation, other MYD88 mutations may be found in 5–10% of patients. Patients with mutations other than *MYD88*$^{MUT}$ were included in Cohort 2 of the ASPEN trial, and because the ERG was concerned about the generalisability of the results based on Cohort 1 only, it requested a scenario with a pooled analysis of both cohorts. The company provided this analysis in response to clarification question B6 and the pooled result was comparable to the result based on Cohort 1 only. A remaining issue is, however, that it is unclear what the mutation status was of patients in the comparator arm of the trials. The ERG considers that the cost effectiveness analysis in Cohort 1 of the ASPEN trial may be a reasonable approximation to cost effectiveness in the overall WM population. However, given that the mutation status of patients in the comparator arms of the model was unclear, the cost effectiveness in the overall population (including patients with other variants) remains difficult to assess.

### 4.2.4 Interventions and comparators

Zanubrutinib was considered within the economic evaluation as per the anticipated licenced indication in WM. Zanubrutinib was, in line with the dosage used in the ASPEN trial[2], modelled with a posology of 160 mg orally twice daily until disease progression or until no longer tolerated by the patient.

The comparators considered in the CS were ibrutinib (420 mg orally once daily), BR (rituximab (375 mg/m², day 1) + bendamustine (90 mg/m², days 1 and 2) IV infused every cycle, repeated every four weeks, until six cycles or disease progression) and DRC (dexamethasone 20 mg IV on day 1, rituximab 375 mg/m² IV on day 1, and cyclophosphamide 100 mg/m² orally twice daily on days 1 through 5, repeated every three weeks, until six cycles or disease progression). Next to BR and DRC, the NICE scope listed the following comparators: fludarabine and rituximab (FR), cyclophosphamide and rituximab (FCR), cladribine and rituximab (Clad-R) and autologous stem cell transplantation (ASCT)

(for patients who have had at least one prior therapy), chlorambucil, rituximab monotherapy and BSC (including blood product transfusions, plasma exchange, granulocyte stimulating factors and intravenous Ig infusions for patients for whom chemo-immunotherapy is unsuitable). However, the company argued that other than BR and DRC, it was not possible to conduct comparisons with chemotherapy regimens or BSC, due to a lack of data in the literature to enable comparison of zanubrutinib with the comparators of interest. Ibrutinib, BR and DRC were also a subsequent treatment option in the economic model.

**ERG comment:** The main concerns of the ERG relate to: a) not modelling all comparators that were included in the NICE scope; b) including ibrutinib as a comparator and subsequent treatment option in the economic model; and c) assuming different relative dose intensities for BR and DRC.

a) In contrast to the NICE scope, the model does not include FR, FCR, Clad-R and ASCT (for patients who have had at least one prior therapy), chlorambucil, rituximab monotherapy and BSC (for patients for whom chemo-immunotherapy is unsuitable) as comparators. The company stated that "Other than BR and DRC, it was not possible to conduct comparisons with chemotherapy regimens or BSC, due to a lack of data in the literature to enable comparison of zanubrutinib with the comparators of interest".[4] In response to clarification question B1, the company stated that it considered conducting exploratory analyses during the clarification stage by relying on certain assumptions, such as (1) assuming equivalent clinical outcomes between BR and other chemo-immunotherapies (e.g., FR/FCR/Clad-R, chlorambucil) specified for patients with relapsed or refractory disease (i.e. adults with WM who have received at least one prior therapy), and (2) applying actual drug costs specific to each comparator regimen (e.g., FR/FCR/Clad-R, chlorambucil). Although the company acknowledges that there was some evidence from the MAIC comparing zanubrutinib to some of these comparator regimens, additional analyses were not performed. The ERG agrees that because MAICs were conducted in a pairwise manner, the results were not informative for a full incremental analysis for the same patient population. However, based on the company's response to question A26, the ERG argues that the company could have done exploratory analyses for the comparisons of zanubrutinib with FR/FCR and rituximab monotherapy.
b) The company included ibrutinib as a comparator and as a subsequent treatment to BR and DRC in the economic model. However, NICE's position statement[44] *"*consideration of products recommended for use in the Cancer Drugs Fund as comparators, or in a treatment sequence, in the appraisal of a new cancer product*"* states that technologies available through the CDF should not be modelled in treatment sequences. The ERG therefore asked the company to explore alternative subsequent treatments in a scenario, but the company did not respond to this request. To justify the inclusion of ibrutinib as a comparator and subsequent treatment, the company states that ibrutinib is considered to be clinically relevant as a subsequent treatment, given that data of the UK WMUK Rory Morrison Registry up to 2018[1] indicates that BTK inhibitors (currently only ibrutinib is available) are an emerging standard of care in patients who have had ≥1 prior therapy, with ibrutinib being the most frequently used treatment in clinical practice. Nevertheless, based on NICE's position statement, the ERG ignores the evidence for the comparison of zanubrutinib with ibrutinib and excluded the possibility of ibrutinib as a subsequent treatment in the ERG base case analysis. Instead, although the ERG acknowledges that this may not perfectly reflect UK clinical practice, the ERG assumed in its base-case that patients initially treated with BR would receive DRC as subsequent treatment and patients initially treated with DRC would receive BR. Patients initially receiving zanubrutinib received subsequent treatment according to the CS base-case (BR for 60.4%

of the patients and DRC for 39.6% of the patients). In addition, the ERG explored a scenario in which all subsequent treatments were excluded from the model.

c) In its base case analysis, the company assumed a higher relative dose intensity for BR and DRC (100%) than for zanubrutinib (97.5%). In response to clarification question B2, the company stated that there was a lack of reported relative dose intensities for BR and DRC and agreed that an alternative would be to use the same estimate between zanubrutinib and the comparators. The company provided a scenario analysis assuming the same relative dose intensity rates for zanubrutinib, BR and DRC, which had very minor impact on the ICER. Given the lack of evidence that the relative dose intensity rates would differ between the treatments, the ERG adopted the assumption of equal relative dose intensity rates of 97.5% for BR, DRC, and zanubrutinib in its base case analysis.

### 4.2.5 Perspective, time horizon and discounting

The analysis is performed from the NHS and Personal Social Services (PSS) perspective. Discount rates of 3.5% are applied to both costs and benefits. The model cycle length is 28 days with a lifetime time horizon (30 years) and a half-cycle correction is applied.

### 4.2.6 Treatment effectiveness and extrapolation

The main sources of evidence on treatment effectiveness used for intervention and comparators are the ASPEN trial for the comparison of zanubrutinib (only Cohort 1) and ibrutinib, and MAICs comparing zanubrutinib with BR and DRC based on the ASPEN trial and Tedeschi et al 2015 and Dimopoulos et al. 2007/Kastritis et al. 2015,[33, 34] respectively (see Section 3.4 for more detail). Patient baseline characteristics used in the model were gender, age and body surface area and were derived from the ASPEN ITT population in the company's base-case analysis and from the ASPEN matched populations in two exploratory scenarios (see Tables B.3.3 and B.3.4 of the CS).

The company used survival analysis on secondary and exploratory endpoints PFS and OS to extrapolate treatment effectiveness for zanubrutinib and comparators beyond the available trial data. Survival analysis on TTD data from ASPEN (ITT Cohort 1 or matched in comparisons with DRC and BR) were used to estimate treatment duration for zanubrutinib and ibrutinib.

The company's survival analysis methods very closely followed recommendations from the NICE DSU TSD 14 [45], that is, the proportional hazards assumption was assessed for each comparison and outcome, parametric distributions were fitted and the most plausible distribution was selected based on statistical goodness of fit (Akaike Information Criterion; AIC, and Bayesian Information Criterion; BIC), visual inspection of 95% CIs and external validity through published estimates and expert opinion.

#### 4.2.6.1 Zanubrutinib versus ibrutinib comparison

*Overall survival*

The company selected dependent exponential models. According to the company, the proportional hazards assumption could be supported. The exponential model had the best statistical fit, and its mean OS was considered plausible by clinical experts. According to the company, the hazard pattern (constant) was also more in line with clinical expectations (increasing hazards) than some of the other distributions.

*Progression-free survival*

The company selected dependent exponential models. According to the company, the proportional hazards assumption could be supported. The exponential model had the best statistical fit.

*Time-to-treatment-discontinuation*

The company selected dependent exponential models. According to the company, the proportional hazards assumption could be supported. The exponential model had the best statistical fit.

Treatment waning was explored in scenario analysis on the ibrutinib comparison by setting the hazard ratios equal to 1.

#### 4.2.6.2 Zanubrutinib versus DRC comparison

*Overall survival*

The company selected dependent gamma models. According to the company, the proportional hazards assumption could be supported. The gamma model had the second best statistical fit (after the exponential), and its mean OS was considered plausible by clinical experts. The monotonically increasing hazard pattern was also more in line with clinical expectations than some of the other distributions (the only other distribution with monotonically increasing hazards was the Gompertz model and this had a very similar statistical fit).

*Progression-free survival*

The company selected dependent exponential models. According to the company, the proportional hazards assumption could be supported. The exponential model had the best statistical fit.

*Time-to-treatment-discontinuation*

The company selected an independent exponential model based on statistical fit and to be in line with the selected model for PFS. No survival analysis was performed for BR and DRC: these treatments were assumed to be administered during a fixed time period of six months or terminated sooner if the patient had progressed or no longer tolerated the treatment.

#### 4.2.6.3 Zanubrutinib versus BR comparison

*Overall survival*

The company selected independent Weibull and exponential models for BR and zanubrutinib respectively based on clinically plausible mean OS for both treatment arms and clinically plausible hazard patterns. The proportional hazards assumption was neither fully supported by inspection of hazard plots nor ruled out. However, the company assessed that independently fitted models exhibited better external validity to UK chemo-immunotherapy second- and third-line treatment patients given that BR patients had a median of two prior treatments; and the clinical expectations for the zanubrutinib arm with a median of one prior treatment. The independent Weibull used for the BR arm also had a clinically plausible hazard pattern.

*Progression-free survival*

The company selected dependent exponential models. According to the company, the proportional hazards assumption could be supported. The exponential model had the best statistical fit (BIC).

*Time-to-treatment-discontinuation*

The company selected an independent exponential model based on statistical fit and to be in line with the selected model for PFS.

**ERG comment:** The main concerns of the ERG relate to: a) data immaturity; b) mortality hazards dropping below those of background mortality; c) the use of only patients with *MYD88*$^{MUT}$; d) assumed lifelong treatment effectiveness; e) choice of some survival models and f) treatment effectiveness being analysed for the different comparisons separately.

a) The main issue in this submission is the immaturity of the available data, as was acknowledged by the company. The partitioned survival analysis chosen by the company relies on estimates for PFS and OS, secondary and exploratory endpoints respectively. Only a small number of PFS and OS events had occurred at the time of this appraisal (for zanubrutinib 12.7% had progressive disease during follow-up and 2% had died, Table B.2.11 of the CS[4]) and many patients were censored (e.g. only one out of 102 patients still at risk of OS event after 30 months, clarification response to B7, Figure 57[6]). It is therefore extremely difficult to make long-term predictions. This difficulty is illustrated in the resulting discrepancy in OS and PFS curves when different independent parametric distributions are fitted (see Figures 42-51 in clarification response to B7[6]). For instance, for the comparison with DRC, Figure 4.2 shows PFS results for the zanubrutinib arm (matched to DRC) with different distributions: at 30 years, PFS may be as high as approximately 90%, or as low as approximately 20%. Similarly, for the comparison with DRC, Figure 4.3 shows that OS may be as high as approximately 95% (Gompertz) or as low as approximately 40% (exponential) at 30 years. It is of note that these extreme differences do not necessary lead to an extreme impact on the ICER, nor on the median and mean survival estimates (for example median OS of 15.91 years for the independent Gompertz, and 15.29 years for the independent exponential in the zanubrutinib arm). This is because the hazards of all survival models fall below background mortality hazards after a certain point in time and background mortality is then assumed to apply. Because of background mortality over-riding the OS distributions, it is likely that the driving factor in the model is short-to-medium term OS and the timepoint background mortality takes over in the zanubrutinib arm, rather than long-term extrapolation. These timepoints differ for each distribution in each comparison but range between seven to 10 years and two to seven years with jointly-fitted models and independently fitted models respectively (for zanubrutinib) for the DRC comparison; and five to 12 years and baseline-19 years with jointly-fitted and independent models respectively for the BR comparison (Tables 43 and 37 in the clarification letter response[6]).

**Figure 4.2: PFS extrapolation (independent) for zanubrutinib matched to DRC**

**Figure 4.3: OS extrapolation (independent) for zanubrutinib matched to DRC**

In the absence of mature evidence, the company relied heavily on external validation efforts. The European Chart Review study used in NICE TA491[46] highlighted considerable country-specific OS differences. In addition, the company also considered the ESMO clinical practice guidelines for WM[47] and the company's phase 1/2 BGB-3111-AU-003 trial of zanubrutinib in their external validation efforts. In response to clarification question B5, the company provided additional external validation with the first WMUK registry report from the Rory Morrison Registry (n=579 from 19 hospitals across the UK [1]), which showed an observed median OS of 14.6 years for BTK

inhibitors, which the company considered similar to the median OS estimates for zanubrutinib based on their base-case assumptions (gamma model for DRC resulting in a median of 14.6 years and exponential for BR comparison resulting in a median of 15.29 years, both after adjusting for background mortality). The company also noted that a lack of reporting on patient characteristics (e.g., year of diagnosis, number of prior line[s] of treatment, other key prognostic factors) and treatment pattern (e.g., proportion of patients treated with BTK inhibitor) specifically for patients aged ≥65 years meant that this comparison should also be interpreted with caution.

The ERG considered that, given the extreme data immaturity and given that median follow-up was longer in BGB-3111-AU-003 than in ASPEN (48 months vs 19 months), BGB-3111-AU-003 in particular may be helpful in adding information on PFS and OS hazard patterns in the long run and requested that the company perform survival analysis using external data, for example using the method by Soikkeli et al..[48] The company provided a comparison of OS and PFS KM curves between BGB-3111-AU-003 and ASPEN, which showed that OS was comparable but that the PFS curve was lower in BGB-3111-AU-003 than in ASPEN (that is patients progressed quicker in BGB-3111-AU-003 than in ASPEN). There were differences in populations of BGB-3111-AU-003 compared with ASPEN that may explain this: a significantly higher proportion of treatment-naïve patients, lower proportion of patients with *MYD88*$^{MUT}$, and a higher proportion of patients with CXCR4 status. It was unclear how these differences would impact time to PFS. The ERG acknowledges that indeed data were still immature also in BGB-3111-AU-003 and that it was difficult to draw conclusions from this. The ERG was not able to resolve the uncertainty caused by data immaturity.

b) The hazards of all survival models fall below background mortality hazards and background mortality is then assumed to apply. The timepoints at which this occurs range between seven to 10 years and two to seven years with jointly-fitted models and independently fitted models respectively (for zanubrutinib) for the DRC comparison; and five to 12 years and baseline-19 years with jointly-fitted and independent models respectively for the BR comparison (Tables 43 and 45 in the clarification letter response[6]). The company could not provide any evidence to support mortality hazards dropping below general population mortality hazards. No expert opinion was provided on this in particular (only that experts expected monotonically increasing hazards). Upon request, the company assessed the impact of summing up model hazards and background mortality hazards in scenario analysis: this increased the ICER substantially in both BR and DRC (Table 58 of clarification letter response[6]) and illustrates that uncertainty about long-term hazards could be a model driver. The ERG questioned whether this is simply an artifact of data immaturity, or whether low mortality hazards in the long run indicate that there is a subgroup of patients with WM that are at particular risk of dying in the first years into the modelled disease trajectory, whilst the average patient has closer to normal life expectancy. No information to support either hypothesis was available.

c) The company used the ITT population for their analyses, which only contained patients with *MYD88*$^{MUT)}$ (L265P point mutation in myeloid differentiation primary response gene 88). The prevalence of this mutation according to the company was 90% of patients with WM. The company also stated that other MYD88 mutations or a wild-type MYD88 gene (*MYD88*$^{WT}$) may be found in 5–10% of patients. Cohort 2 of the ASPEN trial collected data on patients with *MYD88*$^{WT}$ who received treatment with zanubrutinib (no comparison). There are a further 1-2% of patients who have non-L265P mutation variants. The ERG was concerned that results of cost effectiveness analyses based on Cohort 1 only might not be generalisable and requested a scenario with a pooled analysis of both cohorts. The company provided this in response to question B6 by performing new MAICs, updating HRQoL inputs and performing cost effectiveness analyses with these inputs. The

company showed that cost effectiveness results of pooled Cohort 1 and 2 were relatively close to results of Cohort 1, if slightly lower for the pooled analysis. Of note, the company did not state whether the analysis was weighted to reflect the mix of patients in clinical practice (i.e. 90% of *MYD88*$^{MUT}$ and 5-10% of *MYD88*$^{WT}$) and hence the ERG assumes that this weighting did not occur and the weight was instead determined by patient numbers in the cohorts (n=28 in Cohort 2 and n=102 in Cohort 1). Furthermore, the ERG considered that the cost of testing should be added if this had not been standard practice in the England & Wales NHS. The company stated in response to clarification question B6 that "according to clinical experts, testing for MYD88 mutation is the standard of care at the majority of the 24 British WM centres (covering 90% of all WM patients since 2016 in the UK)".[6] The ERG considers that the generalisability of the cost effectiveness analysis in Cohort 1 to the overall population is still unclear, because of the mix of mutations in the comparator arm. In addition, the cost effectiveness in the overall population (for example including patients with other variants) remains difficult to assess.

d) The ERG considered that the assumption of lifelong treatment effectiveness may not be warranted and requested that the company implement treatment effectiveness waning in the model. The company implemented treatment waning for both PFS and OS at different time points, using hazard ratios of 1 from the chosen time point onwards. Results of these scenarios showed that this was highly influential (Table 59 of the clarification response[6]). The ERG, in line with previous appraisals adopted treatment waning at five years in its base-case but acknowledges that this is arbitrary.

e) The company were very clear in their presentation of survival analysis. Nevertheless, the ERG had some concerns regarding some model choices for the base-case. One was that the company used differential OS (dependent gamma) and PFS (exponential) models for the DRC comparison. The company stated in response to the POC letter that there was no clinical reason for which PFS and OS need to follow the same pattern. The ERG considers that this may indeed be the case given the long-term nature of the disease, but in a scenario the ERG explored the use of the exponential model for OS to be in line with PFS because it had the best statistical fit. It has to be noted that this model does not fulfil the criterion of monotonically increasing hazards and therefore is used only in the ERG scenario.

   Second, model choice in the BR comparison was difficult. As with the DRC comparison, statistical fit was similar between models but resulting survival curves looked vastly different. Again, the effect on mean and median OS in the zanubrutinib arm was limited due to background mortality hazards applying once OS model hazards dropped below background mortality, but for BR, the range of median OS estimates was wide ([redacted] years for jointly fitted models and [redacted]-[redacted] for independently fitted models). The company considered mainly expert opinion to inform model choice and cited an expected OS of [redacted] and [redacted] for second-line and third-line treatment chemotherapy respectively (which is in line with evidence from WMUK registry report from the Rory Morrison Registry). It is however unclear whether these survival estimates for chemotherapy apply to BR, DRC or a mix of treatments and whether BR and DRC would be expected to be similar or not. The ERG therefore considered it questionable whether these estimates should be used to guide survival model choice. Second, the company considered which of these estimates was more appropriate: in the BR population, patients had received a median of two prior treatments, but in the ASPEN population a median of one prior treatment. The company decided that it was more appropriate to take ASPEN and one prior treatment line as guiding the choice of OS benchmark, which is possibly conservative but not necessarily when dependent models are fitted. However, the ERG considers that it should probably be the matched population that guides this, though the caveats previously mentioned apply (available estimates are for a mix of different chemotherapy regimens). The company chose the independent Weibull model for BR and the independent

exponential model for zanubrutinib to model OS and explored dependent Weibull and gamma in scenarios (both resulting in slightly lower OS estimates for BR, which may be appropriate). In conclusion, owing to data immaturity and the long-term nature of the disease it is very difficult to choose survival models and different models should be considered. The ERG retained the company's selection in the base-case but also explored the dependent gamma in a scenario.

f) The ERG is concerned that treatment effectiveness was analysed for the different comparisons separately. This meant that no fully incremental cost effectiveness analyses were performed, due to differences in the populations between the BR and DRC MAIC in several characteristics: age, differences in measurement of outcomes, the years in which studies were performed, presence of extramedullary disease, amongst others, which resulted in different estimates of costs and effects for zanubrutinib. Whilst the ERG acknowledges the limitations with the data, it considers the use of different analyses as problematic as these do not allow for fully incremental comparison of zanubrutinib, DRC and BR.

### 4.2.7 Adverse events

In the original CS only AEs of Grade ≥3 which occurred in ≥5% of the trial populations were included in the company submission. The company attempted to comply with the request of the ERG to include AEs of Grade ≥3 which occurred in ≥1% of the population. Table 4.5: Adverse events provides the AEs included after the CQs. The incidence of relevant AEs was applied based on the ASPEN trial [2]. Where incidence for specific AEs were not reported, it was assumed that patients did not experience these AEs.

**Table 4.5: Adverse events**

| | **AE incidence, %** | | | | **Duration, days** |
|---|---|---|---|---|---|
| | [redacted] | **BR (N=71)** | [redacted] | **DRC (N=72)** | **ASPEN Safety Analysis Set (N=199)** |
| **Reference** | [redacted] | Tedeschi et al. 2015[32] | [redacted] | Dimopoulos et al. 2007[33] | ASPEN IPD |
| **Anaemia** | [redacted] | NR | [redacted] | NR | 17.0 |
| **Hypertension** | [redacted] | NR | [redacted] | NR | 20.9 |
| **Neutropenia** | [redacted] | 35.21 | [redacted] | 10.00 | 10.9 |
| **Pneumonia** | [redacted] | 5.63 | [redacted] | NR | 21.3 |
| **Thrombocytopenia** | [redacted] | NR | [redacted] | 0.00 | 28.8 |
| **Nausea** | [redacted] | NR | [redacted] | 0.00 | 5.0 |
| **Vomiting** | [redacted] | NR | [redacted] | 0.00 | 5.0 |
| **Headache** | [redacted] | NR | [redacted] | 2.78 | 6.7 |
| **Hypotension** | [redacted] | NR | [redacted] | 5.56 | 0.0 |
| **Sepsis** | [redacted] | 1.41 | [redacted] | NR | 5.0 |

Source: Response to Clarification Letter
Abbreviations: BR = bendamustine and rituximab; DRC = dexamethasone, rituximab and cyclophosphamide; IPD = individual patient-level data; N = number of patients evaluable; $n_{eff}$ = effective sample size; NR = not reported

**ERG comment:** The main concerns of the ERG relate to the restrictive inclusion criteria for AEs.

In the original company submission the company included only AEs of Grade ≥3 which occurred in ≥5% of the population. The ERG questioned whether this would exclude relevant AEs that would occur

in a smaller percentage of the population. Therefore, the ERG asked the company to provide an updated model including all AEs of Grade ≥3 which occurred in ≥1% of the population in clarification question B10. In response. the company provided an updated model in which the AEs were included with of Grade ≥3 and incidence of ≥1%. However, the ERG is unsure that all relevant AEs were included in the model given that the ERG identified additional AEs in the CS which occurred in more than 1% of the population and are severe:

a) Pleural effusion occurred in ≥1% of the population according to CS Table B.2.23 and is a serious adverse event according to Table B.2.41.
b) Febrile Neutropenia occurred in ≥1% of the population of a treatment arm and is a Grade ≥3 adverse event according to CS Table B.2.23.

It is unclear to the ERG why these were not included in the updated AEs.

The inclusion of AEs occurring in 1–5% of patients in any treatment arms had minimal impact on the ICERs [redacted] vs. [redacted] in the CS base-case for BR and [redacted] vs. [redacted] in the CS base-case for DRC). However, for completeness, the ERG included these updated AEs (i.e. Grade ≥3 AEs in ≥1% of patients; as provided in the company scenario) in its base-case.

### 4.2.8 Health-related quality of life

The utility analysis was performed using EQ-5D-5L data from the ASPEN trial[2]. In total, 193 patients (who completed at least one EQ-5D-5L questionnaire) were included. During the first 48 weeks, data was collected every 12 weeks and following this, every 24 weeks. EQ-5D-5L measurements were mapped to EQ-5D-3L values using the crosswalk described by van Hout.[49] and valued using the Dolan algorithm[50]. The average utility value before progression (0.791) was used as the progression-free health state utility (see Table 4.6).

The company stated that due to the low number of observations from patients who progressed, the valuation of post-progression utility was not feasible based on the ASPEN trial data. Therefore, following the example of the company submissions for TA429 and TA491, a utility decrement of 0.1 was applied to calculate the post-progression health state utility, resulting in a utility of 0.691.

**Table 4.6: Health state utility values**

| Health state | Utility value | Reference | Justification |
|---|---|---|---|
| **Progression-free survival** | 0.791 | ASPEN IPD[2] | |
| **Post-progression survival** | 0.691 | TA491[46] | Assuming a utility decrement of 0.1, in line with previous CSs for TAs |
| Source: CS Table B.3.22 | | | |

A linear mixed effect model for repeated measures was used to model utility considering correlation between repeated measures. Out of three fitted models, one was chosen based on its goodness of fit measured by the AIC and BIC. This model presented treatment, age, sex and number of completed visits to date as independent variables. The model is used to model treatment-specific PFS based on differences between the zanubrutinib and ibrutinib treatment arms in a scenario analysis.

#### 4.2.8.1 Disutility values

AE disutilities were applied in the first model cycle as the sum product of AE disutilities, incidence and duration. For all AEs except for hypertension the source of the disutility value was TA491.[46] For hypertension, based on TA429[51] the disutility was assumed to be equal to pneumonia.

**ERG comment:** The main concerns of the ERG relate to: a) PFS utility higher than general UK population values; b) the value and standard error implemented for post-progression utility is not evidence-based; c) how missing data were handled when estimating utilities from the ASPEN trial; and d) the assumption of equal utilities for comparators.

a) In the original CS, PFS utility was set at 0.791. According to a report by Public Health England [52] health-related quality of life in people aged 65 and over was 0.735 for the general population in the UK and therefore lower than that of patients with WM in the model. When asked to justify why the PFS utility used in the model was higher than that in the average of the general population of the same age-group, the company responded that populations could differ as there are (1) natural differences between clinical trial settings and real-world settings and (2) differences in geographic locations between the ASPEN trial and the UK population. Hence, it is unknown to the ERG to what extent the utility values used in the model are applicable to the UK population.
   In response to clarification question B12, the company did accommodate age-related utility decrease relying on the equation from Ara and Brazier [53]. To this end, the utility of the general population was estimated by age and then used to derive utility multipliers over time, using the following equation: *General population utility value = 0.9508566 + 0.0212126*male – 0.0002587*age in years – 0.0000332*age*$^2$. This adjustment was incorporated in the ERG base-case.
b) The utility for the post-progression population was not evidence-based. Due to a lack of data from the ASPEN trial, the company used the progression utility decrements from TA491[3] and TA502[54] to calculate the quality of life of patients in the post-progression health state. In reference to the company submissions of these appraisals the current model therefore applies a utility decrement of -0.1 to the pre-progression health state to calculate the post-progression health state. The final appraisal documents of both TA491 and TA502 note that the utility decrement for the post-progression state applied in the respective company submissions was too small. The post-progression state was therefore decreased to 0.6 in both previous technical appraisals. In TA502 this corresponds to a utility decrement of 0.18 from the pre-progression health state to the post-progression health state.
   In addition, in the economic model the standard error for the post-progression utility is the same as the standard error for progression free utility. It may be argued that the uncertainty around post-progression utility is larger due to the uncertainty around its estimation. In response to clarification question B13, the company argued that implementing a larger standard error was arbitrary and that the scenario analyses around the post-progression utility that they had conducted were more appropriate. The ERG disagrees with this judgement as the post-progression utility value is already set in an arbitrary manner. Reflecting the resulting uncertainty in the probabilistic analysis is preferrable to just reflecting it in a scenario analysis. Hence, in its base-case analysis, the ERG applied a utility decrement of 0.18 in line with TA491 and TA502 but remains concerned that this is not evidence-based and hence uncertain. The ERG did not modify the associated standard error.
c) In response to clarification question B11d,[6] which requested clarification on how missing data was handled and what the potential implications of the company's approach were, the company answered that they assumed that responses were missing at random and therefore did not impute missing data. This position was restated by the company in response to CQ 11f which questioned

the likely causes of missing data, to which the company responded stating that the causes of missing data were 'non-attendance at scheduled meetings' or 'non-completion of all measurements on the EQ-5D-5L scale'. This response gives little information as the underlying reasons for non-attendance and non-completion are not given. The ERG finds the assumption of data 'missing at random' questionable as an alternative plausible explanation is that missing data is caused by patients feeling unwell and therefore not filling out the EQ-5D. If the company erroneously assumes data missing at random instead of some of the data missing because patients feel unwell, this could bias QALY measurements upwards, thereby benefitting the most effective treatment in the model. To reflect uncertainty around the missing data, in clarification question B11g to j the ERG requested the company to recalculate utility estimates while imputing missing values using multiple imputation, providing new estimates and an updated model.[6] The company did not comply with this request, citing that the missing rate of 0% and 38.4% per measurement point would likely not affect utility estimates. This adds to the unexplored uncertainty around outcomes in this model.

d) Health-related quality of life was assumed to be the same between all comparators. Utility estimates derived from the ASPEN trial were not matched to adjust for possible differences in population between the ASPEN trial and the relevant studies for the comparison to BR and DRC. While, due to a lack of data, matching treatment populations to adjust utility estimates is likely not feasible, this adds to the unexplored uncertainty in the economic model.

### 4.2.9 Resources and costs

Costs included in the model were treatment costs, administration costs, miscellaneous, terminal care and adverse event costs.

Unit prices were based on the National Health Service (NHS) reference prices,[55] British National Formulary (BNF),[56] Personal Social Services Research Unit (PSSRU).[57]

#### 4.2.9.1 Treatment costs

Table 4.7 presents the direct treatment costs. Administration cost differed per treatment as no administration costs were applied for zanubrutinib, ibrutinib and cyclophosphamide as this is given orally and £336.14 per administration of other drugs were applied for bendamustine, rituximab and dexamethasone as this was given intravenously. In the model, BR was supposed to be given for six cycles, but was applied for seven cycles. DRC was applied in the model for five model cycles. However, it was unclear how often DRC was supposed to be given as the underlying reference states in its abstract that the treatment was supposed to be given for six months every three weeks, while the main body of text states that the treatment was given for six courses every three weeks (which amounts to four months).

**Table 4.7: Direct treatment costs**

| Treatment | Drug | Dosage | Price per dose (no vial sharing) | Administration costs per cycle (min. administrations) | Administration costs per cycle (max. administrations) | Total cost per applied treatment cycle (average) |
|---|---|---|---|---|---|---|
| **Zanubrutinib (comparison BR)** | **Zanubrutinib** | 312 mg, Daily | £160.41 | £- | | [redacted] |
| **BR** | **Bendamustine** | 167 mg, 2 treatments every cycle | £541.20 | £672.28 | | £2,671.11 |
| | **Rituximab** | 500 mg, 2 treatments every cycle | £1,734.05 | £336.14 | | |
| **Zanubrutinib (comparison DRC)** | **Zanubrutinib** | 312 mg, Daily | £160.41 | £- | | [redacted] |
| **DRC** | **Dexamethasone** | 20 mg, Once every 3 weeks | £11.99 | £336.14 | £672.28 | £2,443.60 |
| | **Rituximab** | 500 mg, Once every 3 weeks | £1,734.05 | £336.14 | £672.28 | |
| | **Cyclophosphamide** | 372 mg, 5 treatments every 3 weeks | £11.04 | £- | | |

Source: CS Model values

#### 4.2.9.2 Miscellaneous costs

Due to a lack of published studies reporting the use of other healthcare-related costs in patients with WM, the frequency of resource use was based on the resource use implemented for the model in TA491.[3] These costs (Table 4.8) were based on the NHS reference cost of 2018-2019.[55]

**Table 4.8: Miscellaneous costs**

| | Frequency per year | | | Reference | Unit cost, £ | Reference |
|---|---|---|---|---|---|---|
| | Year 1–2 | Year 3–5 | Year 6+ | | | |
| **Full blood count** | 5 | 4 | 3 | NICE TA491 | 2.87 | NHS reference cost 2018-2019, DAPS05 Haematology[55] |
| **Immunoglobulin** | 5 | 4 | 3 | NICE TA491 | 6.72 | NHS reference cost 2018-2019, DAPS06 Immunology[55] |
| **Chemistry** | 5 | 4 | 3 | NICE TA491 | 1.14 | NHS reference cost 2018-2019, DAPS04 Clinical biochemistry[55] |
| **Haematologist** | 5 | 4 | 3 | NICE TA491 | 135.59 | NHS reference cost 2018-2019, WF01A Clinical haematology, consultant-led, non-admitted face to face follow-up[55] |
| **Plasma viscosity** | 5 | 4 | 3 | NICE TA491 | 6.75 | NHS reference cost 2018-2019, DAPS06 Immunology[55] |
| **Paraprotein** | 5 | 4 | 3 | NICE TA491 | 1.13 | NHS reference cost 2018-2019, DAPS04 Clinical biochemistry[55] |

Source: CS Table B.3.27

Abbreviations: DAPS = Directly Accessed Pathology Services; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; TA = technology appraisal; WM = Waldenström's macroglobulinaemia

Patients who progressed were assumed to receive subsequent treatment. Due to immature data in the ASPEN trial, subsequent treatment costs from TA491 were applied for zanubrutinib and ibrutinib. Based on TA491, 86% of progressed patients received subsequent care. For patients receiving BTK inhibitors as a first-line treatment ibrutinib use until death was assumed post-progressions.

Upon death, terminal care costs (£7,978.35) were applied.

#### 4.2.9.3 Event costs

AE costs were applied as a one-off cost in the first cycle based on the sum product of incidence and cost of all included AEs. They were split up in two types of costs - "Infections" and "AEs other than infections". The source of these costs was the NHS reference cost 2018-2019.

**Table 4.9: Adverse event costs**

| AE type | Unit cost, £ | Source |
|---|---|---|
| Infections (mainly sepsis)a | 1,481.76 | NHS reference cost 2018-2019[55] |
| AEs other than infectionsb | 179.94 | NHS reference cost 2018-2019[55] |

Source: CS Table B.3.28
Abbreviation: AE = adverse event; NHS = National Health Service
[a] The cost was estimated based on the weighted average of costs for Infections or other complications of procedures, without interventions, with CC Score 0-<4 (codes: WH07F – WH07G in NHS reference cost 2018-2019)
[b] The cost was estimated based on the weighted average of costs for Non-Admitted Face to Face Attendance - Clinical Haematology (codes: WF01A, WF01B, WF01C, WF01D, WF02A, WF02B, WF02C, WF02D in NHS reference cost 2018-2019)

**ERG comment:** The main concerns of the ERG relate to: a) AE costs not being broken down further than "infections" and "non-infections" and b) mistake in the application of the BR and DRC treatment and c) the application miscellaneous costs after progression.

a) In the CS all AE costs are divided into 'infections' and 'AEs that are not infections'. Even though pneumonia is an infection, it was categorised as an AE that is not an infection in the original CS[4]. Costs are not broken down further, and the 'infection' cost were not applied to any AE in the original CS. The assumption that AE costs are accurately reflected using 'infections' and 'AEs that are not infections' was questioned in clarification question B15.[6] In their response, the company implemented the costs for pneumonia correctly but did not break the costs down any further as requested per clarification question B15b.[6] This imprecision adds to the uncertainty in the model. The impact of this uncertainty is however likely minor.
b) In the Excel model BR and DRC were applied for five and seven model cycles respectively. BR was supposed to be implemented for six cycles, while the underlying reference[33] used for DRC was unclear on the number of treatments that was supposed to be implemented: while the abstract of the reference stated that DRC was supposed to be given every three weeks for six months, the main body of the text stated that DRC was supposed to be given every three weeks for six courses (four months). The ERG corrected this error in their base-case with the assumption of DRC being given for four cycles and BR being given for six cycles.
c) No changes to costs were implemented in the model after treatment progression. The ERG questions whether patients who have progressed require more care. As patients receiving zanubrutinib generally enter the post-progression state after patients on BR or DRC, the company's assumption may underestimate the cost of patients treated with BR or DRC more than costs of patients treated with zanubrutinib and may therefore be conservative.

## 5. COST EFFECTIVENESS RESULTS

### *5.1 Company's cost effectiveness results*

The base case cost effectiveness results amount to [redacted] per QALY in the base-case when compared to BR and [redacted] per QALY when compared to DRC in the deterministic analysis. For the probabilistic analysis, the model results amount to [redacted] per QALY when compared to BR and [redacted] per QALY when compared to DRC. The probabilistic CS base-case analyses (1,000 simulations) indicated cost effectiveness acceptability probabilities of 1% at willingness to pay threshold of £30,000 per QALY gained in the comparison to BR. When compared to DRC, the analysis resulted in a cost effectiveness acceptability probability of 0% at a willingness to pay threshold of £30,000.

**Table 5.1: Company's cost effectiveness results base-case**

| Technologies | Total costs (£) | Total LYs | Total QALYs | Incremental costs (£) | Incremental Lys | Incremental QALYs | ICER (£/QALY) |
|---|---|---|---|---|---|---|---|
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £116,902 | 6.55 | 4.96 | | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £139,102 | 7.81 | 5.88 | | | | |

Overall, the technology is modelled to affect QALYs by:

- Increasing survival (OS)
- Increasing progression-free survival

Overall, the technology is modelled to affect costs by:

- Higher unit price than current treatments

The modelling assumptions that have the greatest effect on the ICER are:

- Assumptions regarding subsequent treatments to BR and DRC in the economic model.
- Because of background mortality over-riding the OS distributions, it is likely that the driving factor in the model is short-to-medium term OS and the timepoint background mortality takes over in the zanubrutinib arm, rather than long-term extrapolation. These timepoints differ for each distribution in each comparison but range between 7-10 years and 2-7 years with jointly-fitted models and independently fitted models respectively (for zanubrutinib) for the DRC comparison; and 5-12 years and baseline-19 years with jointly-fitted and independent models respectively for the BR comparison.
- Assumption of the timepoint at which treatment waning is assumed.

### *5.2 Company's sensitivity analyses*

The company performed and presented the results of deterministic sensitivity analyses (DSA) as well as scenario analyses.

The DSA showed that the results were most sensitive to changes in the percentage of the population that receives subsequent treatment as well as the value used for health-related quality of life during progression-free survival and the distribution of age and gender in the population in both BR and DRC.

The company conducted several scenario analyses for BR and DRC. The results showed costs ranging between [redacted] and [redacted] per QALY gained for BR and costs ranging between [redacted] and [redacted] per QALY gained DRC. The three most influential scenarios for both BR and DRC that increased the ICER were:

1.) A decrease of the time horizon to 10 years from 30 years.
2.) Excluding subsequent treatment costs.
3.) Applying cost and effects of subsequent treatment from the previous technology appraisal for ibrutinib in WM.

**ERG comment:** The main concerns of the ERG relate to: a) the company did not perform a fully incremental analysis; b) the number of PSA simulations; and c) some scenarios which were requested during clarification were not provided by the company.

a) Contrary to the final scope issued by NICE, a full incremental analysis of zanubrutinib, BR, and DRC was not performed (see Section 4.2.6). As mentioned in Section 3.3, the company argued that, due to the lack of RCTs comparing zanubrutinib with any comparator other than ibrutinib, and a lack of common comparators for an anchored indirect treatment comparison (ITC), two separate matching adjusted indirect comparison (MAIC) were necessary. This warrants the questions to what extent the two ICERs (i.e. zanubrutinib vs BR and zanubrutinib vs DRC) are applicable to the same population.
b) The model supplied by the company with the original CS is only able to compute 1,000 simulations. Hence, in clarification question B17, the ERG therefore requested the company run the model with 5,000 simulations and provide convergence plots. The company responded by implementing 5,000 simulations and stating that the results were stable. They did not provide convergence plots, pointing out that they could be created when the ERG runs their analyses. The ERG used 5,000 simulations in its ERG base-case which looked stable.
c) Table 5.2 presents analyses which were not conducted by the company upon request in the clarification questions.

**Table 5.2: Requested analyses which were not provided**

| CQ | Requested | Reference | Analysis | Justification for not complying |
|---|---|---|---|---|
| **B1** | Inclusion of different comparators | Please include FR, FCR, Clad-R, ASCT, chlorambucil, rituximab monotherapy and BSC as a comparator in the model and provide a full incremental analysis. | Not provided | "As detailed in Appendix D, Section D.1.2 of the company submission, there was a lack of data to inform the inclusion of non-BR/DRC comparators in the cost-effectiveness analysis. " |
| **B1** | Full incremental analysis | Please include FR, FCR, Clad-R, ASCT, chlorambucil, rituximab monotherapy and BSC as a comparator in the model and provide a full incremental analysis. | Not provided | "...a full incremental analysis was not applicable because the treatment comparisons relied on pairwise comparisons using an MAIC approach, in which the patient populations vary by treatment comparison." |
| **B3** | Provide an alternative scenario in which alternative subsequent treatments have been explored | Based on your response to B3b above, please provide an alternative scenario in which alternative subsequent treatments have been explored. | Not provided | Summarised - Ibrutinib is a relevant treatment and should therefore be included - therefore the exploration of alternative subsequent treatments is not necessary |
| **B4d** | Provide a state transition model | Please provide a STM to inform health state occupancy, HRQoL, and cost and resource use based on response status from the ASPEN trial. And assist in verifying the plausibility of the PSM extrapolations and to address uncertainties in the extrapolation period (NICE DSU TSD 19, recommendation 11). | Not provided | In summary, based on the available data for the comparators listed in the final scope, a PSM relying on an MAIC was developed to compare zanubrutinib to the comparators of interest, relying heavily on clinical expert opinion on the validity of extrapolated long-term survival (see Section B.3.3.2 of the company submission and the response to clarification question B5), whereas a STM was deemed unfeasible. Although an STM was deemed unfeasible for the comparison of zanubrutinib versus the compactors listed in the final scope, with the patient-level ASPEN data for zanubrutinib, an STM is potentially feasible for the zanubrutinib arm alone. |
| **B11j** | Provide an updated economic model | Please provide an updated economic model as well as scenario analysis incorporating the estimated utility | Not provided | Not provided |

| CQ | Requested | Reference | Analysis | Justification for not complying |
|---|---|---|---|---|
| | incorporating more appropriate ways to deal with missing values. | values in response to sub-questions g and h (i.e. utility values estimated stratified for patients being on treatment or not with and without imputation). | | |
| Source: Company's response to clarification questions.[6]<br>CQ = clarification question | | | | |

## *5.3 Model validation and face validity check*

### 5.3.1 Face validity assessment

It is stated in the CS that the in- and exclusion of comparators and the clinical trials used in the CS were informed by a combination of WM treatment guidelines per ESMO, International Workshop on Waldenström's Macroglobulinemia, and Australia Medical and Scientific Advisory Group, Real-world treatment patterns according to the UK Rory Morrison Registry, and a medical advisory board meeting in the EU.

### 5.3.2 Internal validation

In the CS, it is stated that various parts of the model were subjected to internal validity checks. The company categorized these validity tests into logical tests (e.g., set all utility values to one and see whether QALYs and LYs are equal) and technical implementation tests (e.g., check whether the half-cycle correction was appropriately applied). It is not mentioned in the CS who performed these validity checks. Moreover, the selection of parametric models for PFS, OS, and TTD survival was based on AIC and BIC fit statistics and visual inspection.

In response to CQ B5, the company completed the TECH-VER checklist, which is a comprehensive checklist for the technical verification of decision analytical models.[58]

### 5.3.3 External validation

In the CS, it is stated that the selection of parametric models for PFS, OS, and TTD survival was based on visual inspection assessed by published estimates and clinical experts' opinions on clinical plausibility of the extrapolated survival. As mentioned in Section 4.2.6, the company also considered the ESMO clinical practice guidelines for WM and the company's phase 1/2 BGB-3111-AU-003 trial of zanubrutinib in their external validation efforts. In response to clarification question B5, the company provided additional external validation with the first WMUK registry report from the Rory Morrison Registry (n=579 from 19 hospitals across the UK).

### 5.3.4 Comparisons with other technology appraisals

In the initial CS, no cross-validation to other technology appraisals (e.g. NICE TA491) was performed in terms of outcome parameters (e.g. mortality rates, QALYs, or costs per cycle and over the full-time horizon). However, in response to CQ B21, the company provided a comparisons with TA491 (ibrutinib for treating WM) and TA502 (ibrutinib for treating MCL).

**ERG comment:** The main concerns of the ERG relate to the large discrepancy between the deterministic ICER and the probabilistic ICER

a) Although the company mentions in the CS that "the mean probabilistic results are aligned with the deterministic results for all three treatment comparisons, indicating the model is structurally stable", the ERG noted a large difference between the deterministic ICER and the probabilistic ICER for BR (difference of [redacted]) and DRC (difference of [redacted]). As the company did not provide convergence plots, only pointing out that they could be created when the ERG runs their analyses, the ERG implemented convergence plots in the model which demonstrated relatively stable results after approximately 2,000 runs. However, the resulting ICER were structurally higher compared to the deterministic ICERs. Although the ERG could not exactly pinpoint the cause of this difference, there appears to be a structural difference between the deterministic and probabilistic analysis which questions the internal validity of the modelled results.

**Figure 5.1: Convergence plot displaying incremental costs and QALYs for zanubrutinib vs BR**

**Figure 5.2: Convergence plot displaying incremental costs and QALYs for zanubrutinib vs DRC**

# 6. EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

## *6.1 Exploratory and sensitivity analyses undertaken by the ERG*

Table 6.1 summarises the key issues related to the cost effectiveness categorised according to the sources of uncertainty as defined by Grimm et al. 2020:[59]

- Transparency (e.g. lack of clarity in presentation, description, or justification)
- Methods (e.g. violation of best research practices, existing guidelines, or the reference case)
- Imprecision (e.g. particularly wide confidence intervals, small sample sizes, or immaturity of data)
- Bias and indirectness (e.g. there is a mismatch between the decision problem and evidence used to inform it in terms of population, intervention/comparator and/or outcomes considered)
- Unavailability (e.g. lack of data or insight)

Identifying the source of uncertainty can help determine what course of action can be taken (i.e. whether additional clarifications, evidence and/ or analyses might help to resolve the key issue). Moreover, Table 6.1 lists suggested alternative approaches, expected effects on the cost-effectiveness, whether it is reflected in the ERG base-case as well as additional evidence or analyses that might help to resolve the key issues.

Based on all considerations in the preceding Sections of this ERG report, the ERG defined a new base-case. This base-case included multiple adjustments to the original base-case presented in the previous sections. These adjustments made by the ERG form the ERG base-case and were subdivided into three categories (derived from Kaltenthaler 2016):[60]

- Fixing errors (FE) (correcting the model where the company's submitted model was unequivocally wrong)
- Fixing violations (FV) (correcting the model where the ERG considered that the NICE reference case, scope or best practice had not been adhered to)
- Matters of judgement (MJ) (amending the model where the ERG considers that reasonable alternative assumptions are preferred)

### 6.1.1 ERG base-case

Adjustments made by the ERG, to derive the ERG base-case (using the CS base-case as starting point) are listed below. Table 6.2 shows how individual adjustments impact the results plus the combined effect of all abovementioned adjustments simultaneously, resulting in the ERG base-case. The 'fixing error' adjustments were combined and the other ERG analyses were performed also incorporating these 'fixing error' adjustments given the ERG considered that the 'fixing error' adjustments corrected unequivocally wrong issues.

**Fixing errors**

1. The model of the company submission applied DRC for five and BR for seven model cycles. BR was supposed to be given for six treatment cycles, which was implemented in the ERG base-case. Given that it was unclear to the ERG for how many cycles DRC was supposed to be given, the ERG chose to implement BR for six treatment cycles with the duration of three weeks in the ERG base-case (Section 4.2.9).

**Fixing violation**

2. Ibrutinib was excluded from the model as direct comparator and as subsequent treatment (Section 4.2.4).

**Matters of judgement**

3. Assuming similar relative dose intensity rates of 97.5% for BR, DRC, and zanubrutinib instead of 100% for both (Section 4.2.4).
4. In the CS base-case no treatment waning was assumed. Although the ERG acknowledges the difficulties in empirically assessing treatment-waning, a five-year cut-off, as adopted in prior related appraisals (e.g. TA627) was assumed (Section 4.2.6).
5. Inclusion of all AEs of Grade ≥3 which occurred in ≥1% of the population, instead of ≥5% of the trial populations (Section 4.2.7).
6. The use of age-adjusted utility values instead of (Section 4.2.8).
7. A utility decrement of 0.18 in line with TA491 and TA502 instead of a utility decrement of 0.1 (Section 4.2.8).

### 6.1.2 ERG exploratory scenario analyses

The ERG performed the following exploratory scenario analyses to assess the impact of alternative assumptions conditional on the ERG base-case.

1. OS scenarios: DRC comparison: use dependent exponential for OS (to be in line with PFS), BR comparison: dependent gamma for OS
2. PFS scenarios: DRC comparison: dependent Gompertz for PFS, BR comparison: dependent lognormal for PFS
3. A scenario assuming no subsequent treatments.

### 6.1.3 ERG subgroup analyses

No subgroup analyses were performed by the ERG.

**Table 6.1: Overview of key issues related to the cost effectiveness (conditional on fixing errors highlighted in Section 6.1)**

| Issue nr. | Key issue (Health economic issues) | Section | Source of uncertainty | Alternative approaches | Expected impact on ICER[a] | Resolved in ERG base-case[b] | Required additional evidence or analyses |
|---|---|---|---|---|---|---|---|
| #6 | The choice of a partitioned survival model and its underlying assumptions. | 4.2.2 | Methods | The ERG suggested to explore a STM to validate outcomes of the current model | Unknown | No | Model using STM approach |
| #7 | The model does not include all comparators mentioned in the NICE scope. | 4.2.4 | Methods | Based on the company's response to question A26, the ERG argues that the company could have done exploratory analyses for the comparisons of zanubrutinib with FR/FCR and rituximab monotherapy. | Unknown | No | |
| #8 | Ibrutinib should not be included as direct comparator and/or subsequent treatment option in the economic model. | 4.2.4 | Unavailability | The ERG assumed in its base-case that patients initially treated with BR would receive DRC as subsequent treatment and patient initially treated with DRC would receive BR. Patients initially receiving zanubrutinib received subsequent treatment according to the CS base-case (BR for 60.4% of the patients and DRC for 39.6% of the patients). | Unknown | Partly | The company should provide additional evidence (e.g. expert opinion or clinical trials) that gives insight into possible subsequent treatments in absence of ibrutinib in the UK. |

| Issue nr. | Key issue (Health economic issues) | Section | Source of uncertainty | Alternative approaches | Expected impact on ICER[a] | Resolved in ERG base-case[b] | Required additional evidence or analyses |
|---|---|---|---|---|---|---|---|
| #9 | The partitioned survival analysis chosen by the company relies on estimates for PFS and OS, secondary and exploratory endpoints respectively. Only a small number of PFS and OS events had occurred at the time of this appraisal. | 4.2.6 | Imprecision | The ERG was not able to resolve the uncertainty caused by data immaturity, but explored this using extreme model choices in scenarios | Unknown | No | Collection of long-term follow-up data. |
| #10 | Plausibility of OS hazards falling below background mortality hazards. Hazards of all survival models fall below background mortality hazards and background mortality is then assumed to apply. | 4.2.6 | Bias | Additional evidence to support mortality hazards dropping below general population mortality hazards. | Unknown | No | The ERG is questioning whether low mortality hazards in the long run indicate that there is a subgroup of patients with WM that die sooner, but that the average patient has normal life expectancy, but no such information was available. |
| #11 | The use of only patients with *MYD88*$^{MUT}$. | 4.2.3 / 4.2.6 | Imprecision | A weighted analysis to reflect the mix of patients in clinical practice (i.e. 90% of *MYD88*$^{MUT}$ and 5-10% of *MYD88*$^{WT}$). | Unknown | No | A weighted analysis would require information regarding the mix of mutations in the comparator arm. |

| Issue nr. | Key issue (Health economic issues) | Section | Source of uncertainty | Alternative approaches | Expected impact on ICER[a] | Resolved in ERG base-case[b] | Required additional evidence or analyses |
|---|---|---|---|---|---|---|---|
| #12 | Assumption of lifelong treatment effectiveness. | 4.2.6 | Methods | The ERG, in line with previous appraisals adopted treatment waning at 5 years in its base-case, but acknowledges that this is arbitrary. | Increase | Partly | Long-term follow-up data regarding treatment waning over time. |
| #13 | PFS utility higher than general UK population values | 4.2.8 | Bias | Provide evidence on the justification of the PFS utility values in the model. | Likely increase | No | Additional evidence about the health-related quality of life of WM patients in the UK may help to resolve this issue. |
| #14 | The value and standard error implemented for post-progression utility | 4.2.8 | Bias | The ERG implemented a utility decrement of -.18 to stay in line with previous technical appraisals. | Likely increase | Yes | Additional evidence about the post-progression health-related quality of life of WM patients may help to resolve this issue. |
| #15 | Large discrepancy between the deterministic ICER and the probabilistic ICER. | 5.2.4 | Imprecision | The company should make sure all relevant input parameters are adequately captured in the PSA and demonstrate a minimal difference between both the deterministic and probabilistic results or provide an explanation as to why these difference arise. | Unknown | No | Additional examination of the model by the company. |

| Issue nr. | Key issue (Health economic issues) | Section | Source of uncertainty | Alternative approaches | Expected impact on ICER[a] | Resolved in ERG base-case[b] | Required additional evidence or analyses |
|---|---|---|---|---|---|---|---|
| #16 | A full incremental analysis of zanubrutinib, BR, and DRC was not performed | 4.2.6 | Bias | The ERG questions whether a fully incremental analysis could be performed with these comparisons. | Unknown | No | Present results of a fully incremental analysis. |

[a] Likely conservative assumptions (of the intervention versus all comparators); [b] Explored
ERG = Evidence Review Group; FE = Fixing errors; FV = fixing violations; MJ = matters of judgement; ICER = incremental cost effectiveness ratio

### *6.2 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG*

In Section 6.1 the ERG base-case was presented, which was based on various changes compared to the company base-case. Table 6.2 shows how individual changes impact the results plus the combined effect of all changes simultaneously, the probabilistic ERG base-case is presented in Table 6.3. The exploratory scenario analyses are presented in Table 6.4. These are all conditional on the ERG base-case. The analyses numbers in Tables 6.2 and 6.4 correspond to the numbers reported in Section 6.1. Finally, Table 6.4 provides the results of the scenario analysis (described in Section 6.1.3). The submitted model file contains technical details on the analyses performed by the ERG (e.g. the "ERG" sheet provides an overview of the cells that were altered for each adjustment).

**Table 6.2: Deterministic ERG base-case**

| Technologies | Total costs (£) | Total QALYs | Incremental costs (£) | Incremental QALYs | ICER (£/QALY) |
|---|---|---|---|---|---|
| **ERG base-case (ERG_1 -ERG_7)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £ 53,685 | 4.60 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £ 50,562 | 5.40 | | | |
| **Company's corrected base-case (ERG_1 & ERG_2)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £ 53,842 | 4.96 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £ 50,695 | 5.88 | | | |
| **Matter of judgement: Similar dose-intensities (ERG_1, ERG_2 & ERG_3)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £ 53,685 | 4.96 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £ 50,552 | 5.88 | | | |
| **Matter of judgement: Treatment waning (ERG 1, ERG_2 & ERG_4)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £ 53,842 | 4.96 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £ 50,695 | 5.88 | | | |
| **Matter of judgement: Including additional AEs (ERG_1, ERG_2 & ERG_5)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |

| Technologies | Total costs (£) | Total QALYs | Incremental costs (£) | Incremental QALYs | ICER (£/QALY) |
|---|---|---|---|---|---|
| BR | £ 53,842 | 4.96 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £ 50,705 | 5.88 | | | |
| **Matter of judgement: Age-adjusted utilities (ERG_1, ERG_2 & ERG_6)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £ 53,842 | 4.77 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £ 50,695 | 5.62 | | | |
| **Matter of judgement: Post-progression utility decrement (ERG_1, ERG_2 & ERG_7)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £ 53,842 | 4.78 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £ 50,695 | 5.64 | | | |

**Table 6.3: Probabilistic ERG base-case**

| Technologies | Total costs (£) | Total QALYs | Incremental costs (£) | Incremental QALYs | ICER (£/QALY) |
|---|---|---|---|---|---|
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £53,658 | 4.51 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £50,626 | 5.38 | | | |

**Table 6.4: Deterministic scenario analyses (conditional on ERG base-case)**

| Technologies | Total costs (£) | Total QALYs | Incremental costs (£) | Incremental QALYs | ICER (£/QALY) |
|---|---|---|---|---|---|
| **Scenario: Treatment waning 10 years (ERG_1 - ERG_7, ERG_12)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £ 53,685 | 4.60 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £ 50,562 | 5.40 | | | |
| **Scenario: OS - Dependent exponential DRC (ERG_1 - ERG_7, ERG_13)** | | | | | |

| Technologies | Total costs (£) | Total QALYs | Incremental costs (£) | Incremental QALYs | ICER (£/QALY) |
|---|---|---|---|---|---|
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £ 53,685 | 4.60 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £ 50,561 | 5.44 | | | |
| **Scenario: OS - Dependent gamma BR (ERG_1 - ERG_7, ERG_14)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £ 53,563 | 4.23 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £ 50,562 | 5.40 | | | |
| **Scenario: PFS - Dependent Gompertz DRC (ERG_1 - ERG_7, ERG_15)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £ 53,685 | 4.60 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £ 49,908 | 5.48 | | | |
| **Scenario: PFS - Dependent lognormal BR (ERG_1 - ERG_7, ERG_16)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £ 52,651 | 4.82 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £ 50,562 | 5.40 | | | |
| **Scenario: No subsequent treatment (ERG_1 - ERG_7, ERG_17)** | | | | | |
| Zanubrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | £ 32,039 | 4.60 | | | |
| Zanubrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | £ 24,859 | 5.40 | | | |

### *6.3 ERG's preferred assumptions*

The estimated ERG base-case ICER (probabilistic), based on the ERG preferred assumptions highlighted in Section 6.1, was [redacted] per QALY gained for zanubrutinib compared to BR and [redacted] per QALY gained for zanubrutinib compared to DRC.

The probabilistic ERG base-case analyses indicated cost effectiveness probabilities of 0% at willingness to pay thresholds of £20,000 and £30,000 per QALY gained. The most influential adjustments were the

exclusion of ibrutinib as subsequent treatment from the model and the inclusion of treatment waning at a 5-yeat cut-off. . The ICER increased most in the scenario analysis with alternative assumptions regarding the exclusion of subsequent treatments and the scenario regarding alternative survival curves for PFS (i.e. using dependent Gompertz models for DRC and lognormal for BR).

### *6.4 Conclusions of the cost effectiveness section*

The company's modelling approach consisted of a partitioned survival model. Zanubrutinib was considered within the economic evaluation as per the anticipated licenced indication in WM. The comparators considered in the CS were ibrutinib, BR, and DRC. However, NICE's position statement "consideration of products recommended for use in the Cancer Drugs Fund as comparators, or in a treatment sequence, in the appraisal of a new cancer product" states that technologies available through the CDF should not be modelled in treatment sequences. Hence, ibrutinib was excluded from the model by the ERG both as direct comparator as well as subsequent treatment. In contrast to the NICE scope, the model does not include FR, FCR, Clad-R and ASCT (for patients who have had at least one prior therapy), chlorambucil, rituximab monotherapy and BSC (for patients for whom chemo-immunotherapy is unsuitable) as comparators.

In line with its anticipated marketing authorisation and the final scope issued by NICE, zanubrutinib was considered in the cost effectiveness model for the treatment of adult patients with WM previously treated with at least one prior line of therapy, or who are treatment naïve and unsuitable for chemo-immunotherapy. Contrary to the final scope issued by NICE, a full incremental analysis of zanubrutinib, BR, and DRC was not performed. As mentioned in Section 3.3, the company argued that, due to the lack of RCTs comparing zanubrutinib with any comparator other than ibrutinib, and a lack of common comparators for an anchored indirect treatment comparison (ITC), two separate matching adjusted indirect comparison (MAIC) were necessary. This warrants the questions to what extent the two ICERs (i.e. zanubrutinib vs BR and zanubrutinib vs DRC) are applicable to the same population. Moreover, the company used the ITT of the ASPEN trial population for their analyses, which only contained patients with MYD88 mutation. It should be noted, however, that the final scope issued by NICE does not specify any genetic marker and refers to people with the *MYD88*$^{MUT}$ as a relevant subgroup of the population only. The ERG considers that the generalisability of the cost effectiveness analysis in Cohort 1 to the overall population is still unexplored because of the mix of mutations in the comparator arm, but there is no evidence that effectiveness differs between subgroups. As a consequence, the cost effectiveness in the overall population (for example including patients with other variants) remains difficult to assess.

The main issue in this submission is the immaturity of the available data, as was acknowledged by the company. The partitioned survival analysis chosen by the company relies on estimates for PFS and OS, secondary and exploratory endpoints respectively. Only a small number of PFS and OS events had occurred at the time of this appraisal and many patients were censored. It is therefore extremely difficult to make long-term predictions. This difficulty in making predictions is somewhat mitigated by the life expectancy of patients with WM being similar to the average general population life expectancy, once they survived until a certain timepoint. The company's fitted OS time-to-event distributions exhibited hazards that dropped below general population background mortality hazards several years into the model horizon. It was unclear whether these low mortality hazards in the long run indicate that there is a subgroup of patients with WM that die soon after diagnosis, but that the patients alive at later timepoints have similar to normal life expectancy, but no such information was available. Background mortality hazards over-riding the OS distributions meant that the impact of these distributions on model outcomes (including mean survival) was limited, despite the extreme differences in the fitted curves.

The company could not provide any evidence to support mortality hazards dropping below general population mortality hazards. Hence, the ERG is questioning whether low mortality hazards in the long run indicate that there is a subgroup of patients with WM that die sooner, but that the average patient has normal life expectancy, but no such information was available.

According to the ERG, there are several issues that add to the uncertainty of the ICERs reported in the CS, which the ERG was only partly able to incorporate in the ERG base-case. Consequently, the cost effectiveness results, calculated using the economic model submitted by the company, are subject to a large degree of uncertainty. Key uncertainties in this cost effectiveness assessment are, according to the ERG, 1) the choice of a partitioned survival model and its underlying assumptions; 2) the model does not include all comparators mentioned in the NICE scope; 3) ibrutinib should not be included as direct comparator and/or subsequent treatment option in the economic model; 4) the partitioned survival analysis chosen by the company relies on estimates for PFS and OS (secondary and exploratory endpoints) but only a small number of PFS and OS events had occurred at the time of this appraisal; 5) plausibility of OS hazards falling below background mortality hazards; 6) the use of only patients with *MYD88*$^{MUT}$; 7) assumption of lifelong treatment effectiveness; 8) PFS utility higher than general UK population values; 9) the value and standard error implemented for post-progression utility; 10) it is important to mention that the ERG noted a large difference between the deterministic ICER and the probabilistic ICER for both BR and DRC; and 11) a fully incremental analysis was not performed.

In the company base-case (probabilistic), the model results amount to [redacted] per QALY when compared to BR and [redacted] per QALY when compared to DRC. The ERG has incorporated various adjustments to the CS base-case (using the revised economic model with input parameters from the original CS as starting point). However, the ERG considers that there remains substantial uncertainty about the presented cost effectiveness results.

The individual ERG adjustments had large impact on the ICER, ranging from [redacted] per QALY gained to [redacted] per QALY gained for BR and from [redacted] per QALY gained to [redacted] per QALY gained for DRC (deterministic). The estimated ERG base-case ICER (deterministic) was [redacted] per QALY gained for ibrutinib compared to BR and [redacted] per QALY gained for ibrutinib compared to DRC. To demonstrate the discrepancy between the deterministic and probabilistic results: the ERG base-case ICER (probabilistic) was [redacted] per QALY gained for ibrutinib compared to BR and [redacted] per QALY gained for ibrutinib compared to DRC.

The most influential adjustments were the exclusion of ibrutinib as subsequent treatment from the model and the inclusion of treatment waning at a five-year cut-off. The ICER increased most in the scenario analysis with alternative assumptions regarding the exclusion of subsequent treatments and the scenario regarding alternative survival curves for PFS (i.e. using dependent Gompertz models for DRC and lognormal for BR).

It should be reiterated that some of the above mentioned potential biases, especially related to data immaturity and discrepancy between deterministic and probabilistic results could not be explored by the ERG. Consequently, the ICERs reported are subject to great uncertainty.

## 7. REFERENCES

[1] WMUK. *First UK Waldenström's macroglobulinaemia registry report*, 2018 [accessed August 2020] Available from: https://www.wmuk.org.uk/sites/default/files/2020-01/The%20Rory%20Morrison%20WM%20Registry%20Report%20%281%29%20-%20March%202018.pdf

[2] Tam CS, Opat S, D'Sa S, Jurczak W, Lee HP, Cull G, et al. A randomized phase 3 trial of zanubrutinib versus ibrutinib in symptomatic Waldenstrom macroglobulinemia: the Aspen study. *Blood* 2020.

[3] National Institute for Health Care Excellence. Single Technology Appraisal: Ibrutinib for treating Waldenström's macroglobulinaemia [ID884]. 2016 [accessed August 2020]. Available from: https://www.nice.org.uk/guidance/ta491/documents/committee-papers

[4] BeiGene. *Single technology appraisal. Zanubrutinib for treating Waldenström's macroglobulinaemia [ID1427]. Document B: company evidence submission*: BeiGene, 12th March 2021 [accessed 18.3.21]. 164p.

[5] National Institute for Health and Care Excellence. *Zanubrutinib for treating Waldenström's macroglobulinaemia. Appendix B: final scope [PDF]*. London: NICE, October 2020 [accessed 18.3.21]. 5p.

[6] BeiGene. *Single technology appraisal. Zanubrutinib for treating Waldenstrom's macroglobulinaemia [ID1427]: clarification response (version 1.0) [Word document]*: BeiGene, 26th April 2021 [accessed 4.5.21]. 119p.

[7] BeiGene. Zanubrutinib summary of product characteristics. 2020.

[8] National Institute for Health and Care Excellence. *Zanubrutinib for treating Waldenström's macroglobulinaemia [ID1427]: response to comments on the draft scope [PDF provided by NICE]*. London: NICE, Sept 2020 [accessed 15.4.21]. 18p.

[9] National Institute for Health and Care Excellence. Cancer Drugs Fund [Internet]. London: NICE, n.d. [accessed 15.4.21]. Available from: https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-technology-appraisal-guidance/cancer-drugs-fund

[10] Treon SP, Xu L, Hunter Z. MYD88 Mutations and response to ibrutinib in Waldenstrom's macroglobulinemia. *N Engl J Med* 2015;373(6):584-6.

[11] Grunenberg A, Buske C. Treatment options for relapsed Waldenström's macroglobulinemia. *HemaSphere* 2019;3(S2):65-67.

[12] BeiGene. *Single technology appraisal. Zanubrutinib for treating Waldenström's macroglobulinaemia [ID1427]. Document B: appendices*: BeiGene, 12th March 2021 [accessed 18.3.21]. 70p.

[13] BeiGene. *Single technology appraisal. Zanubrutinib for treating Waldenstrom's macroglobulinaemia [ID1427]: clarification response (version 3.0) [Word document]*: BeiGene, 14th May 2021 [accessed 19.5.21]. 154p.

[14] National Institute for Health and Care Excellence. 5.4 Developing search strategies: limits and filters. In: Developing NICE guidelines: the manual. Process and methods [PMG20] [Internet]. NICE, Updated: 15 October 2020 [accessed 27.5.21]. Available from: https://www.nice.org.uk/process/pmg20/chapter/identifying-the-evidence-literature-searching-and-evidence-submission#developing-search-strategies

[15] National Institute for Health and Care Excellence. Single technology appraisal: User guide for company evidence submission template. Process and methods [PMG24] [Internet]. NICE, Updated 1 April 2017 [accessed 27.5.21]. Available from: https://www.nice.org.uk/process/pmg9/chapter/the-reference-case#synthesis-of-evidence-on-health-effects

[16] National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal 2013 (2018 revision) [Internet]*. London: NICE, 2013 (revised 2018) [accessed 15.9.20]. 93p. Available from: https://www.nice.org.uk/process/pmg9/

[17] Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care [Internet]*. York: University of York, 2009 [accessed 23.3.11] Available from: http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm

[18] Morrison A, Polisena J, Husereau D, Moulton K, Clark M, Fiander M, et al. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. *Int J Technol Assess Health Care* 2012;28(2):138-44.

[19] Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf M, et al. Chapter 4: Searching for and selecting studies. 4.4.5 Language, date and document format restrictions. In: Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al., editors. *Cochrane handbook for systematic reviews of interventions version 6.0 (updated July 2019) [Internet]*: Cochrane, 2019 [accessed 21.7.20]. Available from: https://training.cochrane.org/handbook/current/chapter-04

[20] Egger M, Zellweger-Zähner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. *Lancet* 1997;350(9074):326-9.

[21] Higgins JPT, Green S, eds. *Cochrane handbook for systematic reviews of interventions [Internet]*. Version 5.1.0 [updated March 2011]: The Cochrane Collaboration, 2011 [accessed 23.3.11]. Available from: http://handbook.cochrane.org/

[22] National Institute for Health Care Excellence. *Methods for the development of NICE public health guidance (third edition) [Internet]*. London: NICE, 2012 [accessed February 2021] Available from: https://www.nice.org.uk/process/pmg4/chapter/appendix-f-quality-appraisal-checklist-quantitative-intervention-studies

[23] Tam CS, Opat S, D'Sa S, Jurczak W, Lee HP, Cull G, et al. A Randomized Phase 3 Trial of Zanubrutinib Versus Ibrutinib in Symptomatic Waldenstrom Macroglobulinemia:The Aspen Study. *Blood* 2020.

[24] BeiGene. *Clinical study report: ASPEN BGB-3111-302*, June 2020

[25] Tam CS, LeBlond V, Novotny W, Owen RG, Tedeschi A, Atwal S, et al. A head-to-head Phase III study comparing zanubrutinib versus ibrutinib in patients with Waldenstrom macroglobulinemia. *Future Oncol* 2018;14(22):2229-2237.

[26] BeiGene. *Addendum to clinical study report. A phase 3, randomized, open-label, multicenter study comparing the efficacy and safety of the bruton's tyrosine kinase (BTK) inhibitors BGB-3111 and ibrutinib in subjects with Waldenström's Macroglobulinemia (WM). BGB-3111-302 [PDF supplied by the company]*: BeiGene, 5th February 2021 [accessed 4.5.21]. 462p.

[27] Garcia-Sanz R, Dimopoulos MA, Lee H-P, Trneny M, Varettoni M, Owen RG, et al. Updated results of the ASPEN trial from a cohort of patients with MYD88 wild-type (MYD88WT) Waldenström macroglobulinemia (WM). *J Clin Oncol* 2020;38(15 suppl):e20056-e20056. Available from: https://ascopubs.org/doi/abs/10.1200/JCO.2020.38.15_suppl.e20056

[28] Dimopoulos M. Updated results of the ASPEN trial from a cohort of patients with MYD88 wild-type (MYD88WT) Waldenström macroglobulinemia (WM) [Virtual]. *American Society of Clinical Oncology* 2020.

[29] Trotman J, Opat S, Gottlieb DJ, Simpson D, Marlton P, Cull G, et al. Zanubrutinib for the treatment of patients with Waldenström macroglobulinemia: three years of follow-up. *Blood* 2020.

[30] BeiGene. *Clinical study report: BGB-3111-AU-003*, 2020

[31] Paludo J, Abeykoon JP, Shreders A, Ansell SM, Kumar S, Ailawadhi S, et al. Bendamustine and rituximab (BR) versus dexamethasone, rituximab, and cyclophosphamide (DRC) in patients with Waldenstrom macroglobulinemia. *Ann Hematol* 2018;97(8):1417-1425.

[32] Tedeschi A, Picardi P, Ferrero S, Benevolo G, Margiotta Casaluci G, Varettoni M, et al. Bendamustine and rituximab combination is safe and effective as salvage regimen in Waldenstrom macroglobulinemia. *Leuk Lymphoma* 2015;56(9):2637-42.

[33] Dimopoulos MA, Anagnostopoulos A, Kyrtsonis MC, Zervas K, Tsatalas C, Kokkinis G, et al. Primary treatment of Waldenström macroglobulinemia with dexamethasone, rituximab, and cyclophosphamide. *J Clin Oncol* 2007;25(22):3344-9.

[34] Kastritis E, Gavriatopoulou M, Kyrtsonis MC, Roussou M, Hadjiharissi E, Symeonidis A, et al. Dexamethasone, rituximab, and cyclophosphamide as primary treatment of Waldenström macroglobulinemia: final analysis of a phase 2 study. *Blood* 2015;126(11):1392-4.

[35] Signorovitch JE, Sikirica V, Erder MH, Xie J, Lu M, Hodgkins PS, et al. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. *Value Health* 2012;15(6):940-7.

[36] Tam CS, Wolf MM, Westerman D, Januszewicz EH, Prince HM, Seymour JF. Fludarabine combination therapy is highly effective in first-line and salvage treatment of patients with Waldenstrom's macroglobulinemia. *Clin Lymphoma Myeloma* 2005;6(2):136-9.

[37] Ngan S, Rohatiner AZ, Matthews J, Williams A, Amess JA, Norton A, et al. Waldenstrom's macroglobulinemia: a retrospective analysis of 40 patients from 1972 to 2001. *Semin Oncol* 2003;30(2):236-8.

[38] Tedeschi A, Ricci F, Goldaniga MC, Benevolo G, Varettoni M, Motta M, et al. Fludarabine, cyclophosphamide, and rituximab in salvage therapy of Waldenstrom's macroglobulinemia. *Clin Lymphoma Myeloma Leuk* 2013;13(2):231-4.

[39] Dimopoulos MA, Zervas C, Zomas A, Hamilos G, Gika D, Efstathiou E, et al. Extended rituximab therapy for previously untreated patients with Waldenstrom's macroglobulinemia. *Clin Lymphoma* 2002;3(3):163-6.

[40] Dimopoulos MA, Zervas C, Zomas A, Kiamouris C, Viniou NA, Grigoraki V, et al. Treatment of Waldenstrom's macroglobulinemia with rituximab. *J Clin Oncol* 2002;20(9):2327-33.

[41] Byrd JC, White CA, Link B, Lucas MS, Velasquez WS, Rosenberg J, et al. Rituximab therapy in Waldenstrom's macroglobulinemia: preliminary evidence of clinical activity. *Ann Oncol* 1999;10(12):1525-7.

[42] Treon SP, Agus TB, Link B, Rodrigues G, Molina A, Lacy MQ, et al. CD20-directed antibody-mediated immunotherapy induces responses and facilitates hematologic recovery in patients with Waldenstrom's macroglobulinemia. *J Immunother* 2001;24(3):272-9.

[43] Castillo JJ, Gustine JN, Meid K, Dubeau TE, Severns P, Xu L, et al. Response and survival for primary therapy combination regimens and maintenance rituximab in Waldenstrom macroglobulinaemia. *Br J Haematol* 2018;181(1):77-85.

[44] National Institute for Health and Care Excellence. *Position statement: consideration of products recommended for use in the Cancer Drugs Fund as comparators, or in a treatment sequence, in the appraisal of a new cancer product[Internet]*. London: NICE, January 2019 [accessed 19.5.21]. 1p. Available from: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisal-guidance/cancer-drugs-fund/CDF-comparator-position-statement.pdf

[45] Latimer NR. *Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data (NICE Decision Support Unit Technical Support Document 14) [Internet]*. London: National Institute for Health and Care Excellence (NICE) updated March 2013 [accessed 27.5.21]. 52p. Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf

[46] National Institute for Health Care Excellence. *Ibrutinib for treating Waldenstrom's macroglobulinaemia [TA491]*, 2017 [accessed October 2020] Available from: www.nice.org.uk/guidance/ta491

[47] Kastritis E, Leblond V, Dimopoulos MA, Kimby E, Staber P, Kersten MJ, et al. Waldenstrom's macroglobulinaemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2019;30(5):860-862.

[48] Soikkeli F, Hashim M, Ouwens M, Postma M, Heeg B. Extrapolating survival data using historical trial-based a priori distributions. *Value Health* 2019;22(9):1012-1017.

[49] van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health* 2012;15(5):708-15.

[50] Dolan P, Kahneman D. Interpretations of utility and their implications for the valuation of health*. *The Economic Journal* 2008;118(525):215-234.

[51] National Institute for Health Care Excellence. *Ibrutinib for treating chronic lymphocytic leukaemia [ID749]. Committee Papers*, 2016

[52] Public Health England. Health-related quality of life for people aged 65 and over [Internet]. Updated 5 March 2021 [accessed 18.5.21]. Available from: https://www.ethnicity-facts-figures.service.gov.uk/health/physical-health/health-related-quality-of-life-for-people-aged-65-and-over/latest

[53] Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value Health* 2010;13(5):509-18.

[54] Excellence NIfHC. *Ibrutinib for treating relapsed or refractory mantle cell lymphoma [ID753]. Committee Papers*, 2016

[55] Department of Health. *NHS reference costs 2018-2019 [Internet]*. London: Department of Health, 2020 [accessed 16.10.20] Available from: https://www.england.nhs.uk/national-cost-collection/#ncc1819

[56] Joint Formulary Committee. *British National Formulary [Internet]*. London: BMJ Group and Pharmaceutical Press, 2020 [accessed 6.8.20]. Available from: https://bnf.nice.org.uk/

[57] Personal Social Services Research Unit. *Unit Costs of Health and Social Care*. Canterbury: University of Kent, ???? [accessed ??.??.??]. 237p. Available from: ????

[58] Büyükkaramikli NC, Rutten-van Mölken M, Severens JL, Al M. TECH-VER: a verification checklist to reduce errors in models and improve their credibility. *Pharmacoeconomics* 2019;37(11):1391-1408.

[59] Grimm SE, Pouwels X, Ramaekers BLT, Wijnen B, Knies S, Grutters J, et al. Development and Validation of the TRansparent Uncertainty ASsessmenT (TRUST) Tool for assessing uncertainties in health economic decision models. *Pharmacoeconomics* 2020;38(2):205-216.

[60] Kaltenthaler E, Carroll C, Hill-McManus D, Scope A, Holmes M, Rice S, et al. The use of exploratory analyses within the National Institute for Health and Care Excellence single technology appraisal process: an evaluation and qualitative analysis. *Health Technol Assess* 2016;20(26):1-48.

## Appendix 1: Additional ERG comments regarding the clinical and cost effectiveness searches

### Additional ERG critique regarding the clinical effectiveness search

Each facet of the Embase clinical effectiveness search strategy contained only brief title and abstract synonyms for each of the drug in the intervention facet. The ERG identified additional word and spelling variants which, if included, would have improved strategy recall. CAS Registry numbers would also have aided strategy performance. All ERG additions and amendments are underlined in blue.

**ERG test searches run on 1.4.21 in Embase (Ovid): 1974 to 2021/03/31**

For the Zanubrutinib facet:

| | | |
|---|---|---|
| 1 | zanubrutinib/ | 224 |
| 2 | (zanubrutinib or brukinsa or bgb3111 or bgb3111).ti,ab. | 117 |
| 3 | 1691249-45-2.rn. | 200 |
| 4 | or/1-3 | 228 |

For the Rituximab facet:

| | | |
|---|---|---|
| 1 | rituximab/ | 85318 |
| 2 | (ritixumab or blitzima or mabthera or reditux or ritemvia or ritumax or rituxan or rituxin or rituzena or rixathon or riximyo or ruxience or truxima or tuxella or riabni or zytux or novex or rituzena).ti,ab. | 1011 |
| 3 | (abp-798 or abp-798 or ct-p10 or ctp10 or gp-2013 or gp2013 or lx01 or hlx-01 or idec-102 or idec-c2b8 or idec102 or idecc2b8 or nk-8808 or mk8808 or pf-05280586 or pf05280586 or pf5280586 or r-105 or r105 or rg-105 or rg105 or ro-452294 or ro452294).ti,ab. | 387 |
| 4 | 174722-31-7.rn. | 70400 |
| 5 | or/1-4 | 85648 |

For the Bendamustine facet:

**ERG comment:** The Embase search strategy contains a typographical error (line 75, pg 17). The omission of a space between the OR search operator and the final free text term means this search missed +550 references (see below):

| | | |
|---|---|---|
| 2 | (Bendamustine or belrapzo or bendamustin or bendeka or cytostasan* or levact or ribomustin or ribovact or treanda or?imet*).ti,ab. | 4280 |
| 3 | (Bendamustine or belrapzo or bendamustin or bendeka or cytostasan* or levact or ribomustin or ribovact or treanda or ?imet*).ti,ab. | 4855 |

The ERG conducted further tests on the performance of the ? wildcard character at the start of a word. According to Ovid, the optional wild card '?' character stands for zero or one characters within a word

or at the end of a word. This wildcard should not be used at the start of a word. The ERG identified three word variant synonyms for Bendamustine that should have been searched as full words for the strategy to work correctly.

| | | |
|---|---|---|
| 8 | ?imet.ti,ab. | 156 |
| 9 | (cimet or imet or zimet).ti,ab. | 199 |

The impact on recall was sufficient that the ERG concluded the search phrase '?imet*' may not have performed as the searcher intended.

The ERG identified the appropriate CAS Registry numbers for this drug, which would have increased search sensitivity further:

| | | | |
|---|---|---|---|
| 1 | Bendamustine/ | 7175 | Advanced |
| 2 | (Bendamustine or belrapzo or bendamustin or bendeka or cytostasan* or levact or ribomustin or ribovact or treanda or?imet*).ti,ab. | 4280 | Advanced |
| 3 | or/1-2 | 7562 | Advanced |
| 4 | Bendamustine/ | 7175 | Advanced |
| 5 | (Bendamustine or belrapzo or bendamustin or bendeka or cytostasan* or levact or ribomustin or ribovact or treanda or cimet or imet or zimet).ti,ab. | 4481 | Advanced |
| 6 | (16506-27-7 or 3543-75-7).rn. | 5818 | Advanced |
| 7 | or/4-6 | 7753 | Advanced |

Company syntax

ERG search with correction and CAS RN

For the Fludarabine facet:

| | | |
|---|---|---|
| 1 | fludarabine/ | 28697 |
| 2 | (fludarabine or Fludara or Oforta).ti,ab. | 13720 |
| 3 | 21679-14-1.rn. | 23821 |
| 4 | 1 or 3 | 28700 |

For the dexamethasone facet:

| | | |
|---|---|---|
| 1 | dexamethasone/ | 156300 |
| 2 | (Dextenza or Ozurdex or Neofordex or Decadron or Dexamethasone or Dexasone or Decadron or adrecort or adrenocot or aeroseb-dex or aflucoson or aflucosone or alfalyl or anaflogistico or arcodexan or arcodexane).ti,ab. | 80969 |
| 3 | (artrosone or azium or bidexol or calonat or cebedex or cetadexon or colofoam or corsona or cortastat or cortidex or cortidexason or cortidrona or cortidrone or cortisumman or dacortina fuerte or dacortine fuerte or dalalone or danasone or de-sone la or decacortin or decadeltosona or decadeltosone or decaderm or decadion or decadran or decadronal or decaesadril or decaject or decamethasone or decasone or decaspray or decasterolone or decdan or decilone or decilone forte or decofluor or dectancyl or dekacort).ti,ab. | 10 |
| 4 | (delladec or deltafluoren or deltafluorene or dergramin or deronil or desacort or desacortone or desadrene or desalark or desameton or desametone or desigdron or dexa cortisyl or dexa dabrosan or dexa korti or dexa scherosan or dexa scherozon or dexa scherozone or dexa-p or dexacen or dexachel or dexacort or dexacortal or dexacorten or dexacortin or dexacortisyl or dexadabroson or dexadecadrol or dexadrol or dexagel or dexagen or dexahelvacort or dexakorti or dexalien or dexalocal or dexame or dexamecortin or dexameson or dexamesone).ti,ab. | 66 |
| 5 | (fluoromethylprednisolone or fortecortin or gammacorten or gammacortene or grosodexon or grosodexone or hemady or hexadecadiol or hexadecadrol or hexadiol or hexadrol or isnacort or isopto dex or isopto maxidex or isopto-dex or isoptodex or isoptomaxidex or lokalison or loverine or luxazone or marvidione or maxidex or mediamethasone or megacortin or mephameson or mephamesone or metasolon or metasolone or methazon ion or methazone ion or methazonion or methazonione or metisone lafi or mexasone or millicorten or millicortenol or mk-125 or mk125 or mymethasone or nisomethasona or novocort or nsc-34521 or nsc34521 or oftan-dexa or opticorten or opticortinol or oradexan or oradexon or oradexone or orgadrone or pidexon or policort or posurdex or predni-f or prednisolone or prodexona or prodexone or sanamethasone or santenson or santeson or sawasone or solurex or solurex la or spoloven or sterasone or thilodexine or triamcimetil or vexamet or visumetazone or visumethazone).ti,ab. | 40333 |
| 6 | (dexametason or dexametasone or dexameth or dexamethason or dexamethazon or dexamethazone or dexamethonium or dexamonozon or dexan or dexane or dexano or dexapot or dexascheroson or dexascherozon or dexascherozone or dexason or dexinoral or dexionil or dexmethsone or dexona or dexone or dexpak taperpak or dextelan or dextrasone or dexycu or dezone or dibasona or doxamethasone or esacortene or ex-s1 or exadion or exadione or firmalone or fluormethyl prednisolone or fluormethylprednisolon or fluormethylprednisolone or fluormone or fluorocort or fluorodelta).ti,ab. | 1267 |
| 7 | 50-02-2.rn. | 144167 |
| 8 | or/1-7 | 205467 |
| 9 | 1 or dexmethasone.ti,ab. | 156311 |
| 10 | 8 not 9 | 49165 |

Company search

Additional results identified buy the ERG search terms

For the Cyclophosphamide facet:

| | | |
|---|---|---|
| 1 | cyclophosphamide/ | 221941 |
| 2 | cyclophosphamide.ti,ab. | 77573 |
| 3 | (alkyroxan or b-518 or b518 or carloxan or ciclofosfamida or ciclolen or cicloxal or clafen or cyclo-cell or cycloblastin or cycloblastine or cyclofosamide or cyclofosfamid or cyclofosfamide or cyclophar or cyclophosphamid or cyclophosphamide isopac or cyclophosphamides or cyclophosphan or cyclophosphane or cyclostin or cyclostin n or cycloxan or cyphos or cytophosphan or cytophosphane or cytoxan or cytoxan lyophilized or endocyclo phosphate or endoxan or endoxan asta or endoxan-asta or endoxana or endoxon-asta or enduxan or genoxal or ledoxan or ledoxina or lyophilized cytoxan or mitoxan or neosan or neosar or noristan or nsc-26271 or nsc2671 or procytox or procytoxide or Revimmune or semdoxan or sendoxan or syklofosfamid).ti,ab. | 2633 |
| 4 | 50-18-0.rn. | 207651 |
| 5 | or/1-4 | 229990 |

For the cladribine facet:

| | | |
|---|---|---|
| 1 | cladribine/ | 7348 |
| 2 | (cladribine or leustatin).ti,ab. | 2630 |
| 3 | (biodribin or novaplus or intocel or leustat or leustatine or litak or litax or mavenclad or movectro or mylinax or rwj-26251 or rwj26251).ti,ab. | 52 |
| 4 | 4291-63-8.rn. | 6814 |
| 5 | or/1-4 | 7500 |

Due to the way the search strategy was constructed, the ERG felt it was likely further synonyms, word variants and CAS registry numbers would also have been beneficial in the chlorambucil, alemtuzumab, bortezomib, prednisone and vincristine facets:

**Additional ERG critique regarding the cost effectiveness searches**

The ERG noticed considerable duplication within the combined cost effectiveness, health-related quality of life and cost and healthcare resource identification, measurement and valuation study design filters reported in Table B.5.29 of the appendices to the company submission.

The following lines were affected:

Lines S6, S10-14 are repeated later in the strategy as lines S39-43.

| **S6** | **MESH.EXACT("Economics")** |
|---|---|
| S10 | EMB.EXACT("Economic aspect") |
| S11 | EMB.EXACT("Socioeconomics") |
| S12 | MESH.EXACT("Economics, pharmaceutical") |
| S13 | EMB.EXACT("Health economics") |
| S14 | MESH.EXACT("Costs and cost analysis") |

| **S38** | **MESH.EXACT("Economics")** |
|---|---|
| S39 | EMB.EXACT("Economic aspect") |
| S40 | EMB.EXACT("Socioeconomics") |
| S41 | MESH.EXACT("Economics, pharmaceutical") |
| S42 | EMB.EXACT("Health economics") |
| S43 | MESH.EXACT("Costs and cost analysis") |

Repetition was also noted in lines S78-79.

| **S78** | **MESH.EXACT.EXPLODE("Costs and cost analysis")** |
|---|---|
| S79 | EMB.EXACT("Economics") |

Line S29 is repeated later in the strategy as line S77.

| **S29** | **TI,IF(economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed)** |
|---|---|

| **S77** | **TI,IF(Economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed)** |
|---|---|

The ERG felt that the strategy would have benefitted from streamlining of these lines to reduce repetition. However, the ERG acknowledged that strategy performance was not adversely affected.

# National Institute for Health and Care Excellence

# Centre for Health Technology Evaluation

## ERG report – factual accuracy check and confidential information check

## Zanubrutinib for treating Waldenstrom's macroglobulinaemia [ID1427]

'*Data owners will be asked to check that confidential information is correctly marked in documents created by others in the technology appraisal process before release; for example, the technical report and ERG report.*' (Section 3.1.29, Guide to the processes of technology appraisals).

You are asked to check the ERG 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 **5pm on Friday 18 June 2021** 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 separately highlight information that is submitted as 'commercial in confidence' in turquoise, all information submitted as 'academic in confidence' in yellow, and all information submitted as 'depersonalised data' in pink.

## Issue 1 Inaccurate descriptions of outcomes included for MAICs

| Description of problem | Description of proposed amendment | Justification for amendment | ERG Response |
|---|---|---|---|
| Page 15, Table 1.5 and Table 1.6: "only PFS and OS were considered as outcomes in the MAIC".<br><br>Page 57: “only PFS and OS were considered as outcomes in the MAIC" | Change “PFS, OS, and AEs were considered as outcomes in the MAIC”. | The original descriptions were inaccurate because in addition to PFS and OS, AE was also considered, reported, and applied in the cost-effectiveness analysis. (CS, Section B.3.3.3, Table B.3.19).<br><br>The MAIC results of AE were also reported in the ERG report (Section 4.2.7, Table 4.5). | Agree, change made. |

## Issue 2 Inaccurate descriptions of populations in the studies considered for MAICs

| Description of problem | Description of proposed amendment | Justification for amendment | ERG Response |
|---|---|---|---|
| Page 57: “The populations in the two trials of chlorambucil monotherapy and rituximab monotherapy in the company submission MAICs were solely adults with WM whose disease is untreated and for whom chemo-immunotherapy is unsuitable. The company considered it unfeasible to match the treatment naïve (unsuitable for chemo-immunotherapy) patients in ASPEN to these studies given the small number of such patients in the zanubrutinib arm (n=19), however, MAICs | Either delete this whole paragraph or change to “~~The populations in the two trials of chlorambucil monotherapy and rituximab monotherapy in the company submission MAICs were solely adults with WM whose disease is untreated and for whom chemo-immunotherapy is unsuitable~~ The population in the study of chlorambucil monotherapy identified from the | The original statements were inaccurate in several ways:<br><br>The first sentence from the ERG report implied that MAICs were conducted in the CS. However, the MAICs were conducted during the clarification stage. More accurate descriptions would be either “considered in the company submission MAICs” or “identified from the company submission SLR”. Even if later in the | The first part has been amended using the text suggested.<br><br>And we have amended the last sentence to:<br><br>“The populations in the trials included in the clarification response exploratory MAICs were varied, including both treatment naïve and treatment refractory patients, these results should be interpreted with caution.” |

| | | | |
|---|---|---|---|
| were performed for one of each treatment in the clarification response. The populations in the trails included in the clarification response MAICs were varied, including both treatment naïve and treatment refractory patients." | company submission SLR was adults with WM with unknown prior treatment history. The population in the study of rituximab monotherapy identified from the company submission SLR included both treatment-naïve (suitable for chemo-immunotherapy) and relapsed/refractory patients. The company considered it unfeasible to match the treatment naïve (unsuitable for chemo-immunotherapy) patients in ASPEN to these studies given the small number of such patients in the zanubrutinib arm (n=19), however, MAICs were performed for one of each treatment in the clarification response. ~~The populations in the trails trials included in the clarification response MAICs were varied, including both treatment naïve and treatment refractory patients.~~" | paragraph, ERG mentioned that "MAICs were performed for one of teach treatment in the clarification response", the first sentence could still be confusing and misleading.<br><br>The first sentence from the ERG reported indicated the comparator study populations were consistently "adults with WM *whose disease is untreated and for whom chemo-immunotherapy is unsuitable*". However, this was inaccurate. The study of chlorambucil monotherapy (Kyle 2000) included 46 patients with *unknown prior treatment history*, likely from a single centre in the US, whereas the study for rituximab monotherapy (Gertz 2004/2009) included *34 treatment-naïve patients (suitable for chemo-immunotherapy) and 35 relapsed/refractory patients* from multiple centres in the US. It is important to be clear in the ERG report as to the definition of study population because it is directly related to the conclusion of the feasibility assessment of indirect treatment comparisons. Specifically, due to the discrepancies in the study populations between the comparator studies (i.e., Kyle 2000 and Gertz 2004/2009) and the population per NICE scope | |

| | | | |
|---|---|---|---|
| | | (*treatment-naïve patients unsuitable for chemo-immunotherapy*), the Company did not include these studies in the company submission MAICs. Despite that the company included these studies in MAICs (adjusting the overall zanubrutinib arm to match the overall comparator study populations) during the clarification stage for exploratory purposes and for transparency, the results should be interpreted with extreme cautions given the discrepancies in the study populations. | |

## Issue 3 Inaccurate descriptions of the studies considered for MAICs

| Description of problem | Description of proposed amendment | Justification for amendment | ERG Response |
|---|---|---|---|
| Page 60: “Table 3.19 shows an overview of MAIC trials” “Table 3.19: Overview of MAIC trials in the clarification response”<br><br>Page 62: “The MAICs were conducted using single-arm trials to compare zanubrutinib with comparator”<br><br>Page 63: “MAICs allow for patients in different trials to be made more comparable” | Change “trials” to “studies” | The studies identified from the clinical SLR (Company Submission) and presented in Table 3.19 included a mix of prospective interventional studies and observational retrospective studies. The term “trial” could be misleading to audience that all the studies fall under the former category. Considering that the study design was taken into account during the MAIC feasibility assessment, it is suggested to use the relatively broader “studies” to be more accurate. | Agree, changes made. |

## Issue 4 Inaccurate numbers

| Description of problem | Description of proposed amendment | Justification for amendment | ERG Response |
|---|---|---|---|
| Page 25, Table 1.19 \| Page 104, Table 6.3 \| (in general, throughout the ERG report): None of the probabilistic base case result estimates in the ERG report can be replicated in the ERG model that the company received (with ERG base case set as default). | Update the PSA results (throughout the ERG report) to be in line with those in the ERG model. | It is acknowledged that the exact PSA results are different in each run. However, considering that there are a lot of ERG comments in the ERG report specifically about the differences between the deterministic and probabilistic base | Agree, changes made. |

| | | | |
|---|---|---|---|
| On the other hand, the PSA results shown in the ERG model (which, based on the way the model was set up, model users are able to tell that the PSA results were generated based on the deterministic base case analysis in the model) could not be found anywhere in the ERG report. | | case analysis results, it is suggested to ensure that the PSA results (reported in the ERG report) can be replicated, for transparency. At the time of this factual check, the company was not able to do factual check of the PSA results presented in the ERG report based on the ERG report and the ERG model that were shared with the company. | |
| In the ERG model, the duration of sepsis (6.7 days) was inconsistent with that reported in Page 84, Table 4.5 (5.0 days). | Either make a note in the ERG report or update the ERG model (which is currently populated with 6.7 days) to be aligned with what's reported in the ERG report (5.0 days); no proposed amendment for the ERG report. | The duration of sepsis (5.0 days per ASPEN IPD) was reported in the Response to Clarification Letter and subsequently in the ERG report. However, in the ERG model, a duration of 6.7 days was applied, which was inconsistent with that in the ERG report.<br><br>That said, it is acknowledged that it may lead to a different inconsistency issue (i.e., even if "6.7" is replaced with "5.0" in the ERG report, all the results throughout the ERG report were still based on "6.7"), especially when this minor correction would have very minimal impact on the results. | Not a factual inaccuracy. This parameter was not updated or added by the ERG. It was already part of the initial company submission. Any discrepancy between the CS and the company's model is not caused by any changes made by the ERG. |

## Issue 5 Inaccurate descriptions of the scope of SLR

| Description of problem | Description of proposed amendment | Justification for amendment | ERG Response |
|---|---|---|---|
| Page 56, Section 3.3: " the SLR aimed to identify the efficacy of treatment options for previously untreated patients published evidence for current and future treatment options for patients with WM, as well as the safety and tolerability of treatment options for patients with WM." | Change to “the SLR aimed to identify the efficacy of treatment options for ~~previously untreated~~ patients with WM and published evidence for current and future treatment options..." | The original statement was inaccurate because the SLR was not restricted to *previously untreated* patients but WM patients overall, including both *previously untreated* and *relapsed or refractory* patients. | Agree, change made.<br>However, this was taken from appendix D of the CS (page 15). Therefore, the company might also want to correct the statement there. |

## Issue 6 Inaccurate descriptions of the analyses requested by the ERG

| Description of problem | Description of proposed amendment | Justification for amendment | ERG Response |
|---|---|---|---|
| Page 82: “The ERG considered that, given the extreme data immaturity and given that median follow-up was longer in BGB-3111-AU-003 than in ASPEN (48 months vs 19 months), BGB-3111-AU-003 in particular may be helpful in adding information on PFS and OS hazard patterns in the long run and requested that the company perform survival analysis using external data, for example using the method by Soikkeli et al..48 The company did not provide this, stating that data immaturity would also be an issue with BGB-3111-AU-003.” | The last sentence may be re-written, given that the company did conduct the analyses using BGB-3111-AU-003 data, using the method by Soikkeli et al. | Despite that the company asked for an extension of the timeline for submitting the response for clarification question B5b due to the time required for internal request of patient-level data of BGB-3111-AU-003, the company did conduct the analyses requested by the ERG and submit the associated results to NICE on May 14, 2021. Despite that data immaturity is also an issue with BGB-3111-AU-003, the survival analysis results (using BGB-3111-AU-003 data with relatively longer follow-up) supported the conclusion of the model selection (based on the ASPEN data) applied in the CS base case. | This sentence has been deleted. |

## Issue 7 Typographic errors

| Description of problem | Description of proposed amendment | Justification for amendment | ERG Response |
|---|---|---|---|
| Page 57: “The populations in the trails included in the clarification response MAICs were varied, including both | Change it to “The populations in the ~~trails~~ trials included in the clarification response MAICs were varied, | A typographic error | Corrected. |

| | | | |
|---|---|---|---|
| treatment naïve and treatment refractory patients." | including both treatment naïve and treatment refractory patients." | | |
| Pages 81-82: "…which showed an observed median OS of 14.6 months for BTK inhibitors, which the company considered similar to the median OS estimates for zanubrutinib based on their base-case assumptions (gamma model for DRC resulting in a median of 14.6 months and exponential for BR comparison resulting in a median of 15.29 months, both after adjusting for background mortality)." The time unit should not be months, but years. In addition, 14.6 and 15.29 years are median survival for zanubrutinib (match DRC) and zanubrutinib (match BR), rather than DRC and BR." | Change to "…which showed an observed median OS of 14.6 ~~months~~ years for BTK inhibitors, which the company considered similar to the median OS estimates for zanubrutinib based on their base-case assumptions (gamma model for DRC resulting in a median of 14.6 ~~months~~ years and exponential for BR comparison resulting in a median of 15.29 ~~months~~ years, both after adjusting for background mortality)." The time unit should not be months, but years. In addition, 14.6 and 15.29 years are median survival for zanubrutinib (match DRC) and zanubrutinib (match BR), rather than DRC and BR. | Typographic errors | Corrected. |
| Page 80: "e.g. only one out of 102 patients still at risk of OS event after 30 months, clarification response to B7, Figure 34". | Change "clarification response to B7, Figure 34" to "clarification response to B7, Figure 57" | Figure 34 in the clarification letter was not about OS or under response to B7. Figure 57 was. | Corrected. |
| Page 80: "…with jointly-fitted and independent models respectively for the BR comparison (Tables 43 and 45 in the clarification letter response…" | Change "Table 45" to "Table 37" | Table 45 in the clarification letter response was about utilities rather than survival extrapolation. | Corrected. |
| Page 82: "Upon request, the company assessed the impact of summing up model hazards and background mortality hazards in scenario analysis: this increased the ICER substantially in both | Change "Table 47" to "Table 58". | Table 47 in response to clarification letter was about utilities rather than the mortality hazards or the associated ICER results. Table 58 in response to | Corrected. |

| | | | |
|---|---|---|---|
| BR and DRC (Table 47 of clarification letter response)" | | clarification letter reported the associated ICER results. | |
| Page 85: "Respiratory failure occurred in ≥1% of the population of a treatment arm according to Table B.2.23." | Change "Table B.2.23" to "CS Table B.2.26".<br>Delete this sentence. | Respiratory failure was not reported in Table B.2.23 (CS) but in Table B.2.26 (CS).<br>Even after the correction of cross-reference, this statement was inaccurate. Specifically, in Table B.2.26 (CS), it was reported that respiratory failure occurred in 1/101 (0.99%) of the patients in the zanubrutinib arm in cohort 1 (i.e., the modelled population for zanubrutinib). In addition, Therefore, the original statement was not correct for the zanubrutinib arm. It was not correct for any comparator arm, based on the publicly available information from the comparator study population. | Agree, this sentence has been deleted. |
| Page 85, Table 4.6: "Source: CS Table B.3.22" | Change "Table B.3.22" to "Table B.3.24". | Neither Table B.3.24 nor Table B.3.22 was incorrect, but Table B.3.24 is technically more relevant and specific. | Not a factual inaccuracy. |
| Page 105: "The estimated ERG base-case ICER (probabilistic), based on the ERG preferred assumptions highlighted in Section 6.1, was … for ibrutinib compared to BR and … for ibrutinib compared to DRC." | Change to "The estimated ERG base-case ICER (probabilistic), based on the ERG preferred assumptions highlighted in Section 6.1, was … for zanubrutinib compared to BR and … for zanubrutinib compared to DRC." | No comparison was conducted between ibrutinib and BR or between ibrutinib and DRC. | Corrected. |

| | | | |
|---|---|---|---|
| In general (throughout the ERG report), comma and dot were used as 1000 separator in an exchangeable manner (e.g., “1,000” versus “1.000”). | Use “comma” or “dot” consistently. | Although it is a relatively minor issue (or lies with region/country-specific norms), it is suggested to be consistent throughout the ERG report in case of misinterpretation in audiences who may have little knowledge of the context. | We have checked the ERG report and made changes accordingly.<br>The whole report should now only use UK style punctuation. |

## Issue 8 Incorrect or Missing ACIC Marking

| Location of incorrect marking | Description of incorrect marking | Amended Marking | ERG Response |
|---|---|---|---|
| Page 64, Figure 3.4; Page 65 Figure 3.5 | The results presented in Figure 3.4 were partially marked as academic in confidence in CS. As such, the entire Figure 3.4 in the ERG reported may be marked as academic in confidence. | Not applicable because this is about figures rather than texts | We have marked Figure 3.4 as AIC. However, we can’t see the marking in the document. Therefore, we have marked the word ‘Figure’ as well. |
| Page 81, Figure 4.2 and Figure 4.3 | In general, the extrapolated survival parametric curves were marked as academic in confidence in CS. As such, the entire Figure 4.2 and Figure 4.3 in the ERG | Not applicable because this is about figures rather than texts | Corrected. |

| | | | |
|---|---|---|---|
| | reported may be marked as academic in confidence. | | |
| Page 84, Table 4.5, columns "Zanubrutinib adjusted to match BR ([redacted])" and "Zanubrutinib adjusted to match DRC ([redacted])" | Such information was marked as academic in confidence in CS. | Both the exact effective sample sizes (from the column headers) and the numbers in both columns need to be redacted. | Corrected. Although it should be noted that these numbers are not marked as AIC in the clarification to response. |
| Section 3.4 P64 | This statement was not marked in the ERG report but needs to be marked as AIC. | "Still, the results for the MAICs in the company submission (BR: Tedeschi (2015);[32] DRC: Dimopoukos (2007)[33]/Kastritis (2015)[34]) [redacted]" | This is our interpretation of the evidence, as such we do not think this should be AIC. No change made. |
| Section 3.2.7 P54-55 | This section has not been marked up in the ERG report however the pooled safety data has yet to be presented in the public domain. BeiGene consider it to be AIC | **3.2.7.2 Pooled adverse events**<br>The CS reported pooled safety data for all Waldenström's macrobulinaemia (WM) patients (n=253) across four studies (the ASPEN trial, BGB-3111-210, BGB-3111-AU-003, and BGB-3111-1002) (no comparison group, CS pages 68-77,[4] and Table 3.16 below). The median duration of exposure in all WM patients was [redacted]. A summary of AEs is included in Table B.2.37 in the CS[4] (Table 3.16 below). [redacted] reported one AE, [redacted] reported at least one serious AE, [redacted] had an AE leading to discontinuation, and [redacted]. The only event that led to treatment discontinuation in >1 patient was [redacted]. | Changes made. |

**Table 3.16: Overview of AEs for all WM patients (four studies)**

| Event |
|---|
| AEs, n (%) |
| Grade ≥3 |
| SAEs |
| AEs leading to death |
| AEs leading to discontinuation |
| TRAEs, n (%) |
| AESIs, n (%) |
| Grade ≥3 |
| Serious |

Source: CS, Table B.2.37, page 69.[4]
AE = adverse event; AESI = adverse event of special interest; n = num
number of patients evaluable; SAE = serious adverse event; TRAE = t
= Waldenström's macroglobulinaemia

The CS reported the most common AEs (CS, pages 70-71), with [redacted] The most frequent [redacted] in the All WM group were: [redacted] TRAEs were reported in the CS on page 71, with the most common (>10%) being [redacted] The most frequent SAEs were [redacted] The CS does not highlight [redacted] in their summary of the SAEs. Also, the CS does not report that neoplasms benign, malignant (including cysts and polyps) were [redacted]; the CS only includes one

| | | | |
|---|---|---|---|
| | | row below this general category 'basal cell carcinoma' which affected [redacted]<br>At least one adverse events of special interest (AESIs) were reported by [redacted] of zanubrutinib-treated patients. AEs within the categories of [redacted] were reported most frequently. Events that met the criteria for seriousness and/or were Grade ≥3 were reported in [redacted] and [redacted] of patients, respectively.<br>The CS also reports that a total of [redacted] reported at least one occurrence of treatment-emergent anaemia. [redacted] patients with treatment-emergent anaemia received red blood cell transfusion within 30 days of onset.<br>**3.2.7.3 Safety conclusions**<br>Overall, zanubrutinib has a comparable safety and tolerability profile compared with ibrutinib. Neutropenia was consistently more prevalent in the zanubrutinib groups compared with the ibrutinib group [redacted]. The zanubrutinib treated patients had a lower rate of several AEs compared with ibrutinib, such as [redacted] There were also fewer AEs leading to [redacted] with zanubrutinib compared with ibrutinib.<br>In a pooled analysis of 253 patients with WM, the most common AEs reported by zanubrutinib treated patents were [redacted] | |

<table>
<tr>
<td>Section 3.2.6<br>p48-50</td>
<td>This section has not been marked up in the ERG report however the Data Cut off (DCO) 31st August 2020 has yet to be presented in the public domain. BeiGene consider it to be AIC</td>
<td>
3.2.6 Updated efficacy results of the ASPEN trial (2020 data)<br><br>
Overall, these data represent 12 months of additional follow-up from the initial data cut-off date (31 August 2019). The median follow-up times on study for patients in Cohort 1, treated with ibrutinib and zanubrutinib, were 31.24 months and 30.78 months, respectively.<br>
A total of [redacted] patients were continuing study treatment: [redacted] patients on the ibrutinib arm and [redacted] patients on the zanubrutinib arm) as of the 2020 data cut-off date.<br>
The most common reasons for treatment discontinuation for patients in Cohort 1 were adverse events [redacted] ibrutinib-treated patients versus [redacted] zanubrutinib-treated patients) and progressive disease [redacted] ibrutinib-treated patients versus [redacted] zanubrutinib-treated patients).<br>
Patient disposition for Cohort 1 (ITT analysis set, 2020 data) are summarised in Table 3.12 below.<br>
Table 3.12: Patient disposition: Cohort 1 (ITT analysis set, 2020 data)
<table>
<tr><th>Category</th><th>Zanubrutinib (n=102)</th><th>Ibrutinib (n=99)</th></tr>
<tr><td>Randomised, not treated, n (%)</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>AE</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Progressive disease</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Treated, n (%)</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>On treatment, n (%)</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Discontinued, n (%)</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>AE</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Progressive disease</td><td>[redacted]</td><td>[redacted]</td></tr>
</table>
</td>
<td>Changes made.</td>
</tr>
</table>

| | | | | | |
|---|---|---|---|---|---|
| Investigator's discretion | [redacted] | [redacted] | | 8 (4.0) |
| Withdrawal by patient | [redacted] | [redacted] | | 6 (3.0) |
| Other | [redacted] | [redacted] | | 5 (2.5) |
| Median study follow-up (months) | [redacted] | [redacted] | | 19.45 |
| Min, Max | [redacted] | [redacted] | | 0.4, 31.2 |

Source: Response to Clarification, Question A10.[6]
AE = adverse event; ITT = intention-to-treat; n = number of patients in the category; N = number of patients evaluable

While close to a [redacted] of patients in Cohort 1 missed at least one efficacy assessment [redacted] in the ibrutinib group, and [redacted] in the zanubrutinib), most patients [redacted] missed assessments at a single visit only. Based on a review of the study data associated with the missed visits, including safety assessments, no important protocol deviations were identified. The impact of missed central lab IgM assessments was minimized by allowing response assessments by the investigator using the local lab IgM results as available across the study, and there was no significant impact to the study efficacy assessments. Taken together, no major impact to study conduct and no impact to study conclusion related to COVID-19 were observed as of the data cut-off date of 31 August 2020 according to the company.[26]

**3.2.6.1 IRC-assessed VGPR/CR rate (primary endpoint) and secondary endpoints**

All 2020 efficacy analyses presented by the company are based on assessments by the investigator.[26] Therefore, IRC-assessed VGPR/CR rate (primary endpoint), IRC-assessed duration of response, IRC-assessed progression-free survival and IRC-assessed time to response (secondary endpoints) are not reported for the data cut-off date of 31 August 2020.

**3.2.6.2 Overall survival (exploratory endpoint, 2020 data)**

In Cohort 1, the median follow-up time was 31.7 months for patients on the ibrutinib arm and 31.5 months for patients on the zanubrutinib arm. The median overall survival had not been reached overall or in treatment naïve patients or patients with relapsed/refractory disease in either treatment arm as of the data cut-off date (Table 3.13 and Figure 3.3).

A total of [redacted] deaths occurred as of the data cut-off date: [redacted] deaths on each of the ibrutinib and zanubrutinib arms. The event-free rate for patients overall on the zanubrutinib arm versus the ibrutinib arm was [redacted] versus [redacted] at 12 months, [redacted] versus [redacted] at 18 months, [redacted] versus [redacted] % at 24 months, and [redacted] versus [redacted] at 30 months. Among patients with relapsed/refractory disease, the event-free rate was higher for patients on the zanubrutinib arm versus those on the ibrutinib arm at 12 months [redacted] through 30 months [redacted]

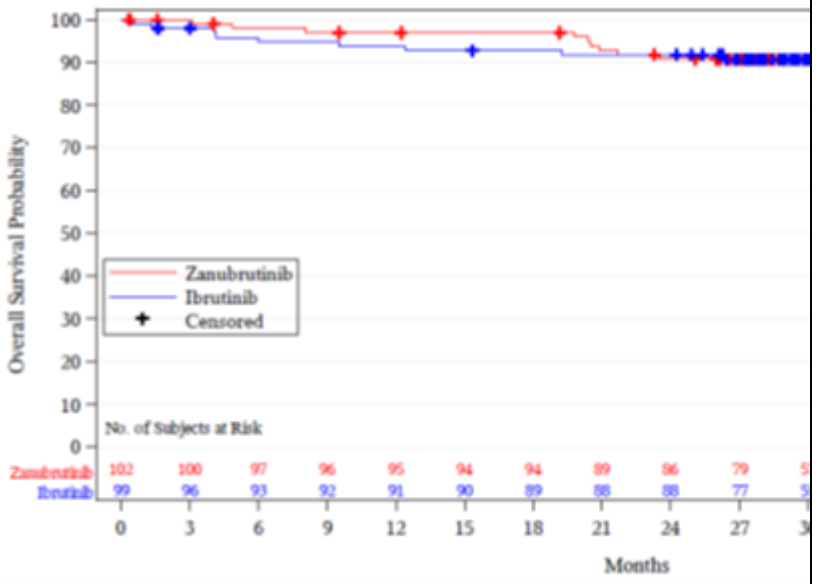
**Table 3.13: OS in Cohort 1 (ITT analysis set)**

| **Assessment** | **Zanubrutinib (n=102)** |
|---|---|
| Median follow-up, months (95% CI) | [redacted] |
| Median OS, months (95% CI) | [redacted] |
| Event-free rate at, % (95% CI) | |
| 12 months | [redacted] |
| 18 months | [redacted] |
| 24 months | [redacted] |
| 30 months | [redacted] |
| Source: Response to Clarification, Question A10.[6] | |

CI = confidence interval; ITT = intention-to-treat; N = number of patients evaluable; NE = not evaluable; OS = overall survival

OS at 30 months was [redacted] among patients treated with zanubrutinib and [redacted] among patients treated with ibrutinib (Figure 3.2).[24]

**Figure 3.3: OS in Cohort 1 (ITT analysis set, 2020 data)**




Source: Response to Clarification, Question A10.[6]

No. = number; OS = overall survival

**3.2.6.3 Other exploratory endpoints**

[redacted] were not reported for the data cut-off date of 31 August 2020.

**3.2.6.4 Adverse events**

| | | | |
|---|---|---|---|
| | | [redacted] were also not reported for the data cut-off date of 31 August 2020 in the document provided by the company.[26]<br>**3.2.6.5 Conclusion**<br>All 2020 efficacy analyses presented by the company are [redacted] | |

NICE National Institute for Health and Care Excellence

# Technical engagement response form

## Zanubrutinib for treating Waldenström's macroglobulinaemia [ID1427]

As a stakeholder you have been invited to comment on the evidence review group (ERG) report for this appraisal.

Your comments and feedback on the key issues below are really valued. The ERG report and stakeholders' responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

## Information on completing this form

We are asking for your views on key issues in the ERG report that are likely to be discussed by the committee. The key issues in the ERG report reflect the areas where there is uncertainty in the evidence, and because of this the cost-effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the ERG report.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the ERG report that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this appraisal, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

NICE National Institute for Health and Care Excellence

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 under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

Deadline for comments by **5pm** on **Friday 18 February 2022**. 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 during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.**

**Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.**

NICE National Institute for Health and Care Excellence

# About you

**Table 1 About you**

| | |
|---|---|
| **Your name** | Dr Robert Mulrooney, General Manager, UK and Ireland |
| **Organisation name: stakeholder or respondent**<br>(if you are responding as an individual rather than a registered stakeholder, please leave blank) | BeiGene UK |
| **Disclosure**<br>Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry. | None |

# Company cover statement

The Company would like to thank NICE and the ERG for the opportunity to review and respond to the technical questions for engagement. The Company's responses are presented in the table below. In addition to the response the Company has presented a revised base-case (Appendix A) which considers the evidence presented as part of this response. Results are presented with a simple discount of [redacted]% on the zanubrutinib list price. Deterministic, probabilistic and key scenario analyses accompany the revised base-case.

# Key issues for engagement

**All**: Please use the table below to respond to the key issues raised in the ERG report.

**Table 2 Key issues**

| Key issue | Does this response contain new evidence, data or analyses? | Response |
|---|---|---|
| **Key issue 1:**<br>The comparators are not in line with the NICE scope | Yes | Given the large variety of treatment options in Waldenström's macroglobulinaemia (WM), it is challenging to identify and validate an appropriate treatment pathway for patients with WM in the United Kingdom (UK).<br>This challenge is not uncommon in blood cancer. Choice of therapy for patients with WM is often highly personalised, and determined by patient's age, fitness, $MYD88^{MUT}$ status, prior therapies, and existing comorbidities.<br>As discussed in Document B, Section B.1.3.5.2, in order to establish a treatment pathway which is reflective of UK clinical practice, the Company utilised real-world evidence from the 2018 UK Rory Morrison Registry report - a research project managed by WM UK, which collects data from over 20 hospitals, comprising 926 patients with WM patients.[1] It was found that: |

| | | |
|---|---|---|
| | | • The two most widely used treatment options in the first-line setting were dexamethasone plus rituximab plus cyclophosphamide (DRC) (16.2%) and bendamustine plus rituximab (BR) (13.1%).<br>• The three most widely used treatments in the second-line setting were ibrutinib (18.2%), DRC (6.7%) and BR (6.1%).<br>Therefore, it was deemed appropriate to include ibrutinib, DRC and BR within the decision problem.<br>While the Company acknowledge that this approach does not include all treatments within the National Institute for Health and Care Excellence (NICE) scope, we consider that the included comparators reflect standard of care for the vast majority of patients with WM in the UK.<br>The Company acknowledge the ERG decision to remove ibrutinib from the model in line with the NICE guidance for appraisals. However, the Company would like to highlight that whilst ibrutinib is not routinely commissioned it is standard of care for patients with relapsed/refractory (R/R) WM, with 65% of patients recorded to have received ibrutinib between 2017 and 2020, as reported by the 2021 UK Rory Morrison Registry report.[2]<br>Ibrutinib is currently under review by NICE following its time within the Cancer Drugs Fund (CDF), and whilst the Company acknowledge that ibrutinib cannot be included as a comparator in this appraisal until routine commissioning is recommended by the NICE, we request that the evolving ibrutinib appraisal is monitored by the NICE team in relation to this appraisal.<br>***BR and DRC represent established clinical practice in the UK***<br>Focusing on the inclusion of non-Bruton Tyrosine Kinase Inhibitors (non-BTKi) comparators, the inclusion of DRC and BR is supported by the 2018 European Society of Molecular Oncology (ESMO) WM guideline, the 2021 British Society for Haematology (BSH) WM guideline, and the UK 2021 Rory Morrison Registry report:<br>• The 2018 ESMO WM guideline recommends the use of both DRC and BR across all symptomatic treatment naïve patients (both fit and unfit, and those |

with high or low tumour burden). The guidelines note that *"Anti-CD20-based (rituximab-based) combinations are the mainstay of first-line treatment"*. For R/R patients the guideline recommends that (aside from ibrutinib) rituximab-based regimen are considered for all patients.[3]

- The 2021 BSH WM guideline recommends the use of DRC and BR as front-line treatments, noting that *"rituximab combination therapies are the cornerstone of first-line treatment in WM with response rates typically over 80%. The two most commonly used first-line regimens are dexamethasone, rituximab and cyclophosphamide and rituximab–bendamustine".* For R/R disease the guidelines recommend that rituximab-containing regimens are considered.[4]
- The UK 2021 Rory Morrison Registry report (Table 1) indicates that of the treatments list when the NICE scope 85% of patients received either a bendamustine-based regimen (i.e., BR) or DRC between 2015 and 2020 in the first-line setting. When considering the second-line setting, 77% of patients received either BR or DRC between 2017 and 2020.[1]

Furthermore, clinical expert opinion obtained by the Company during this Technical Engagement stage supports the inclusion of BR and DRC as the two main treatments (aside from ibrutinib) within UK clinical practice for patients with WM.

**Table 1: First- and second-line treatment regimens extracted from the 2021 UK Rory Morrison report**

| Treatment regimen | N=158 | N=13 | Weighted average |
|---|---|---|---|
| | First-line | Second-line | |
| BR | 40% | 54% | 41% |
| DRC | 45% | 23% | 43% |
| FCR | 0% | 0% | 0% |
| FR | 0% | 0% | 0% |
| Clad-R | 0% | 0% | 0% |
| SCT | 0% | 0% | 0% |
| Chlorambucil monotherapy | 4% | 8% | 5% |

| | | |
|---|---|---|
| Rituximab monotherapy | 11% | 15% | 11% |

Abbreviations: BR, bendamustine plus rituximab; Clad-R, cladribine plus rituximab; DRC, dexamethasone + rituximab + cyclophosphamide; FCR, fludarabine/cyclophosphamide/rituximab; FR, fludarabine/rituximab; SCT, Stem cell transplant. Note. Regimens adjusted to reflect only those included within the NICE scope

***European and British guidelines, combined with registry data highlight the lack of use of remaining scoped comparators***

A review of the guidelines from ESMO, BSH and the UK 2021 Rory Morrison Registry report was conducted, and validation was sought from a UK clinical expert. Clinical feedback obtained indicated that the alternative treatments listed in the NICE scope were either outmoded (e.g purine analogues [e.g. fludarabine or cladribine]), were of limited use in UK clinical practice (e.g rituximab monotherapy) or could be considered emerging treatment options which were not firmly established in UK clinical practice (e.g analogous stem cell transplant [ASCT]) at present. In addition, the clinical expert acknowledged the heterogeneity of evidence across these alternative comparators, highlighting the challenge of constructing a robust comparison with zanubrutinib. This clinical feedback is supported by ESMO and BSH WM guidelines, and the 2021 treatment data presented in the Rory Morrison Registry report (please refer to Table 7, Appendix B for further details). Based on the above findings, the Company consider that the remaining comparators included within the NICE scope are not currently established standard of care within UK clinical practice, and hence are not relevant comparators for zanubrutinib in this appraisal.

***A standard of care comparator comprising BR and DRC is in line with the Physician's Choice comparator in the appraisal of ibrutinib for R/R WM (TA491)***

Within the NICE appraisal of ibrutinib (TA491), a pooled basket of treatment options ("Physician's Choice") formed the comparator arm for the appraisal. Clinical efficacy data for which were obtained from a European Chart Review.[5]

The chart review included n=454 patients over five lines of therapy. A matched adjusted indirect comparison (MAIC) was conducted to produce comparative effectiveness estimates of ibrutinib relative to the "Physician's Choice" comparator arm. The

| | | |
|---|---|---|
| | | breakdown of treatments received by patients included in the chart review was not reported in TA491. Within the economic model Janssen assumed (based on clinical expert opinion) that the majority of the second-line patients would be receiving either BR or DRC (78%).[5] This assumption was accepted by NICE and the ERG, indicating that both must have been suitability satisfied with this assumption, indicating that it could not have been too far away from the true distribution of treatments within the chart review.<br><br>Additional treatments included within the chart review (fludarabine/cyclophosphamide/rituximab [FCR] and cladribine plus rituximab) are (according to European and British guidelines, and UK patient numbers) not widely used in clinical practice and have known tolerability issues. The inclusion of less effective and tolerable treatments in the basket of "Physicians Choice" may therefore bias cost-effectiveness results in favour of ibrutinib.<br><br>That being said, and as noted in Key Issue 16, the Company do acknowledge the importance of comparing zanubrutinib with standard of care in a single comparison and propose this is aligned with the methodology adopted in TA491, with the exception that comparators not used in routine clinical practice (as previously specified) are not included within the definition of standard of care.<br><br>In light of this, the cost-effectiveness results of zanubrutinib versus standard of care (consisting of 49% BR and 51% DRC based on the UK 2021 Rory Morrison Registry report, adjusted to sum to 100%) have been weighted to produce an overall incremental cost-effectiveness ratio (ICER) of zanubrutinib versus standard of care.<br><br>For the Company's revised base-case (see Appendix A), when compared to standard of care, zanubrutinib is associated with £[redacted] additional costs and [redacted] additional quality adjusted life years (QALY), corresponding to an ICER of £20,054 per QALY gained. |
| **Key issue 2:**<br><br>Patients with cardiovascular disease and those taking warfarin | No | ***Patients taking warfarin are not within the target population of this appraisal***<br><br>The Company would like to highlight that the exclusion of patients taking warfarin in the ASPEN trial is in line with the European licensed population for zanubrutinib: |

| | | |
|---|---|---|
| were excluded from the ASPEN trial | | • *"Warfarin or other vitamin K antagonists should not be administered concomitantly with BRUKINSA."* [6]<br><br>As such, these patients would not be eligible for treatment with zanubrutinib in UK clinical practice and hence do not fall within the target population of this appraisal.<br><br>***The exclusion of cardiovascular disease patients from ASPEN was aligned to the exclusion criteria typically applied to clinical studies with investigational drugs ie currently active, clinically significant cardiovascular disease, QTcF prolongation, and active, clinically significant ECG abnormalities.***<br><br>Exclusion criteria of ASPEN were formed according to the BTKi class safety profile which was known at the start of the study. Atrial fibrillation, atrial flutter, and cases of ventricular tachyarrhythmia and cardiac failure have been reported in patients treated with ibrutinib. Cardiac arrhythmia, mainly presented as atrial fibrillation and flutter, is a potential risk in treatment with BTKi and patients with risk factors (hypertension, acute infections, history of atrial fibrillation, pre-existing cardiovascular disease) are more likely to experience it.<br><br>In the ASPEN clinical trial, risk factors for atrial fibrillation were balanced across the study arms in Cohort 1. These included prior medical history of atrial fibrillation or flutter, hypertension, or diabetes mellitus. Of the 12 zanubrutinib-treated patients with a history of atrial fibrillation/flutter (10 in Cohort 1 and 2 in Cohort 2), there were no cases in which their history of atrial fibrillation/flutter worsened and became an adverse event while on treatment. Conversely, of the 8 patients with a history of atrial fibrillation randomised to ibrutinib treatment, 3 (37.5%) developed an adverse event of atrial fibrillation. In the ASPEN clinical trial zanubrutinib has demonstrated a favourable cardiac safety profile compared to ibrutinib. |
| **Key issue 3:**<br><br>The evidence for treatment naïve patients is based on small | Yes | The Company acknowledge that there are only a small number of treatment naïve patients within the ASPEN trial, but do not agree that it limits the generalisability of the ASPEN trial data to the UK. |

<table>
<tr>
<td>numbers of patients and has limited generalisability</td>
<td></td>
<td>
Historically, treatment naïve patients have a better prognosis than patients with R/R WM:
<ul>
<li>Data from a European chart review demonstrates a trend in decreasing progression-free survival (PFS) with each line of therapy (see Table 2).[5]</li>
<li>PFS and overall survival (OS) landmark rates in Castillo et al. 2021 (treatment naïve ibrutinib WM trial) are greater than PFS and OS landmark rates in Treon et al. 2021 (R/R ibrutinib WM trial).[7,8]</li>
<li>Evidence from ASPEN indicates a comparable treatment effect for zanubrutinib across both treatment naïve patients and R/R patients, with a similar proportion of patients achieving a very good partial response (VGPR) (treatment naïve – 26% vs. R/R – 29%).[9]</li>
</ul>
Given the historical improved prognosis of treatment naïve patients and consistent treatment effect of zanubrutinib across all patients, the Company consider the current cost-effectiveness results for zanubrutinib to be conservative.<br><br>
If the treatment naïve population was larger in the ASPEN trial, as it may well be in the UK, it is anticipated that the cost-effectiveness of zanubrutinib would improve, and as such the Company consider that the only limitation in terms of generalisability may be that the cost-effectiveness of zanubrutinib is underestimated.<br><br>
Table 2: European chart review - Median PFS by treatment line
<table>
<tr><th>Treatment line</th><th>Median PFS (months)</th></tr>
<tr><td>Front-line (n=454)</td><td>29 (25-31)</td></tr>
<tr><td>2nd line (n=387)</td><td>23 (20-26)</td></tr>
<tr><td>3rd line (n=160)</td><td>16 (10 -18)</td></tr>
<tr><td>4th line (n=61)</td><td>11 (8-15)</td></tr>
<tr><td>5th line (n=26)</td><td>14 (7-29)</td></tr>
</table>
</td>
</tr>
</table>

| | | |
|---|---|---|
| | | Abbreviations: PFS, progression-free survival. Source: CDF review of ibrutinib TA491 ACM public committee slides[5] |
| **Key issue 4:**<br>Survival data for zanubrutinib are immature | Yes | The Company acknowledges that survival data from the ASPEN are immature. Nonetheless, this is common in oncology, with a recent paper reporting that at least 41% of NICE's cancer technology appraisals use immature data to inform reimbursement decisions.[10]<br><br>Furthermore, given that historically patients with WM have reached a median survival of 18.5 years, it is expected that data from a median follow-up of 19.5 months would remain immature for any intervention assessed in the area.[1]<br><br>Nonetheless, among patients treated with zanubrutinib in ASPEN, OS was 97.0% and 89.5% at 12 months and 24 months, respectively. Comparatively, among patients treated with ibrutinib, OS was 93.9% and 91% at 12 months and 24 months, respectively. This provides an early indication that treatment with a BTKi is beneficial, and that zanubrutinib has more favourable survival estimates compared to ibrutinib.<br><br>In response to ERG clarification questions (question B5), the Company explained that the immaturity of publicly available survival data (median follow-up 37-months) from the Phase 2 Study 118E of ibrutinib did not allow for an informative validation of zanubrutinib long-term survival extrapolation.<br><br>Since this response, extended long-term follow-up has become publicly available (follow-up 59-months) from Study 118E for ibrutinib in patients with R/R W/M.[8] As such, to further address uncertainty in survival data, Kaplan-Meier (KM) plots obtained from Study 118E were digitised and long-term survival extrapolated as demonstrated in Figure 2 and Figure 3 in Appendix C. This analysis indicates that long-term survival as a result of treatment with ibrutinib should be expected, with mean extrapolated undiscounted survival ranging from 18.40 to 18.88 years (considering all-cause mortality, Table 8, Appendix C).<br><br>A further comparison of OS KM data for ibrutinib (from Study 118E), BR (from Tedeschi et al. 2015[11]) and DRC (from Kastritis et al. 2015[12]) demonstrates a clear difference in |

| | | |
|---|---|---|
| | | survival of patients on a BTKi compared to chemotherapy (Figure 4). Whilst the OS data for zanubrutinib is immature, it is comparable to the long-term ibrutinib OS data.<br><br>The long-term survival benefits for ibrutinib are further supported by feedback from clinical experts at the recent CDF Appraisal Committee Meeting (ACM) of ibrutinib (TA491), who indicated that at least two-thirds of patients whose disease progressed while on ibrutinib achieve a good response to further lines of chemotherapy.[13] They also noted that the median time between disease progression and death in clinical practice is much longer than the year modelled in the TA491 CDF review Company submission. Accordingly, this suggests that OS is prolonged in the long-term, given that patients who receive earlier, effective treatments are more likely to survive longer.<br><br>The Company expect the long-term survival benefit following treatment with zanubrutinib will demonstrate improvements compared to ibrutinib, given that:<br><br>1. Zanubrutinib is a second-generation BTKi with improved selectively and less off target effects.<br><br>2. The ASPEN trial has demonstrated comparable efficacy and a superior safety/tolerability profile to first-generation BTKi, ibrutinib. UK clinical expert feedback indicated that practically an improved tolerability profile will allow patients to remain on the full dose for longer, hence allowing patients to obtain more benefit from treatment with zanubrutinib.<br><br>3. KM OS results at 12-months and 24-months are more favourable for zanubrutinib compared to ibrutinib in the ASPEN study.<br><br>Based on the evidence available from Study 118E and the feedback from clinical experts in attendance at the CDF review of TA491 (ibrutinib in WM), the Company strongly ascertain that long-term survival benefits compared to comparator chemotherapies should be expected following treatment with zanubrutinib.[8,13] |
| **Key issue 5:**<br><br>The indirect comparisons with rituximab and bendamustine (BR) and dexamethasone, rituximab | Yes | Matched adjusted indirect treatment (MAICs) are commonly adopted within oncology appraisals, where there is a need to estimate comparative effectiveness of treatments in the absence of head-to-head evidence. |

| | | |
|---|---|---|
| and cyclophosphamide (DRC) are unreliable | | Within similar appraisals for BTKis in blood cancers (TA491, ibrutinib for WM[14]; TA689, acalabrutininb for chronic lymphocytic leukaemia [CLL][15]) the methodology has been adopted and accepted by NICE and has led to positive recommendations in both cases, despite the known limitations of the MAIC method.<br><br>In order to address this issue the Company have engaged with a statistical expert from The School of Health and Related Research's Technology Assessment Group (SCHARR) ERG to validate and ratify the methods used for the indirect treatment comparisons.<br><br>**The MAIC analyses provide robust evidence to compare zanubrutinib with comparator treatments.**<br><br>While the Company acknowledges the potential risk of bias in the MAICs, it strongly believes that with no head-to-head studies and no means to conduct a network meta-analyses the MAIC analyses represent a robust method to compare zanubrutinib with comparator treatments.<br><br>The ERG note that "*the MAIC analyses represent the best use of the available evidence to compare zanubrutinib with comparator treatments (other than ibrutinib)*." Furthermore, the ERG note that there is "*little evidence to suggest substantial bias in favour of zanubrutinib*", indicating that the MAICs could be underestimating the effectiveness of zanubrutinib, and hence the results could be deemed conservative. This has indeed been proven when considering alternative, potentially less biased methods, as explained later on.<br><br>In response to ERG clarification A21 (request for a comparison of baseline characteristics of patients in the three studies included in the MAIC), the Company provided additional MAICs in treatment naïve patients, thereby addressing further bias in the results. These included three additional analyses for BR and three additional analyses for DRC. The majority of the results from the Company's submission are in line with the additional MAICs from the clarification response, which included only treatment naïve patients. |

| | | |
|---|---|---|
| | | **Comparison with treatment naïve populations only will likely favour comparator treatments.**<br><br>The Company acknowledge the ERG's concern surrounding prior treatment status, and how it may lead to further risk of bias. However, ASPEN enrolled both relapsed and naïve patients, so comparison with treatment naïve populations with ASPEN will likely favour comparator treatments due to the improved prognosis of treatment naïve patients (see response to Key Issue 3).<br><br>**The analyses matched on all feasible variables within the limits of the published comparator trial data**<br><br>The ERG raised additional issues related to bias that had been caused by the reduction of continuous variables to binary or categorical variables.<br><br>However, as highlighted by the ERG itself, the variables (and how they were expressed) which were matched across the MAICs were restricted by data reported in the comparator trials.<br><br>It is not feasible for the Company to use continuous values for various matching variables such as age, IgM concentration, platelet count and number of prior therapies, as the data for such variables in published comparator trials are either expressed as the median, or in binary or categorical terms.<br><br>Furthermore, the Company sought to identify and match on all feasible prognostic and treatment effect modifiers. The ERG's suggestion to consider variables such as socio-economic status, year of study, location of study and general health of patients, is considered unreasonably challenging by the Company given that trials do not commonly record such variables within their baseline patient characteristics.<br><br>**Simulated treatment comparisons (STCs) may produce less biased results as they rely on extrapolation rather than reweighting** |

| | | |
|---|---|---|
| | | NICE DSU Technical Support Document 18 states that an important property of population adjustment methods is that they require sufficient overlap between the population of the individual patient data (IPD) study and the aggregate study.[16]<br><br>Despite this, the MAICs reported in the original Company submission led to a decrease in the effective sample size from 102 to approximately 50 (50 and 53 for comparisons with BR and DRC, respectively).<br><br>According to Phillippo et al, regression-based approaches such as STC are not restricted to scenarios with sufficient overlap, and as such, STCs may provide a more robust alternative to MAICs, leading to less bias.[17]<br><br>Therefore, to complement the MAIC, the Company performed an additional indirect treatment comparison utilising the STC methodology. STC's were performed to indirectly compare zanubrutinib with BR and with DRC separately. Please see Appendix H for details of the methodology.<br><br>Results of the STC's (Table 11 and Table 12, Appendix H for BR and DRC, respectively) demonstrate that compared with BR, zanubrutinib was associated with statistically significantly improved PFS (hazard ratio [HR]: [redacted] and statistically significantly improved OS (HR: [redacted]).<br><br>Compared with DRC, zanubrutinib was associated with statistically significantly improved PFS (HR: [redacted] and statistically significantly improved OS (HR: [redacted]).<br><br>As highlighted in Appendix H, time points used to predict the survival probability were determined based on where there was at least one event in the zanubrutinib IPD arm. As a result, not all prognostic factors and effect modifiers could be considered in each STC since there were cases were no events occurred in patients with certain baseline characteristics.<br><br>Even though this may increase the risk of bias, following consultation with a statistical expert from the SCHARR ERG, the Company believes the STC to produce more reliable and less biased results than the MAIC given the increased sample size. Based on HRs obtained from the STC, when compared to standard of care in the revised |

| | | |
|---|---|---|
| | | Company base case, zanubrutinib is associated with £[redacted] additional costs and [redacted] additional QALYs, corresponding to an ICER of £20,054 per QALY gained.<br><br>Irrespective of the methods used to determine the comparative effectiveness of zanubrutinib with BR and DRC, both show that zanubrutinib is associated with statistically significant improved PFS and improved OS. Moreover, the STC shows statistically significant longer OS for zanubrutinib. Considering the above points, the Company believes that through following NICE DSU guidance, the Company has provided the most reliable and unbiased analysis based on the evidence and methods available. |
| **Key issue 6:**<br><br>The choice of a partitioned survival model and its underlying assumptions | Yes | The Company considered both the partitioned survival model (PSM) and state-transition model (STM) structure during the model conceptualisation phase. Based on the reasoning presented in Document B, Section B.3.2, the PSM was selected as the most appropriate structure and was therefore used for the cost-effectiveness model.<br><br>**The PSM approach is widely used in oncology and understood by health economist and clinicians**<br><br>This PSM approach is consistent with the approaches adopted in the majority of economic evaluations submitted to the NICE for the health technology assessment (HTA) of treatments for lymphoma.[18–20] PSMs are extensively and routinely used to model the costs and outcomes of oncology treatments in the UK and globally across HTA bodies. In a recent review by NICE, it was found that 73% of 30 recent oncology appraisals assessed by NICE used a PSM.[21]<br><br>PSMs are well understood by both clinicians and health economists due to their straightforward approach and their implementation of commonly used and well understood endpoints taken directly from the trials (OS and PFS).<br><br>The ERG evaluating this submission, Kleijnen Systematic Reviews Ltd, had the same issue about the Company choosing a PSM over a STM in the technical appraisals for tepotinib and nivolumab.[22,23] In both cases, the Committee deemed the Company's model structure, PSM, acceptable for decision-making.[24,25] |

| | | |
|---|---|---|
| | | **The STM methodology has its own limitations**<br><br>Although the NICE DSU Technical Support Document 19 recommends presenting a STM alongside the PSM to assist in verifying the plausibility of the PSM extrapolations, there are key limitations associated with STM and the Company do not believe constructing a STM would alleviate the associated uncertainty.[26]<br><br>Firstly, STMs have unclassified endpoints to model transitions such as post-progression survival. As such, the model transitions are highly prone to bias due to the selection effects and informative censoring. Given that the data available from ASPEN, the pivotal, phase III study supporting this submission, is based on a small number of patients, the use of such data to inform post-progression survival could be misleading.<br><br>Extrapolating outcomes from a group of patients who no longer have comparable characteristics and based on patients who are progressing early would be biased against the zanubrutinib arm. This is not an issue when using OS directly as time to death from randomisation for all patients contributes to the extrapolation.<br><br>Secondly, STMs do not negate the need to extrapolate data, therefore extrapolating immature data (such as post-progression survival) produces uncertain estimates for those particular transition probabilities and hence creating uncertain OS projections from the final model outputs.<br><br>**Creating an STM based on the primary endpoint of ASPEN would rely on estimation of surrogacy between response and OS**<br><br>The Company performed survival analysis on secondary and exploratory endpoints, PFS and OS, to extrapolate treatment effectiveness for zanubrutinib and comparators beyond the available trial data. In the clarification questions, the ERG asked the Company to justify why health state occupancy was based on secondary and exploratory endpoints from ASPEN rather than the primary endpoint, response rate.<br><br>Response rates are used in STMs to generate transition probabilities and subsequently the long-term treatment effectiveness. However, recent oncology publications have |

| | | |
|---|---|---|
| | | assessed the validity of using response rates as surrogates for PFS and OS and found no correlation between the endpoints.[27,28] PSMs makes use of the PFS and OS data directly from the trial, ensuring that estimated survival outcomes versus observed outcomes are matched.<br><br>**Development of an STM would be unnecessarily complex and would increase uncertainty**<br><br>In addition to the above, further limitations to a STM involve underlying data availability and complexity of the approach to allow for all possible transitions within the cost-effectiveness model itself.<br><br>For a STM, the development of a three-health state model using time dependencies in event rates for each possible transition would add significant complexity based on the number of tunnel states that would be required to accurately model the transitions.<br><br>This would create unnecessary computational complexity that would potentially make the model burdensome to run. The most recent and relevant NICE appraisal for this submission is TA491, ibrutinib for the treatment of patients with R/R WM.[14] Within TA491, the Company used a STM in their submission which the ERG considered "*too complex*" and introduced "*considerable uncertainty*". The Committee also noted in the recent appraisal consultation document (ACD), that the STM produced clinically implausible outputs.[29] This critique implies that a STM would be an inappropriate model choice for zanubrutinib.<br><br>Based on the above points, the Company believe that providing a STM would not offer the decision makers any additional certainty concerning the cost-effectiveness of zanubrutinib and therefore have not generated a STM. |
| **Key issue 7:**<br><br>The model does not include all comparators mentioned in the NICE scope | Yes | *Please refer to response to Key Issue 1* |

| | | |
|---|---|---|
| **Key issue 8:**<br>Ibrutinib should not be included as direct comparator and/or subsequent treatment option in the economic model | Yes | As highlighted in response to Key Issue 1, the Company acknowledge the challenge of including ibrutinib as a direct comparator within the model.<br><br>However, when considering the subsequent treatment pathway of patients with WM in the UK, ibrutinib is clearly part of the treatment paradigm:<br><br>• When subsequent treatments are applied within the model, it is considering a patient that is approximately 8 years after their time in the progression-free health state.<br><br>• Ibrutinib is under review by NICE for routine commissioning within the NHS for patients with R/R WM. The ACD for ibrutinib stated that clinicians considered ibrutinib a step change in the management of WM.[29]<br><br>• Data from the 2021 Rory Morrison Registry report highlights that the majority of patients (65%) are receiving ibrutinib in the second-line setting in the UK.[2] These data clearly demonstrate that ibrutinib is standard of care for R/R patients in the UK.<br><br>• Given the vast uptake of ibrutinib in clinical practice, the Company consider it unrealistic to ignore the existence of ibrutinib, and whilst it may not be a comparator (given the scope of this appraisal), as of today, it is a subsequent treatment for WM patients following PFS.<br><br>Therefore, the ERG's decision to remove ibrutinib from the modelled subsequent treatment pathway is not in line with UK clinical practice.<br><br>Furthermore, there is a technical issue with the ERG's preferred base case methodology, in that only the costs of ibrutinib subsequent treatment following progression on BR or DRC are removed, whilst no attempt is made to remove the benefits (i.e. survival benefit) of ibrutinib subsequent treatment.<br><br>Given that ibrutinib has been shown to be delay progression and death in the R/R WM setting,[8] the ERG's decision to not reduce survival in tandem with a reduction in costs will produce results bias in the favour of the BR and DRC arms. |

| | | |
|---|---|---|
| | | To reflect the impact of this bias, the Company have ran exploratory scenario analyses which reduces the post-progression survival across the BR and DRC treatments for the ERG-corrected Company corrected base case **[Note this is not the Company's preferred base case]**. These scenarios demonstrate that zanubrutinib is more cost-effective versus a true standard of care arm when more reflective post-progression survival is modelled for BR and DRC (see Table 5).<br><br>Given the arbitrary nature of the exploratory analyses and the inherent methodological issues of the ERG's preferred base case, ibrutinib subsequent treatment should not be removed from the treatment paradigm – aligning with current UK clinical practice. This is reflected in the Company's revised base case (Appendix A). |
| **Key issue 9:**<br><br>The partitioned survival analysis chosen by the company relies on estimates for progression-free survival (PFS) and overall survival (OS), secondary and exploratory endpoints respectively. Only a small number of PFS and OS events had occurred at the time of this appraisal. | No | *Please refer to response to Key Issue 4 and 6.* |
| **Key issue 10:**<br><br>Plausibility of OS hazards falling below background mortality hazards. | Yes | The Company acknowledge the concern raised by NICE and the ERG surrounding the plausibility of the OS hazards within the model.<br><br>The Company would like to first highlight that the modelled OS hazards across treatment arms **never fall below** the background mortality hazard as this is prevented from occurring within the economic model. However, in order to address this concern, the Company have engaged with a clinical expert, reviewed the observed OS hazards across treatment arms, and explored of more flexible survival analyses. |

| | | |
|---|---|---|
| | | **Effective management of patients with WM can result in patients achieving normal life expectancy.**<br><br>Compared with other types of cancer, WM is a relatively slow progressing disease. Historically patients with WM have reached a median survival of 18.5 years,[1] even in the absence of targeted and selective therapies (such as BTKis).<br><br>In addition, in Committee Papers for TA491 it was stated "*nearly half of people diagnosed with WM die from causes unrelated to WM*".[14] The ERG allude to this type of survival trajectory within their report stating:<br><br>*"The ERG questioned whether this is simply an artifact of data immaturity, or whether low mortality hazards in the long run indicate that there is a subgroup of patients with WM that are at particular risk of dying in the first years into the modelled disease trajectory, whilst the average patient has closer to normal life expectancy".*<br><br>In order to explore this hypothesis, the Company sought clinical expert opinion on the plausibility of patients with WM achieving normal life expectancy. Feedback obtained indicated that for a patient who was diagnosed at approximately 70 years (the majority of patients in the UK are diagnosed between 60-70 years,[2] aligning with the baseline mean age in ASPEN) it would be clinically reasonable for this patient to achieve a normal life expectancy. The Company acknowledge for the minority of patients diagnosed at a younger age, achieving a normal life expectancy may be less likely. However, note that this submission should consider the average age of a patient with WM.<br><br>**The use of more flexible models may better align with the observed OS hazard functions**<br><br>To understand and hence model the OS hazard for zanubrutinib, BR, and DRC, observed hazard functions were generated by treatment arm:<br><br>• Zanubrutinib OS (matched to BR) (Figure 15) – the smoothed observed hazard appears to [redacted] The observed hazard should be interpreted with caution given that only a low number of events have occurred. |

| | | |
|---|---|---|
| | | • BR OS (Figure 16) – the smoothed observed hazard for DRC initially steeply increases up to approximately 25 months, before decreasing up until just before 50 months.<br>• Zanubrutinib OS (matched to BR) (Figure 17) – the smoothed observed hazard appears to [redacted] The observed hazard should be interpreted with caution given that only a low number of events have occurred.<br>• DRC OS (Figure 18) – the smoothed observed hazard for DRC initially increases up to just before 50 months, before decreasing just to after 50 months. Following which a clear steep increase in the hazard is observed up until around 100 months, before the hazard begins to decrease.<br>It is clear from Figure 15 to Figure 18 that the standard parametric analyses may not appropriately capture and hence reflect the observed hazard functions for OS. Therefore, the Company have explored additional flexible extrapolation methods in order to present an alternative to standard parametric modelling.<br>As per the methods discussed in Royston and Parmar 2002,[30] three types of flexible survival models were fitted to the OS endpoint: Hazards, Odds and Normal. Up to three knots (k) were evaluated (k=1, 2 and 3). The models for k=0 were not included given that these models are equivalent to the Weibull, log-logistic and log-normal distributions.<br>**Flexible models for zanubrutinib vs. BR do not provide a better fit than standard parametric models**<br>• All models were successfully fitted for BR OS. There were issues with convergence and optimisation in the zanubrutinib (matched to BR) arm, which can most likely be attributed to low numbers of events. This meant that models for k=2 and 3 were not produced for zanubrutinib (matched to BR).<br>• Zanubrutinib (matched to BR) and BR OS curves intersected for k=1 models with zanubrutinib OS falling below BR OS. This was deemed clinically implausible and as such flexible survival models were not considered any further for the comparison versus BR. |

| | | |
|---|---|---|
| | | **Flexible models for zanubrutinib vs. DRC provide clinically realistic alternatives to standard parametric models**<br>• All models were successfully fitted for DRC OS. There were issues with convergence and optimisation in the zanubrutinib (matched to DRC) arm, which can most likely be attributed to low numbers of events. This meant that models for Hazards (k=2 and 3), Odds (k=2) and Normal (k=1 and 2) were not produced for zanubrutinib (matched to DRC).<br>• The following models were successfully fitted across both zanubrutinib (matched to DRC) and DRC: Hazards (k=1), Odds (k=1 and 3) and Normal (k=1).<br>• The Odds k=3 curve resulted in a clinically implausible extrapolation for zanubrutinib and hence was not considered further. Of the two remaining models the Hazards k=1 model had the lowest Akaike Information Criterion (AIC) (combined across treatment arms), however there was less than 1 AIC point between the two (476.43 vs. 477.02).<br>The extrapolated hazard function for both models are presented in Figure 19 and Figure 20, respectively. The zanubrutinib extrapolated hazard function is relatively similar between model choices. The DRC hazard function plateaus after an initial increase for the Hazards k=1 model, in comparison to declining after the same initial increase in the Odds k=1 model. The zanubrutinib OS hazard appears consistent across both the Hazards k=1 and Odds k=1 models. The Odds k=1 model better captures the observed decrease in hazard function for the DRC arm. Therefore, the Odds K=1 model was selected to model OS for both treatment arms.<br>When considering the Odds k=1 curve for the extrapolation of OS for the DRC comparison, when compared to standard of care in the Company revised base case (weighted pairwise analyses – see Key Issue 1 and 16), zanubrutinib is associated with £[redacted] additional costs and [redacted] additional QALYs, corresponding to an ICER of £20,054 per QALY gained. |
| **Key issue 11:** | Yes | The Company acknowledge that their base-case analysis includes data from patients with $MYD88^{MUT}$ only. This was deemed appropriate given that this population |

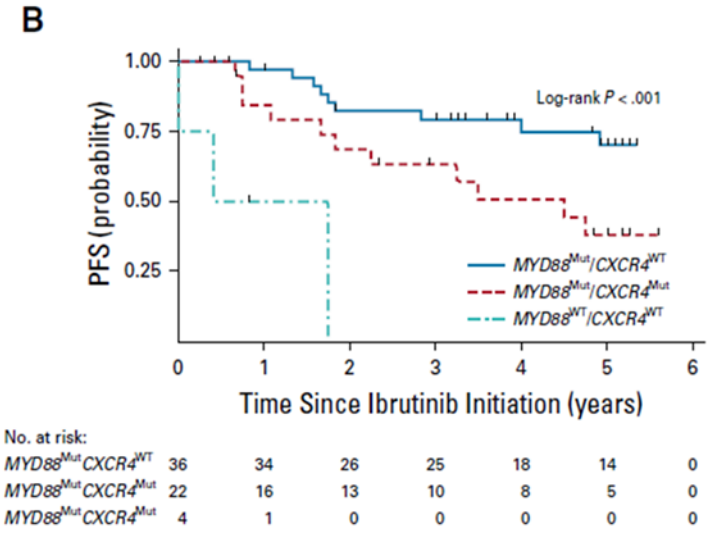
| | | |
|---|---|---|
| The use of data from patients with *MYD88*$^{MUT}$ only. | | represented the randomised cohort of the ASPEN trial (with MYD88$^{WT}$ patients included in a non-randomised single arm cohort of the ASPEN trial only) and are reflective of 90% of the UK WM population (ERG report page 20 – *"90% of MYD88$^{MUT}$ and 5-10% of MYD88$^{WT}$"*).<br><br>This decision was further endorsed within the NICE Technical Engagement meeting (28th January 2022) in which the meeting Chair and the ERG acknowledge that MYD88$^{MUT}$ patients represent the majority of patients with WM in the UK.<br><br>In response to ERG clarification B6 a cost-effectiveness analysis was presented which included all patients from the ASPEN trial (Cohort 1 [MYD88$^{MUT}$] and Cohort 2 [MYD88$^{WT}$]). The weighting of patients for this analysis was derived from the patient numbers enrolled in ASPEN. The results of this analysis demonstrated that the cost-effectiveness of zanubrutinib is consistent across ASPEN Cohort 1 and the pooled analyses.<br><br>Whilst the Company acknowledge the ERG's request to perform a weighted analysis based on the anticipated proportion of MYD88$^{MUT}$ to MYD88$^{WT}$ patients in UK clinical practice (90-95% vs. 5-10%), this would be unnecessary for the following reasons:<br><br>• Across both zanubrutinib arms in Cohort 1 and Cohort 2 the ASPEN clinical trial included approximately an 80%:20% split of MYD88$^{MUT}$:MYD88$^{WT}$ patients.<br><br>• MYD88$^{WT}$ patients historically have a poorer prognosis than MYD88$^{MUT}$ patients with:<br><br>o A lower proportion of Cohort 2 patients achieving a VGPR or partial response (PR) compared to Cohort 1 (Table 3).<br><br>o PFS for MYD88$^{WT}$ patients is shorter than for MYD88$^{MUT}$ patients in Study 118E (ibrutinib in R/R WM) (Figure 1).<br><br>• A weighted analysis would rebalance the ASPEN data to include slightly fewer MYD88$^{WT}$ patients, hence would improve the clinical outcomes of the pooled patient population. This is expected to improve the cost-effectiveness results of zanubrutinib. Therefore, the cost-effectiveness analysis results from Cohort 1 and the current pooled analysis can be considered conservative. |

**Table 3: MYD88$^{MUT}$ vs. MYD88$^{WT}$ outcomes - ASPEN trial**

| | n | CR | VGPR | PR | Ratio of mut:wt in ASPEN |
|---|---|---|---|---|---|
| **MYD88$^{MUT}$ Cohort 1** | | | | | |
| Zanubrutinib | 102 | 0% | 28% | 49% | 79.7% |
| **MYD88$^{WT}$ Cohort 2** | | | | | |
| Zanubrutinib | 26 | 0% | 27% | 23% | 20.3% |

Abbreviations: CR, complete response; PFS, progression-free survival; PR, partial response; VGPR, very good partial response.

**Figure 1: PFS by mutation type - Study 118E**




Abbreviations: PFS, progression-free survival. Source: Treon et al. 2021[8]

| | | |
|---|---|---|
| **Key issue 12:**<br>Assumption of lifelong treatment effectiveness. | Yes | There is evidence to suggest that the treatment effectiveness of zanubrutinib will persist whilst receiving treatment, and there is no evidence to suggest any waning of treatment effect.<br><br>**Zanubrutinib clinical data demonstrates that treatment effect will persist whilst on treatment**<br>The clinical effectiveness of zanubrutinib relative to BR and DRC is derived from an indirect treatment comparison.<br>In order to predict outcomes over the model time horizon, KM data (matched adjusted for zanubrutinib) were extrapolated over 30 years using standard parametric survival techniques. When selecting the base-case survival models, consideration was given to visual fit, clinically plausibility, statistical fit and the appropriateness of the proportional hazards assumption.<br>Based on an assessment, it can be concluded that the proportional hazard assumption cannot be rejected for the endpoint OS between zanubrutinib and DRC (Schoenfeld residual test p-value=[redacted]), and zanubrutinib and BR (Schoenfeld residual test p-value=[redacted]) (see Appendix D for further details).<br>Therefore, it is appropriate to assume that the treatment effect is proportional over time, which supports the conclusion that the treatment effect of zanubrutinib will persist whilst on treatment.<br>**The ERG's decision to implement a treatment waning assumption is arbitrary and not evidence based.**<br>The Company strongly disagree with the ERG's decision to implement a 5-year treatment waning assumption within the economic model.<br>The visual impact of the zanubrutinib five-year treatment waning assumption on the rate of progression and death is presented in Appendix E, Figure 7 to Figure 10. The four graphs presented show that the five-year treatment waning assumption is extremely pessimistic and results in an unrealistic, sudden loss of treatment benefit in |

| | | |
|---|---|---|
| | | which the rate of progression and death for zanubrutinib is assumed to be equivalent to that of BR or DRC.<br><br>The sudden change in survival trajectory is further demonstrated by the kinked survival analysis curves shown in Figure 11 to Figure 14 in Appendix E. This assumption is highly unrealistic given that treatment with BTKi has resulted in long-term OS benefit in WM (see response to Key Issue 4) and in similar blood cancers.[31] Indeed, no such "kink" has ever been observed before in real-world settings. Furthermore, feedback obtained from a UK clinical expert deemed this assumption clinically unrealistic.<br><br>The ERG refers to NICE appraisal TA627 (lenalidomide plus rituximab in follicular lymphoma) to justify why treatment waning should be applied in the base case.[18] NICE TA627 implemented treatment waning after 5 years as this was consistent with previous NICE submissions in the same disease area (TA472 [obinutuzumab plus bendamustine] and TA137 [rituximab]);[18] treatment waning in TA472 and TA137 was 5.5-years and 5-years, respectively.[18,32,33]<br><br>In the TA627 clarification questions, the ERG acknowledged the differences between TA627, TA472, and TA137, in particular the different populations, and asked for the submitting Company to justify why a 5-year treatment waning was appropriate. Neither TA472 nor TA137 presented evidence to support the treatment waning assumption and TA627 referenced past precedent yet highlighted the limited evidence available.<br><br>TA672, TA472 and TA137 consider patients with follicular lymphoma rather than WM. Furthermore, they all evaluate types of chemotherapies as opposed to highly selective treatment agents, such as a BTKi.<br><br>The most recent and relevant NICE appraisal to the population within this appraisal is TA491, ibrutinib for the treatment of patients with R/R WM. Within TA491, treatment waning was not applied in the Company's base case and they received no criticism from the ERG.<br><br>Furthermore, treatment waning was not applied, nor criticised by the ERG, in **six previous BTKi CLL and mantle cell lymphoma (MCL) appraisals**, see Table 9 in the Appendix E for further details. |

| | | |
|---|---|---|
| | | In TA429 for CLL, the Company justify the exclusion of treatment waning as median PFS had not been reached after a median of 30 months and 3 years in the key clinical trials informing the submission (RESONATE and 1102/1102).[14] The Company argued that limiting ibrutinib's benefit to only 6 years was unreasonable, which the Committee did not refute. Median PFS has not been reached in ASPEN after a median follow-up of 33 months therefore a 5-year treatment waning should not be implemented in this submission.[34]<br><br>The ERG's decision to implement an arbitrary 5-year treatment waning within their base case is not evidence based, and instead relies on past appraisals in different populations which consider less efficacious treatment options.<br><br>Given the clinical trial data for zanubrutinib, the past precedent of more relevant BTKi NICE appraisals, the feedback obtained from a UK clinical expert, and the long-term survival benefit observed in ibrutinib trials in WM and CLL, the evidence clearly demonstrates no evidence of treatment waning, and that the treatment effectiveness of zanubrutinib should persist over time. |
| **Key issue 13:**<br><br>PFS utility higher than general UK population values. | No | The Company acknowledge that the ERG is concerned that the progression-free health-related quality-of-life (HRQoL) estimate based on ASPEN data is higher than the average for the general population.<br><br>In response to ERG clarification questions (question 12), the Company indicated that this may be attributed to differences in clinical and real-world settings as well as differences in geographical locations between ASPEN and UK.<br><br>The Company would like to note that at the recent ACM for ibrutinib (CDF review of TA491) in the treatment of R/R WM, a patient expert stated that he experienced extreme fatigue on a daily basis, which can be severe, disabling, and significantly impair his everyday life. According to Spronk et al, fatigue is not sufficiently addressed by the existing EQ-5D domains, suggesting it does not provide an accurate assessment of HRQoL losses.[35]<br><br>Since the EQ-5D-5L instrument was used to measure health utilities in the ASPEN trial, the true HRQoL of a patient with WM is uncertain. Indeed, a similar issue was raised within appraisal TA689 (CLL), in which the PFS trial EQ-5D utility value was higher than |

| | | |
|---|---|---|
| | | that of the general population, highlighting the potential challenges in capturing the true HRQoL of patients with blood cancer.[15]<br><br>The equivalent age-gender matched utility value is 0.7891, only slightly lower than EQ-5D-5L PFS utility value from the ASPEN trial (0.7910). When considering the impact of adverse events the utility value applied in the PFS health state is lower than the age-gender matched value (0.7886 vs. 0.7891 for zanubrutinib).<br><br>Furthermore, the model one-way sensitivity analysis (OWSA) highlights that the results are not particularly sensitive to this parameter. Given that the model is not sensitive to the PFS utility value adopted and that when you consider adverse events the value is not above the general population utility the Company considers that this issue does not warrant further concern. |
| **Key issue 14:**<br><br>The value and standard error implemented for post-progression utility is not evidence-based. | No | The progressed disease (PD) health state utility value (HSUV) was modelled to align with existing literature, identified in a review of relevant NICE appraisals comparable to the decision problem at hand.<br><br>Due to a lack of EQ-5D-5L data collection following progression in the ASPEN trial, it was not possible to calculate a PD HSUV and hence the Company had to rely on external data sources.<br><br>To achieve the PD HSUV utility a decrement of 0.1 was applied to the PFS utility value upon progression. This is equivalent to a 12.6% decrease in quality-of-life.<br><br>As detailed in Table 10 (Appendix G), which contains several existing NICE appraisals across the BTKi treatment class in similar blood cancer populations (R/R WM, CLL and MCL), HSUV values ranged between 0.665 and 0.763.[14,36,37] The base-case PD HSUV applied within the model (0.691) is of a similar magnitude to the values applied within the past NICE appraisals.<br><br>Furthermore, the percentage decrease in quality-of-life due to progression (12.6%) is aligned with that adopted in TA491 (12.8%), TA502 (12.8%) and TA429 (11.4%).<br><br>The ERG's preferred base-case value of 0.611 represents a 22.4% decrease in quality-of-life due to progression, which is notably almost two times greater than that modelled and accepted by NICE in previous BTKi technology appraisals. |

| | | |
|---|---|---|
| | | Finally, variation of this utility value in the OWSA highlights that is not a key driver of the model. However, to alleviate the concerns raised by the ERG on this issue the Company have included the ERG's preferred assumption for the PD utility value within their revised base case. |
| **Key issue 15:**<br>Large discrepancy between the deterministic incremental cost-effectiveness ratio (ICER) and the probabilistic ICER. | No | The observed discrepancy between the deterministic and probabilistic results was driven by large variation of the survival curves across treatment arms. The programming of the probabilistic survival analysis coefficients has been reviewed and updated. Please see section "Summary of changes to the company's cost-effectiveness estimate(s)" for further details. As a result of these model updates the probabilistic results (for n=5,000 simulations) are now line with the deterministic results (Appendix A) for the Company's revised base-case. |
| **Key issue 16:**<br>Treatment effectiveness being analysed for the different comparisons separately. | Yes/No | The Company acknowledges the concern from NICE and the ERG on the presentation of pairwise comparisons for zanubrutinib versus BR and versus DRC.<br><br>As discussed in the Technical Engagement clarification meeting (28th January 2022) it could be expected, given that treatment choice is often driven by patient and disease characteristics,[3] that patients eligible for BR may be different to the patients eligible for DRC. The presentation of pairwise comparisons was also driven by the need to adopt pairwise indirect treatment comparisons by comparator due to the lack of a connected network across all treatments.<br><br>Nonetheless, the Company acknowledge the importance of comparing zanubrutinib with standard of care in a single comparison, and propose a comparison which is aligned with the methodology adopted in TA491, with the exception that comparators not used in routine clinical practice (as previously specified – see Key Issue 1) are not included within the definition of standard of care.[5]<br><br>In light of this, the cost-effectiveness results of zanubrutinib versus standard of care (consisting of 49% BR and 51% DRC based on the UK 2021 Rory Morrison Registry report) have been weighted to produce an overall ICER of zanubrutinib versus standard of care. For the Company's revised base case (see Appendix A), when compared to |

| | | |
|---|---|---|
| | | standard of care, zanubrutinib is associated with £[redacted] additional costs and [redacted] additional QALYs, corresponding to an ICER of £20,054 per QALY gained. |

# Additional issues

**All:** Please use the table below to respond to additional issues in the ERG report that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this appraisal (for example, at the clarification stage).

**Table 3 Additional issues from the ERG report**

| Issue from the ERG report | Relevant section(s) and/or page(s) | Does this response contain new evidence, data or analyses? | Response |
|---|---|---|---|
| Additional issue 1: Insert additional issue | Please indicate the section(s) of the ERG report that discuss this issue | Yes/No | Please include your response, including any new evidence, data or analyses, and a description of why you think this is an important issue for decision making |
| Additional issue 2: Insert additional issue | Please indicate the section(s) of the ERG report that discuss this issue | Yes/No | Please include your response, including any new evidence, data or analyses, and a description of why you think this is an important issue for decision making |
| Additional issue N: Insert additional issue | | | [INSERT / DELETE ROWS AS REQUIRED] |

# Summary of changes to the company's cost-effectiveness estimate(s)

**Company only:** If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

**Table 4 Changes to the company's cost-effectiveness estimate**

| Key issue(s) in the ERG report that the change relates to | Company's base case before technical engagement | Change(s) made in response to technical engagement | Impact on the company's base-case incremental cost-effectiveness ratio (ICER) |
|---|---|---|---|
| Key issue 1 | Pairwise comparisons of zanubrutinib versus BR, and versus DRC | Estimation of a weighted ICER to reflect standard of care in line with Ibrutinib NICE appraisal TA491 | Company submitted base case following ERG clarification questions:<br>ICER vs. BR: [redacted]<br>ICER vs. DRC: [redacted]<br>Revised Company submitted base case<br>ICER vs. Standard of Care: £20,054 |
| Key issue 5 | MAIC analyses for vs. BR and DRC for PFS and OS endpoints | Addition of STC analyses vs. BR and DRC for PFS and OS endpoints. STC HR applied to baseline BR/DRC curves to generate zanubrutinib curves. | Company submitted base case following ERG clarification questions:<br>ICER vs. BR: [redacted]<br>ICER vs. DRC: [redacted]<br>Revised Company submitted base case<br>ICER vs. Standard of Care: £20,054 |
| Key issue 11 | Standard parametric modelling for PFS and OS endpoints | Addition of flexible survival analyses for PFS and OS for | Company submitted base case following ERG clarification questions: |

| | | | |
|---|---|---|---|
| | | zanubrutinib (matched BR), BR, zanubrutinib (matched DRC) and DRC treatment arms. | ICER vs. BR: [redacted]<br>ICER vs. DRC: [redacted]<br>Revised Company submitted base case<br>ICER vs. Standard of Care: £20,054 |
| Key issue 15 | Probabilistic coefficients for standard parametric survival analyses determined via hard coded Choleskey decomposition matrix and inverse normal of the rand() function. | Probabilistic coefficients for standard parametric survival analyses programmed through distribution covariance matrix, rand() function and Choleskey decomposition VBA function.<br>Restriction on dependent Gamma OS treatment effect covariate from varying to values greater than 0.<br>Variation of the Weibull OS independent BR curve programmed using the "Norm.Inv" function to prevent extreme variation of the scale and shape parameters which lead to almost an vertical OS curves for BR. | N/A – Impacts probabilistic results only |
| Company's base case following technical engagement (or revised base case) | Incremental QALYs vs Standard of Care: [redacted] | Incremental costs vs. Standard of Care: £[redacted] | ICER vs. Standard of Care: £20,054 |

Abbreviations: BR, bendamustine plus rituximab; DRC, dexamethasone, rituximab plus cyclophosphamide; ICER, incremental cost-effectiveness ratio; OS, overall survival; PFS, Progression-free survival; QALY, Quality adjusted life-year; QoL, Quality of life; STC, simulated treatment comparison

# Appendix A. Revised cost-effectiveness analysis

**Table 4: Revised Company base-case settings**

| # | Model setting | Included within revised Company base case | Rationale |
|---|---|---|---|
| 1 | BR given for 6 treatment cycles (4 model cycles), DRC given for 6 model cycles | Y | - |
| 2 | Ibrutinib excluded as direct comparator and subsequent treatment | N | See response to Key Issue 8 |
| 3 | Assuming similar relative dose intensity rates for Zanubrutinib, BR and DRC | Y | - |
| 4 | Implementation of treatment waning 5 years | N | See response to Key Issue 12 |
| 5 | Including AEs of Grade ≥3 which occurred in ≥1% of the population and implementation of hypotension | Y | - |
| 6 | Include Age-adjusted utilities | Y | - |
| 7 | Adjust post-progression utility decrement from -.1 to -.18 | Y | - |
| 8 | Flexible Odds k=1 model for OS extrapolation of zanubrutinib (matched to DRC) and DRC | Y | See response to Key Issue 10 |
| 9 | Application STC HRs to generated PFS and OS for zanubrutinib | Y | See response to Key Issue 5 |

Abbreviations: AE, adverse events; BR, bendamustine plus rituximab; DRC, dexamethasone, rituximab plus cyclophosphamide

**Table 5: Revised cost-effectiveness results**

| # | Details | Vs. Standard of Care | | |
|---|---|---|---|---|
| | | **Inc. costs** | **Inc. QALYs** | **ICER** |
| **1** | **Revised Company base case** (Implementation of #1, #2, #5-9 from Table 4) | [redacted] | [redacted] | **£20,054** |
| ***Illustrative scenario analyses on ERG-corrected Company base case*** | | | | |
| 2 | ERG-corrected Company base case with reduced PPS for comparator arms (25% reduction) | [redacted] | [redacted] | £43,243 |
| 3 | ERG-corrected Company base case with reduced PPS for comparator arms (50% reduction) | [redacted] | [redacted] | £38,215 |
| 4 | ERG-corrected Company base case with reduced PPS for comparator arms (75% reduction) | [redacted] | [redacted] | £34,235 |

Abbreviations: Inc., incremental; ICER, incremental cost-effectiveness ratio; PPS, post-progression survival; QALYs, quality adjusted life years

## Sensitivity analyses around revised base case

Probabilistic sensitivity analyses results for the Company revised base-case for zanubrutinib versus standard of care are presented in Table 6. Zanubrutinib is associated with [redacted] additional QALYs and £[redacted] additional costs, with a corresponding ICER of £21,023 per QALY gained. The results demonstrate that the analysis is robust to parameter uncertainty with the probabilistic results lying close to the deterministic results (Table 5).

**Table 6. Probabilistic results: zanubrutinib vs. Standard of Care**

| Treatment | Total costs (£) | Total QALYs | Inc. costs (£) | Inc. QALYs | ICER (£) |
|---|---|---|---|---|---|
| SoC | [redacted] | [redacted] | | | |
| Zanubrutinib | [redacted] | [redacted] | [redacted] | [redacted] | 21,023 |

Abbreviations: Inc., incremental; ICER, incremental cost-effectiveness ratio; QALYs, quality adjusted life years; SoC, Standard of Care

# Appendix B. Assessment of comparators included within the NICE scope

**Table 7: Assessment of comparators included within the NICE scope**

| Treatment within NICE scope | 2018 ESMO WM guidelines[3] | 2021 BSH WM guidelines [4] [38] | 2021 Rory Morrison registry report[2] | Conclusion |
|---|---|---|---|---|
| BR | • The 2018 ESMO WM guideline recommends the use of BR across all symptomatic treatment naïve patients (both fit and unit, and those with high or low tumour burden). | • The 2021 BSH WM guideline recommends the use of BR as front-line treatments, noting that *"The two most commonly used first-line regimens are dexamethasone, rituximab and cyclophosphamide (DRC) and rituximab–bendamustine (BR)".*<br>• For R/R disease the guidelines recommend that rituximab-containing regimens are considered. | • 41% of patients (weighted by treatment naïve and R/R, and adjusted to reflect treatments within the NICE scope), were recorded to have received treatment with a bendamustine-based regimen | • BR is recommended as standard of care by both European and British treatment guidelines.<br>• A large proportion of patients in the UK across both first and second-line therapies have been recorded to have received treatment with a bendamustine-based regimen.<br>• BR represents a main stay treatment option for patients with WM in the UK and hence is a relevant comparator for this appraisal. |

| Treatment within NICE scope | 2018 ESMO WM guidelines[3] | 2021 BSH WM guidelines [4] 38 | 2021 Rory Morrison registry report[2] | Conclusion |
|---|---|---|---|---|
| DRC | • The 2018 ESMO WM guideline recommends the use of DRC across all symptomatic treatment naïve patients (both fit and unfit, and those with high or low tumour burden). | • The 2021 BSH WM guideline recommends the use of DRC as front-line treatments, noting that "*The two most commonly used first-line regimens are dexamethasone, rituximab and cyclophosphamide (DRC) and rituximab–bendamustine (BR)".*<br>• For R/R disease the guidelines recommend that rituximab-containing regimens are considered. | • 43% of patients (weighted by treatment naïve and R/R, and adjusted to reflect treatments within the NICE scope), were recorded to have received treatment with DRC. | • DRC is recommended as standard of care by both European and British treatment guidelines.<br>• A large proportion of patients in the UK across both first and second-line therapies have been recorded to have received treatment with DRC.<br>• DRC represents a main stay treatment option for patients with WM in the UK and hence is a relevant comparator for this appraisal. |

| Treatment within NICE scope | 2018 ESMO WM guidelines[3] | 2021 BSH WM guidelines [4] [38] | 2021 Rory Morrison registry report[2] | Conclusion |
|---|---|---|---|---|
| FCR/FR | • FCR is a more intensive chemotherapy which can induce high response rates but with significant toxicity and are therefore not primary options for first-line treatment of WM.<br>• Rituximab with nucleoside analogues (FR, FCR) is an active but also toxic combination and therefore should be used cautiously. | • FCR/FR is not highlighted as a treatment option within the 2021 BSH guidelines. | • 0% of patients recorded to have received FCR or FR as a first-line (2015-2020) or second-line therapy (2017-2020).<br>• *"Over recent years, as we learnt more about the additions to the treatment arsenal (DRC and bendamustine), there have been concurrent* ***concerns about toxicity*** *of existing agents in widespread use (chlorambucil and fludarabine)."* | • FCR/FR is not widely used for the treatment of WM due to high level of toxicity, hence it is not established standard of care in UK clinical practice. Therefore, FCR/FR is not a relevant comparator in this appraisal. |
| Cladribine and rituximab | • Cladribine is not highlighted as a treatment option within the ESMO guidelines. | • Cladribine is not highlighted as a treatment option within the BSH guidelines. | • 0% of patients recorded to have received cladribine plus rituximab as a first-line (2015-2020) or second-line (2017-2020) therapy. | • Cladribine plus rituximab is not recommended as a treatment option in either the ESMO or BSH guidelines.<br>• No use of cladribine plus rituximab as recorded by the UK |

| Treatment within NICE scope | 2018 ESMO WM guidelines[3] | 2021 BSH WM guidelines [4] [38] | 2021 Rory Morrison registry report[2] | Conclusion |
|---|---|---|---|---|
| | | | | Rory Morrison Registry.<br>• WM not included within the licensed indication of cladribine, hence it is not established standard of care in UK clinical practice.[39]<br>• Therefore, cladribine plus rituximab is not a relevant comparator in this appraisal. |
| Autologous stem cell transplantation | • *"The role of allogeneic stem cell transplantation (alloSCT) is limited outside clinical trials and should be considered only in highly selected young patients with aggressive disease, who have failed or are resistant to BTK inhibitors."* | • *"The lack of prospective comparative trials makes it challenging to provide high-quality recommendations on the role of stem cell transplant (SCT) in WM. For WM patients who are potential autologous SCT candidates it is important to avoid the use of stem-cell-toxic therapeutic drugs for first-line* | • 0.4% of patients received an allograft stem cell transplant, across all recorded first-line treatment options. | • Limited research on the benefit of SCT.<br>• General consensus from guidelines is that other therapies should be considered over SCT due associated risk of death.<br>• It is not established standard of care in UK clinical practice, hence SCT is not a relevant comparator in this appraisal. |

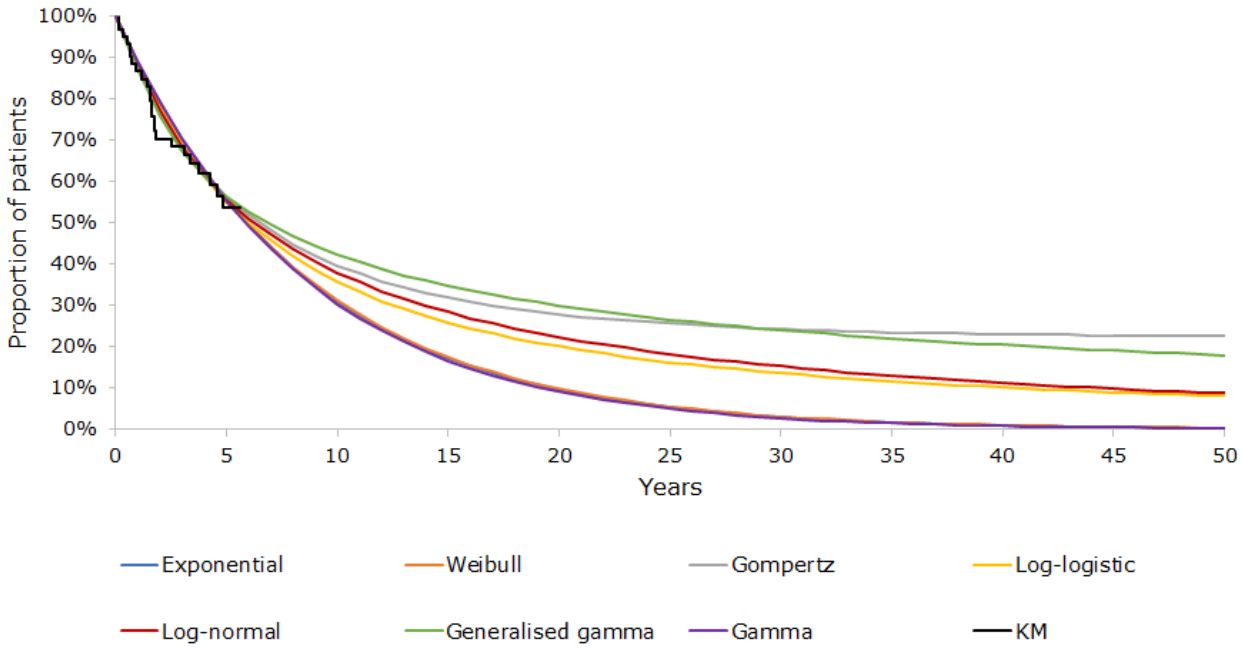
| Treatment within NICE scope | 2018 ESMO WM guidelines[3] | 2021 BSH WM guidelines [4] 38 | 2021 Rory Morrison registry report[2] | Conclusion |
|---|---|---|---|---|
| | | *therapy to reduce risk for stem cell harvest failure"*<br>• *"The place of allogeneic SCT (alloSCT) in the treatment algorithm of WM has become more controversial especially in the era of new agents even for younger patients. There is a high non-relapse mortality (NRM) and the use of allo-SCT is therefore limited to highly selected patients."*<br>• *"Whilst there are case series detailing positive outcomes for autologous SCT (ASCT) for WM as part of first-line therapy, this cannot be recommended outside a clinical trial due to lack of strong evidence"* | | |

| Treatment within NICE scope | 2018 ESMO WM guidelines[3] | 2021 BSH WM guidelines [4] [38] | 2021 Rory Morrison registry report[2] | Conclusion |
|---|---|---|---|---|
| Chlorambucil | • Chlorambucil monotherapy is included within the ESMO treatment algorithm for only unfit treatment naïve patients with a low tumour burden. The guidelines note that it has limited clinical benefit.<br>• Chlorambucil monotherapy is not recommended by ESMO in the treatment algorithm for R/R patients. | • *"Single-agent chlorambucil has a very limited role in contemporary first-line therapy"* | • 5% of patients (weighted by treatment naïve and R/R, and adjusted to reflect treatments within the NICE scope), were recorded to have received treatment with chlorambucil. | • Limited use of chlorambucil within the first-line setting and is not recommended for R/R patients.<br>• It is not established standard of care in UK clinical practice, hence chlorambucil is not a relevant comparator in this appraisal. |
| Rituximab monotherapy | • "*Rituximab has low toxicity but is associated with modest response rates as a monotherapy"*<br>• Rituximab monotherapy is not included within the treatment pathway for treatment naïve or R/R patients | • *"Rituximab monotherapy is generally well tolerated but associated with modest response rates and relatively short PFS"* | • 11% of patients (weighted by treatment naïve and R/R, and adjusted to reflect treatments within the NICE scope), were recorded to have received treatment with rituximab monotherapy | • Limited use of rituximab within the first-line setting and is not recommended for R/R patients.<br>• It is not established standard of care in UK clinical practice; hence rituximab is not a relevant comparator in this appraisal. |

Abbreviations: NICE, National institute for health and care excellence; ESMO, European society for medical oncology; WM, Waldenström's Macroglobulinaemia; BSH, British society for haematology; FCR, Fludarabine, cyclophosphamide and rituximab; FR, Fludarabine and cyclophosphamide; UK, United Kingdom; SCT, Stem cell transplantation; ASCT, autologous stem cell transplantation; AlloSCT, allogeneic stem cell transplantation; R/R, relapsed/refractory; NRM, Non-relapse mortality

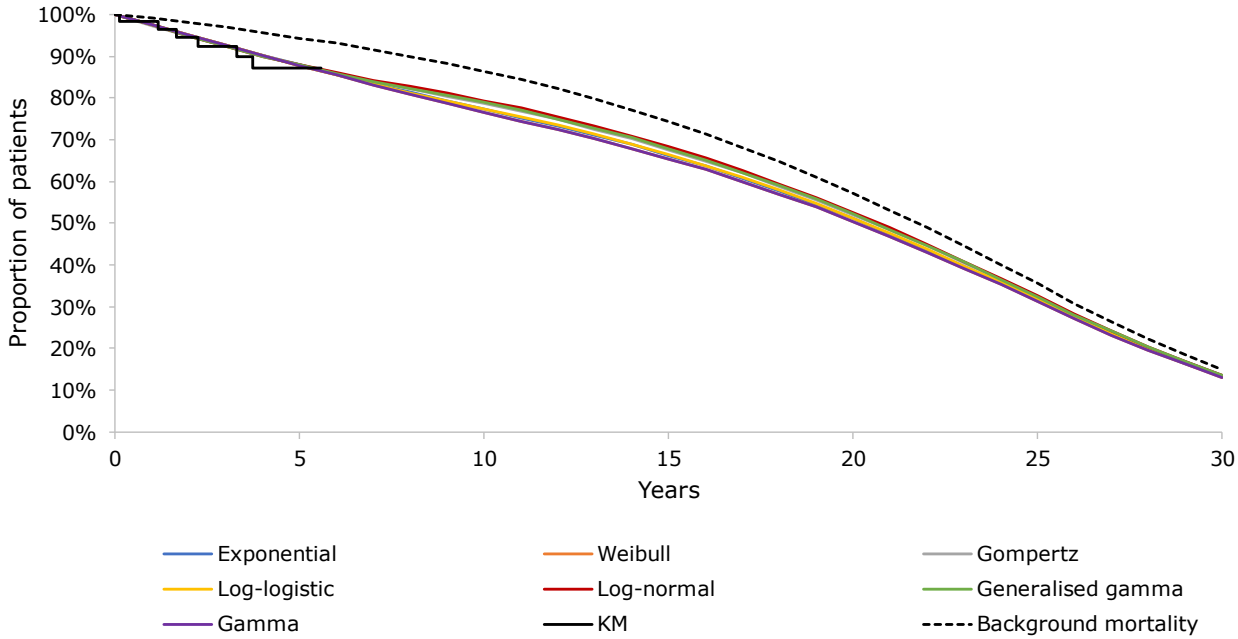
# Appendix C. Long-term ibrutinib survival data

**Figure 2: Study 118E ibrutinib PFS in relapsed and refractory WM**



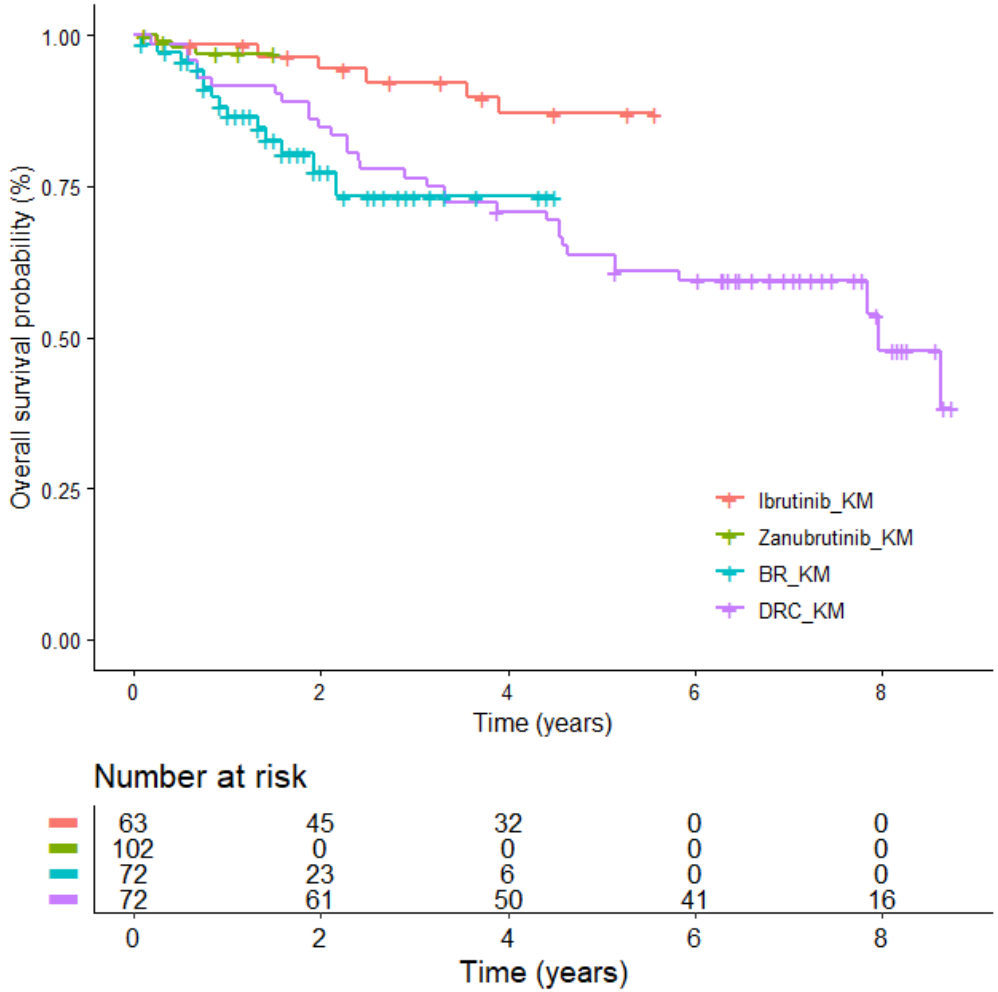

Abbreviations: KM, Kaplan-Meier; PFS, progression-free survival; WM, Waldenström's macroglobulinaemia
Source. Digitised KM data from Treon et al. 2021[8]

**Figure 3: Study 118E ibrutinib OS in relapsed and refractory WM – restricted by background mortality**




Abbreviations: KM, Kaplan-Meier; OS, overall survival; WM, Waldenström's macroglobulinaemia
Source. Digitised KM data from Treon et al. 2021[8]

**Figure 4. OS KM comparison**




Abbreviations: OS, overall survival; KM, Kaplan-Meier
Source. Ibrutinib - digitised KM data from Treon et al. 2021[8]; BR – digitised KM data from Tedeschi 2015[11]; DRC – digitise KM data from Kastritis et al. 2015[12]; zanubrutinib - ASPEN ITT trial[9]

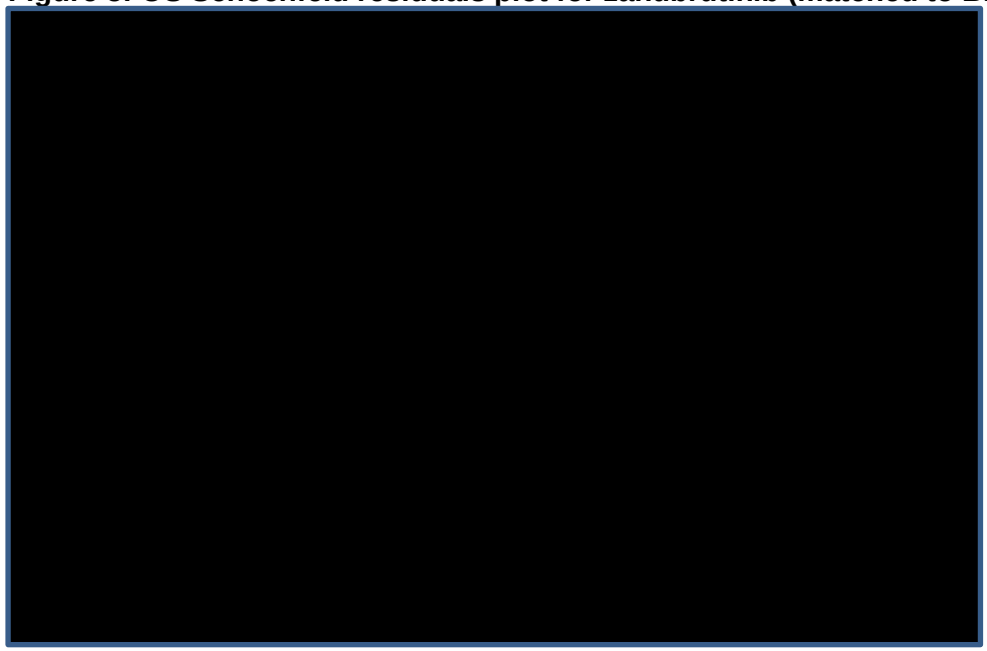
**Table 8: Mean survival for extrapolated Study 118E ibrutinib OS data**

| Distribution | Exponential | Weibull | Gompertz | Log-logistic | Log-normal | Generalised Gamma | Gamma |
|---|---|---|---|---|---|---|---|
| Mean OS | 18.54 | 18.41 | 18.77 | 18.56 | 18.88 | 18.80 | 18.40 |

Abbreviations: OS, overall survival. Notes. Undiscounted mean OS over 30 years, capped by all-cause mortality

# Appendix D. Proportional hazards assessment

**Figure 5: OS Schoenfeld residuals plot for zanubrutinib (matched to BR) and BR**



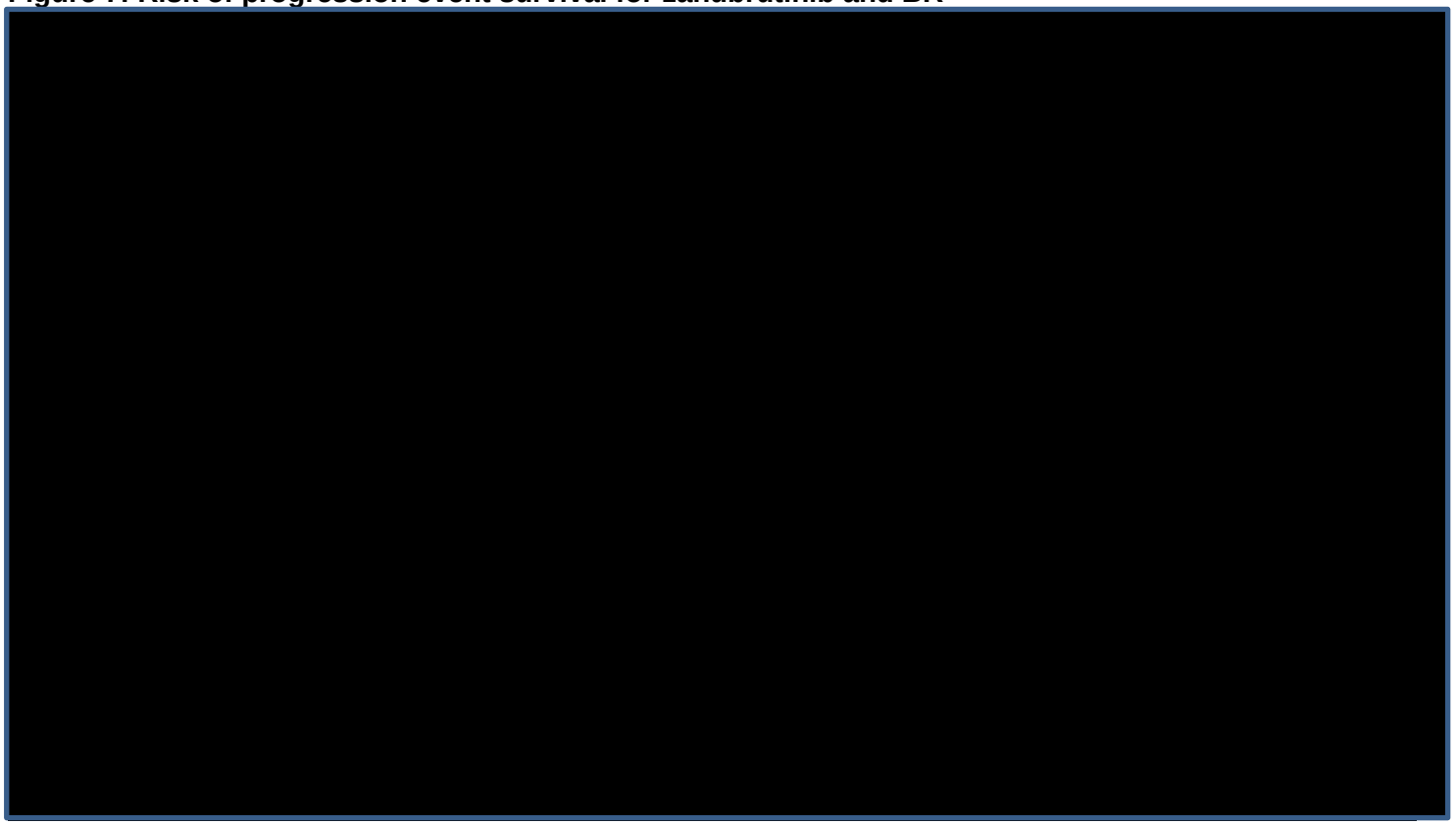
Abbreviations: BR, bendamustine plus rituximab; OS, overall survival; tx, treatment.

**Figure 6: OS Schoenfeld residuals plot for zanubrutinib (matched to DRC) and DRC**

Abbreviations: BR, bendamustine plus rituximab; OS, overall survival; tx, treatment.

# Appendix E. Lifelong treatment effectiveness

**Figure 7: Risk of progression event survival for zanubrutinib and BR**



Abbreviations: BR, bendamustine + rituximab; ERG, Evidence review group; PFS, Progression-free survival.

**Figure 8: Risk of death event for zanubrutinib and BR**

Abbreviations: BR, bendamustine + rituximab; ERG, Evidence review group; OS, Overall survival

**Figure 9: Risk of progression event for zanubrutinib and DRC**

Abbreviations: DRC, Dexamethasone + rituximab + cyclophosphamide; ERG, Evidence review group; PFS, Progression-free survival.

**Figure 10: Risk of death event for zanubrutinib and DRC**

Abbreviations: DRC, Dexamethasone + rituximab + cyclophosphamide; ERG, Evidence review group; OS, Overall survival.

**Figure 11: PFS curve for zanubrutinib and BR with ERG five-year treatment waning assumption applied**

Abbreviations: BR, bendamustine + rituximab; ERG, Evidence Review Group; PFS, Progression-free survival.

**Figure 12: OS curve for zanubrutinib and BR with ERG five-year treatment waning assumption applied**

Abbreviations: BR, bendamustine + rituximab; ERG, Evidence Review Group; OS, Overall survival.

**Figure 13: PFS curve for zanubrutinib and DRC with ERG five-year treatment waning assumption applied**

Abbreviations: DRC, Dexamethasone + rituximab + cyclophosphamide; PFS, Progression-free survival.

**Figure 14: OS curve for zanubrutinib and DRC with ERG five-year treatment waning assumption applied**

Abbreviations: DRC, Dexamethasone + rituximab + cyclophosphamide; ERG, Evidence Review Group; OS, Overall survival.

**Table 9: Past precedent for treatment waning in previous NICE submissions in leukaemia**

| Data source | Population | Base case | Scenarios | ERG comments |
|---|---|---|---|---|
| TA491[14] | R/R WM | Not applied | - | • Proportional hazards assumption holds for PFS between ibrutinib and European chart review cohort.<br>• Consequence of this assumption is to assume that the treatment effect is maintained for the lifetime of patients.<br>• No criticism from the ERG regarding treatment waning |
| TA429[36] | R/R CLL | Not applied | 6-year and 7-year treatment waning | • No criticism from the ERG |
| TA689[15] | R/R CLL | Not applied | - | • No criticism from the ERG |
| TA502[37] | R/R MCL | Not applied | - | • No criticism from the ERG |
| TA677[20] | R/R MCL | Not applied | - | • The company justified exclusion of treatment waning as "Not appropriate as CAR T-cell therapies are given as a single dose".<br>• References TA502 where treatment waning was not applied.<br>• No criticism from the ERG |
| TA663[19] | CLL | Not applied | - | • No criticism from the ERG |
| TA627* [18] | FL | 5-year treatment waning | 3-year and 10-year treatment waning. | • ERG required company to provide justification as to why the 5-year treatment waning was considered appropriate.<br>• Justifications of 5-year treatment based on previous FL appraisals (TA137 and TA472)<br>• The ERG considers the company's choice of time point to be rather arbitrary and a shorter or longer duration of treatment effectiveness may be equally likely. |
| TA137[33] | R/R FL | 5-year treatment waning | 2-year treatment waning | • Full ERG report unavailable but FAD mentions that the committee considered a 5-year treatment benefit reasonable.<br>• Sensitivity analyses should include the effect of varying the time horizons and varying the assumed duration of treatment benefit, within the range 1500 days to 30 years. |

| Data source | Population | Base case | Scenarios | ERG comments |
|---|---|---|---|---|
| TA472[19]** | R/R FL | 5.5-year treatment waning | 7-years<br>4-years and 25-years | • The committee considered that availability of more mature OS data from the GADOLIN trial was likely to resolve uncertainty around treatment effect and may produce more robust cost-effectiveness estimates.<br>• The committee considered it plausible that treatment effect was longer than modelled in the company's base case. |

Abbreviations; CLL, chronic lymphocytic leukaemia; ERG, Evidence Review Group; FAD, Final appraisal determination; FL, follicular lymphoma; MCL, mantle cell lymphoma; OS, overall survival; WM, Waldenström's macroglobulinaemia; TN, treatment naïve; R/R, relapsed and refractory
Note: *TA the ERG references in our appraisal; **TA that T627 references

# Appendix F. Hazard function and flexible survival analyses

**Figure 15. Zanubrutinib (matched to BR) observed and extrapolated (standard parametric models) OS hazard function**

Abbreviations; BR, bendamustine plus rituximab; OS, overall survival

**Figure 16. BR observed and extrapolated (standard parametric models) OS hazard function**

Abbreviations; BR, bendamustine plus rituximab; OS, overall survival

**Figure 17. Zanubrutinib (matched to DRC) observed and extrapolated (standard parametric models) OS hazard function**

Abbreviations; DRC, Dexamethasone + rituximab + cyclophosphamide; OS, overall survival

**Figure 18. DRC observed and extrapolated (standard parametric models) OS hazard function**

Abbreviations; DRC, Dexamethasone + rituximab + cyclophosphamide; OS, overall survival

**Figure 19. Hazard rate for zanubrutinib (matched to DRC) and DRC OS - Hazards k=1 model**

Abbreviations; DRC, Dexamethasone + rituximab + cyclophosphamide; OS, overall survival

**Figure 20. Hazard rate for zanubrutinib (matched to DRC) and DRC OS - Odds k=1 model**

Abbreviations; DRC, Dexamethasone + rituximab + cyclophosphamide; OS, overall survival

# Appendix G. Progressed disease utility

**Table 10: Progressed disease utility value comparison to previous HTA submissions**

| Source | Population | Treatment arm | PD HSUV | Instrument | Details |
|---|---|---|---|---|---|
| TA491 (ibrutinib appraisal) [14] | R/R WM | Non-treatment specific | 0.665 | EQ-5D-5L | Utility decrement of 0.098 (12.8% applied), following Beusterien et al. (2010). |
| TA429 (ibrutinib appraisal)[36] | R/R CLL | Non-treatment specific | 0.763 | Standard gamble | Utility decrement of 0.098 (11.4%) |
| TA502 (ibrutinib appraisal)[37] | R/R MCL | Non-treatment specific | 0.680 | EQ-5D | Utility decrement of 0.1 (12.8%) |

Abbreviations: EQ-5D, EuroQoL-Five Dimensions; EQ-5D-5L, EuroQoL-Five Dimensions-Five Levels; HTA, Health technology assessment; PD, Progressed disease; HSUV, Health state utility value; R/R, Relapsed/refractory; WM, Waldenström's macroglobulinemia; CLL, Chronic lymphocytic leukaemia; MCL, Mantle cell lymphoma.

# Appendix H. Simulated treatment comparison (STC)

A simulated treatment comparison was performed to indirectly compare zanubrutinib with BR and with DRC separately. As outcomes are time-to-event data the following method was used:

**Step 1**: A Cox regression model was fitted to the zanubrutinib ASPEN IPD to determine the relationship between the included covariates and survival (PFS and OS). The following effect modifiers and prognostic variables were used as covariates in the models:

- *Zanubrutinib vs BR:* For both PFS and OS outcomes, Age <=72 years,<=2 prior lines of therapy, IgM <=38.15 g/L, IPSS-High risk and Splenomegaly/Adenopathy were included as covariates.
- *Zanubrutinib vs DRC:* For PFS, Age >69 years, Platelet <100x$10^9$/L, Hemoglobin <100 g/L, Lymphadenopathy and Splenomegaly were included as covariates. For OS, Platelet <100x109/L, Hemoglobin <100 g/L, Lymphadenopathy were included as covariates.
    - *For consistency covariates selected for the MAIC were included within the STC where possible. Age categories were collapsed to a binary covariate for the BR comparison to ensure at least one event was observed for patients within the categories. For the DRC OS comparison covariates for age and splenomegaly/adenopathy were removed from the model as no death event was observed in patients with age <=69 or patients with splenomegaly.*

**Step 2:** The time points used to predict the survival probability were determined based on where there was at least one event in the BR/DRC treatment arms.

**Step 3:** The survival probability for the time points determined in Step 2 is predicted for the comparator population (BR or DRC) when treating with zanubrutinib.

**Step 4:** Population adjusted treatment effects of zanubrutinib on the BR/DRC trial population were obtained by fitting a cox regression model to the reported KM curve from BR/DRC study and predicted zanubrutinib KM curve for BR/DRC study from Step 3.

**Figure 21: OS zanubrutinib vs BR from STC**

Abbreviations: BR, bendamustine-rituximab; OS, overall survival; STC, simulated treated comparison

**Figure 22: PFS zanubrutinib vs BR from STC**

Abbreviations: BR, bendamustine-rituximab; PFS, progression-free survival; STC, simulated treated comparison

**Figure 23: OS zanubrutinib vs DRC from STC**

Abbreviations: DRC, dexamethasone, rituximab and cyclophosphamide; OS, overall survival; STC, simulated treated comparison

**Figure 24: PFS zanubrutinib vs DRC from STC**

Abbreviations: DRC, dexamethasone, rituximab and cyclophosphamide; PFS, progression-free survival; STC, simulated treated comparison

**Table 11: Zanubrutinib vs BR - STC results**

| Zanubrutinib vs BR | HR (95% CI) | p-value |
|---|---|---|
| PFS | [redacted] | [redacted] |
| OS | [redacted] | [redacted] |

Abbreviations: BR, bendamustine-rituximab; CI, confidence interval; HR, hazard ratio; OS, overall survival; PFS, progression-free survival; STC, simulated treated comparison

**Table 12: Zanubrutinib vs DRC - STC results**

| Zanubrutinib vs DRC | HR (95% CI) | p-value |
|---|---|---|
| PFS | [redacted] | [redacted] |
| OS | [redacted] | [redacted] |

Abbreviations: DRC, dexamethasone, rituximab and cyclophosphamide; CI, confidence interval; HR, hazard ratio; OS, overall survival; PFS, progression-free survival; STC, simulated treated comparison

## Appendix I. References


1. Brown, R. *et al.* First UK Waldenström's Macroglobulinaemia Registry Report. 52 (2018).
2. The Rory Morrison Registry. *Second UK Waldenström's Macroglobulinaemia Registry Report 2021*. https://wmuk.org.uk/resource/the-rory-morrison-wmuk-registry-report-2021/ (2021).
3. ESMO guidelines commitee *et al.* Waldenstrom's macroglobulinaemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology* **29**, 41–50 (2018).
4. Pratt, G. *et al.* Guidelines on the diagnosis and management of Waldenström macroglobulinaemia—A British Society for Haematology guideline. *British journal of haematology* (2021).
5. Michael Holmes and Steve Edwards. Ibrutinib for treating Waldenström's macroglobulinaemia CDF review TA491: Lead team presentation. (12/07//21).
6. European Medicines Agency. *Summary of product characteristics: BRUKINSA (zanubrutinib)*. https://www.ema.europa.eu/en/documents/product-information/brukinsa-epar-product-information_en.pdf (2021).
7. Castillo, J. J. *et al.* Long-term follow-up of ibrutinib monotherapy in treatment-naive patients with Waldenstrom macroglobulinemia. *Leukemia* **36**, 532–539 (2022).

8. Treon, S. P. *et al.* Long-Term Follow-Up of Ibrutinib Monotherapy in Symptomatic, Previously Treated Patients With Waldenström Macroglobulinemia. *J Clin Oncol* **39**, 565–575 (2021).
9. Tam, C. S. *et al.* A randomized phase 3 trial of zanubrutinib vs ibrutinib in symptomatic Waldenström macroglobulinemia: the ASPEN study. *Blood* **136**, 2038–2050 (2020).
10. Tai, T.-A., Latimer, N. R., Benedict, Á., Kiss, Z. & Nikolaou, A. Prevalence of Immature Survival Data for Anti-Cancer Drugs Presented to the National Institute for Health and Care Excellence and Impact on Decision Making. *Value in Health* **24**, 505–512 (2021).
11. Tedeschi, A. *et al.* Bendamustine and rituximab combination is safe and effective as salvage regimen in Waldenström macroglobulinemia. *Leuk Lymphoma* **56**, 2637–2642 (2015).
12. Kastritis, E. *et al.* Dexamethasone, rituximab, and cyclophosphamide as primary treatment of Waldenström macroglobulinemia: final analysis of a phase 2 study. *Blood* **126**, 1392–1394 (2015).
13. National Institute for Health and Care Excellence. Project information | Ibrutinib for treating Waldenstrom's macroglobulinaemia (CDF Review of TA491) [ID3778] | Guidance | NICE. https://www.nice.org.uk/guidance/indevelopment/gid-ta10827.
14. National Institute for Health and Care Excellence. *Single Technology Appraisal: Ibrutinib for treating Waldenström's macroglobulinaemia [ID884]*. https://www.nice.org.uk/guidance/ta491/documents/committee-papers (2016).

15. National Institute for Health and Care Excellence. Overview | Acalabrutinib for treating chronic lymphocytic leukaemia | Guidance | NICE. https://www.nice.org.uk/guidance/ta689.
16. Phillippo, D.M., Ades, A.E., Dias, S., Palmer, S., Abrams, K.R., Welton, N.J. NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submission to NICE. (201AD).
17. Assessing the performance of population adjustment methods for anchored indirect comparisons: A simulation study - Phillippo - 2020 - Statistics in Medicine - Wiley Online Library. https://onlinelibrary.wiley.com/doi/full/10.1002/sim.8759.
18. National Institute for Health and Care Excellence. *Single Technology Appraisal: Lenalidomide for treated follicular lymphoma and marginal zone lymphoma [ID1374] TA627*. https://www.nice.org.uk/guidance/ta627/evidence/committee-papers-pdf-8708812813 (2020).
19. National Institute for Health and Care Excellence. *Single Technology Appraisal: Venetoclax with obinutuzumab for untreated chronic lymphocytic leukaemia [ID1402]*. https://www.nice.org.uk/guidance/ta663/evidence/committee-papers-pdf-8952456925 (2020).
20. National Institute for Health and Care Excellence. *Single Technology Appraisal: Autologous anti-CD19-transduced CD3+ cells for treating relapsed or refractory mantle cell lymphoma [ID1313]*. https://www.nice.org.uk/guidance/ta677/evidence/committee-papers-pdf-9016513021 (2021).

21. Floquet, A. *et al.* Progression-free survival by local investigator versus independent central review: Comparative analysis of the AGO-OVAR16 Trial. *Gynecologic Oncology* **136**, 37–42 (2015).
22. National Institute for Health and Care Excellence. *Single Technology Appraisal: Tepotinib for treating advanced non-smallcell lung cancer with MET gene alterations [ID3761]]*. https://www.nice.org.uk/guidance/gid-ta10630/documents/committee-papers (2022).
23. National Institute for Health and Care Excellence. *Single Technology Appraisal: Nivolumab with ipilimumab for untreated unresectable malignant pleural mesothelioma [ID1609]*. https://www.nice.org.uk/guidance/gid-ta10498/documents/committee-papers (2021).
24. National Institute for Health and Care Excellence. *Appraisal consultation document: Tepotinib for treating advanced non-small-cell lung cancer with MET gene alterations*. https://www.nice.org.uk/guidance/gid-ta10630/documents/129 (2022).
25. National Institute for Health and Care Excellence. *Appraisal consultation document: Nivolumab with ipilimumab for untreated unresectable malignant pleural mesothelioma*. https://www.nice.org.uk/guidance/gid-ta10498/documents/129 (2021).
26. Woods, B., Sideris, E., Palmer, S., Latimer, N. & Soares, M. NICE DSU TECHNICAL SUPPORT DOCUMENT 19: Partitioned Survival Analysis for decision modelling in health care: a critical review. 72.

27. Sheth, M. & Ko, J. Exploring the relationship between Overall Survival (OS), Progression Free Survival (PFS) and Objective Response Rate (ORR) in patients with advanced melanoma. *Cancer Treatment and Research Communications* **26**, 100272 (2021).
28. Cooper, K., Tappenden, P., Cantrell, A. & Ennis, K. A systematic review of meta-analyses assessing the validity of tumour response endpoints as surrogates for progression-free or overall survival in cancer. *British Journal of Cancer* **123**, 1686–1696 (2020).
29. National Institute for Health and Care Excellence. *Appraisal consultation document: Ibrutinib for treating Waldenstrom's macroglobulinaemia*. https://www.nice.org.uk/guidance/ta491/documents/appraisal-consultation-document (2016).
30. Royston, P. & Parmar, M. K. B. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Stat Med* **21**, 2175–2197 (2002).
31. Munir, T. *et al.* Final analysis from RESONATE: Up to six years of follow-up on ibrutinib in patients with previously treated chronic lymphocytic leukemia or small lymphocytic lymphoma. *Am J Hematol* **94**, 1353–1363 (2019).
32. National Institute for Health and Care Excellence. *Obinutuzumab with bendamustine for treating follicular lymphoma refractory to rituximab (CDF Review of TA472) [ID1583]*. https://www.nice.org.uk/guidance/ta629/documents/committee-papers (2020).

33. National Institute for Health and Care Excellence. Rituximab for the treatment of relapsed or refractory stage III or IV follicular non-Hodgkin's lymphoma. *NICE* https://www.nice.org.uk/guidance/ta137/documents/follicular-lymphoma-rituximab-appraisal-consultation-document (2008).
34. An, G. *et al.* A Phase II Trial of the Bruton Tyrosine-Kinase Inhibitor Zanubrutinib (BGB-3111) in Patients with Relapsed/Refractory Waldenström Macroglobulinemia. *Clin Cancer Res* **27**, 5492 (2021).
35. Spronk, I., Polinder, S., Bonsel, G. J., Janssen, M. F. & Haagsma, J. A. Adding a fatigue item to the EQ-5D-5L improves its psychometric performance in the general population. *Journal of Patient-Reported Outcomes* **6**, 1 (2022).
36. National Institute for Health and Care Excellence. *Single Technology Appraisal: Ibrutinib for treating chronic lymphocytic leukaemia [ID749]*. https://www.nice.org.uk/guidance/ta429/documents/committee-papers (2016).
37. National Institute for Health and Care Excellence. *Single Technology Appraisal: Ibrutinib for treating relapsed or refractory mantle cell lymphoma [ID753]*. https://www.nice.org.uk/guidance/ta502/documents/committee-papers (2016).
38. Pratt, G. *et al.* 2. Guidelines on the diagnosis and management of Waldenström macroglobulinaemia—A British Society for Haematology guideline. *British journal of haematology* (2021).
39. medicines.org.uk. EMC: MAVENCLAD 10 mg tablets. *EMC: MAVENCLAD 10 mg tablets* https://www.medicines.org.uk/emc/product/8435/smpc#gref (2022).

NICE National Institute for Health and Care Excellence

# Clinical expert statement and technical engagement response form

## Zanubrutinib for treating Waldenstrom's macroglobulinaemia [ID1427]

Thank you for agreeing to comment on the evidence review group (ERG) report for this appraisal, and for providing your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature. The ERG report and stakeholder responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

## 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 for your views on key issues in the ERG report that are likely to be discussed by the committee. The key issues in the ERG report reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the ERG report [Section 1.1]. You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

A clinical perspective could help either:

- resolve any uncertainty that has been identified OR
- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.

In part 3 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 under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

**Please note, part 1** can be completed at any time. We advise that **part 2** is completed after the expert engagement teleconference (if you are attending or have attended). At this teleconference we will discuss some of the key issues, answer any specific questions you may have about the form, and explain the type of information the committee would find useful.

Deadline for comments by **5pm** on **Thursday 17th February**. 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 during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.**

Clinical expert statement

NICE National Institute for Health and Care Excellence

**Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.**

# Part 1: Treating Waldenstrom's macroglobulinaemia and current treatment options

**Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality**

| | |
|---|---|
| **1. Your name** | Dima El-Sharkawi |
| **2. Name of organisation** | Royal Marsden Hospital I am submitting on behalf of BSH/ RCPath |
| **3. Job title or position** | Haematology Consultant |
| **4. Are you (please tick all that apply)** | ☒ An employee or representative of a healthcare professional organisation that represents clinicians?<br>☒ A specialist in the treatment of people with Waldenstrom's macroglobulinaemia?<br>☒ A specialist in the clinical evidence base for Waldenstrom's macroglobulinaemia or technology?<br>☒ Other (please specify): Trustee for WMUK |
| **5. Do you wish to agree with your nominating organisation's submission?**<br>(We would encourage you to complete this form even if you agree with your nominating organisation's submission) | ☒ Yes, I agree with it<br>☐ No, I disagree with it<br>☐ I agree with some of it, but disagree with some of it<br>☐ 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.**<br>(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.** | Nil |

| | |
|---|---|
| **8. What is the main aim of treatment for Waldenstrom's macroglobulinaemia?**<br>(For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability) | Main aim of treatment is to control the disease, to prolong life and to lead to better quality of life by reducing some of the symptoms of the disease and its complications.<br><br>Any treatment choice should take into consideration that a lot of morbidity and mortality associated with WM is not due to the WM itself but other causes which may be indirectly related, e.g. infection risk, complications of treatment. (Castillo et al BJHaem 2015 169: 81-89) |
| **9. What do you consider a clinically significant treatment response?**<br>(For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount) | Response in clinical trials is based on international working group guidelines on response assessment (Owen 2013 160:171-176) primarily based on reduction in paraprotein level. However on a clinical day-to-day perspective I would say there are two aspects to this question that are clinically significant and equally important. Firstly, the indication for which the treatment was being given in the first place and its resolution- e.g. if treatment was commenced for hyperviscosity symptoms related to a high paraprotein, then reduction in paraprotein is very important, however, if it was for symptomatic anaemia, then the more clinically relevant factor is the improvement in haemoglobin rather than level of paraprotein reduction. Given the clinical symptoms and indications for treatment can be very varied, this makes this aspect quite difficult to summarise for all |

| | |
|---|---|
| | patients due to the number of rare complications that can occur all of which can be an indication for treatment.<br><br>The second aspect of clinically significant treatment response is length of time to next treatment. There is some evidence that depth of response with chemoimmunotherapy is predictive of progression-free survival and time to next treatment, however this may not be the case with all therapies, for example, whilst those who achieve a PR with ibrutinib have been reported to have a better PFS than those who achieve less than a PR, achieving a better response (very good partial response) did not result in further improvement in PFS in one retrospective study (Castillo et al BJHaem 2021 Feb;192(3):542-550.) |
| **10. In your view, is there an unmet need for patients and healthcare professionals in Waldenstrom's macroglobulinaemia?** | Yes there is with current treatments available to patients on the NHS.<br><br>The only option available is chemoimmunotherapy which can be effective for some patients, but may of our patients are older and frailer and thus may not be suitable for chemoimmunotherapy. Toxicity can be a concern with chemoimmunotherapy including risk of infection and secondary malignancies. We know that giving multiple lines of different chemotherapeutic regimens can lead to shorter times to next line of therapy with increasing concern about toxicity. |

| | |
|---|---|
| | I reiterate that many of these patients will die due to other causes rather than WM directly and so it is important to be able to give a treatment that could provide may patients a well tolerated oral option that can lead to meaningful and durable responses to their disease but minimise toxicity. |
| **11. How is Waldenstrom's macroglobulinaemia currently treated in the NHS?**<br>• Are any clinical guidelines used in the treatment of the condition, and if so, which?<br>• 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.)<br>• What impact would the technology have on the current pathway of care? | Waldenstrom Macroglobulinaemia (WM) is a rare B cell lymphoproliferative disorder. Patients are typically elderly (median age approx. 70 years at diagnosis) and symptoms occur as a consequence of bone marrow failure due to lymphoma infiltration, due to nodal disease or due to specific complications related to the IgM monoclonal protein produced by the lymphoma cells. The most common symptoms requiring therapy are anaemia, peripheral neuropathy and hyperviscosity syndrome. WM typically follows a relapsing and remitting course over many years and as a consequence patients will receive many different forms of chemotherapy.<br><br>There is no consensus on standard of care for initial therapy in WM. Internationally, and where available, choices of therapy include rituximab monotherapy, chemoimmunotherapy regimens, proteasome inhibitor containing regimens and BTK inhibitors. In the UK, frontline, the two most frequently used chemoimmunotherapy regimens used at present are R-bendamustine based on Rummel et al 2013 Lancet 1203-1210 and DRC (dexamethasone, rituximab and |

| | |
|---|---|
| | cyclophosphamide) based on a phase 2 study (JCO 2007 25(22):3344-9). There are no prospective trials comparing the two regimens, however clinical practice and retrospective evidence seem to suggest that R-bendamustine is associated with quicker, deeper and perhaps more prolonged responses but with added potential toxicity risks both short term (eg. Infections) and longer term (e.g. secondary MDS).<br><br>Treatment in the relapsed/ refractory setting is more varied and depends on again disease related factors, previous treatment, length of time of response to prior therapy, patient related factors. However, the majority of patients have been commenced on ibrutinib since its availability on the CDF. Prior to this, it would have been alternative chemoimmunotherapy regimens.<br><br>British Society of Haematology Guidelines have recently been published, BJHaem (Epub ahead of print, 2022).<br><br>There are also ESMO guidelines (Kastritis et al 2018 Annals of Oncology 29 (S4): iv41-iv50 and international consensus guidelines (Castillo et al Lancet Haematology 2020 e827-837) both of which include ibrutinib (with or without rituximab) as a treatment option for patients with symptomatic WM requiring therapy. |

| | |
|---|---|
| **12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?**<br>• How does healthcare resource use differ between the technology and current care?<br>• In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic)<br>• What investment is needed to introduce the technology? (for example, for facilities, equipment, or training) | Yes, we have vast experience in the use of this class of drug through use of ibrutinib in the CDF in patients with WM, and also these class of drugs are used in other cancers such as CLL. It is an oral therapy and so compared to current standard of care I would anticipate that this technology would use less healthcare resource as opposed to intravenous chemoimmunotherapy options which require daycare space and nursing time as well as intravenous access. This treatment lends itself also to virtual monitoring, and patients can be reviewed for some of their consultations virtually. It would be used in secondary care clinics i.e. haematology clinics in hospital for prescribing, monitoring of efficacy and toxicity. However primary care should be alerted to potential toxicity concerns for support in monitoring and management of toxicities if and when they occur. E.g. hypertension, potential drug interactions.<br>No new investment would be required. |
| **13. Do you expect the technology to provide clinically meaningful benefits compared with current care?**<br>• Do you expect the technology to increase length of life more than current care?<br>• Do you expect the technology to increase health-related quality of life more than current care? | Yes- this is a different way of treating WM compared to chemoimmunotherapy and so I believe that having this “extra line” of therapy available to our patients will lead to an increase in length of life and improvement in QoL whilst they are responding to the treatment due to reduction in disease burden. |

| | |
|---|---|
| **14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?** | No to the best of my knowledge. There is of course interest in understanding whether there are predictive markers for response to this treatment but nothing definitive has been identified as of yet. |
| **15. 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?**<br>(For example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed) | As above, I believe this technology will be easier to use than current care and cost saving in terms of "chair time" and "nursing time" due to lack of intravenous treatment required. This treatment lends itself also to virtual monitoring, and patients can be reviewed for some of their consultations virtually. It would be used in secondary care clinics i.e. haematology clinics in hospital for prescribing, monitoring of efficacy and toxicity. However primary care should be alerted to potential toxicity concerns for support in monitoring and management of toxicities if and when they occur. E.g. hypertension, potential drug interactions.<br>Prophylactic medication may be used alongside this technology in some centres to reduce the risk of infection, but this may be used in patients having chemoimmnotherapy too. |
| **16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?** | No different to standard of care |
| **17. Do you consider that the use of the technology will result in any substantial health-related benefits that** | Yes. This is oral therapy, compared to standard of care. |

| | |
|---|---|
| **are unlikely to be included in the quality-adjusted life year (QALY) calculation?**<br>• 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 | |
| **18. 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?**<br>• Is the technology a 'step-change' in the management of the condition?<br>• Does the use of the technology address any particular unmet need of the patient population? | Yes. It is imperative to have a BTK inihibitor in the treatment armamentarium that we have for treatment of WM, it provides an effective treatment with manageable toxicity profile. It allows us to be able to effectively treat patients who are either unlikely to benefit from chemoimmunotherapy as they have already had it previously or who could not have it for toxicity concerns. |
| **19. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?** | As the trial data has demonstrated, the toxicity profile of zanubrutinib is manageable and the quality of life data shows that the QoL improves on treatment as the burden of the disease reduces. |
| **20. Do the clinical trials on the technology reflect current UK clinical practice?**<br>• If not, how could the results be extrapolated to the UK setting?<br>• What, in your view, are the most important outcomes, and were they measured in the trials?<br>• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?<br>• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently? | Yes, I think the majority of my patients with WM would have fulfilled the eligibility criteria for this trial and indeed many UK centres participated in the ASPEN study and so I think this can be extrapolated. |

| | |
|---|---|
| **21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?** | no |
| **22. How do data on real-world experience compare with the trial data?** | Zanubrutinib has not been used sufficiently in the real world in patients with WM, to have experience in this yet. |
| **23. NICE considers whether there are any equalities issues at each stage of an appraisal. 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.**<br><br>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.<br>Please state if you think this appraisal could<br>• exclude any people for which this treatment is or will be licensed but who are protected by the equality legislation<br>• lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population<br>• lead to recommendations that have an adverse impact on disabled people.<br>Please consider whether these issues are different from issues with current care and why. | no |

| | |
|---|---|
| More information on how NICE deals with equalities issues can be found in the NICE equality scheme.<br>Find more general information about the Equality Act and equalities issues here. | |

# Part 2: Technical engagement questions for clinical experts

We welcome your comments on the key issues below, but you may want to concentrate on issues that are in your field of expertise. If you think an issue that is important to clinicians or patients has been missed in the ERG report, please also advise on this in the space provided at the end of this section.

The text boxes will expand as you type. Your responses to the following issues will be considered by the committee and may be summarised and presented in slides at the appraisal committee meeting.

For information: the professional organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the ERG report. These will also be considered by the committee.

**Table 2 Issues arising from technical engagement**

| | |
|---|---|
| **Key issue 1:**<br>The comparators are not in line with the NICE scope | I believe the comparators are in scope with what is used most frequently in the UK and what has been advised by the latest BSH guidelines (BJ Haem 2022 epub). In these guidelines we have stated that purine analgoues such as fludarabine whilst efficacious, are not recommended due to toxicity concerns such as long term risk of MDS and AML. This is reflective of real world practice where 0% of patients requiring 2nd line regimens for WM received purine analogues as per the 2nd Rory Morrison Registry report ( Rory-Morrison-Report-2021-2-11-21-Final-Version.pdf (wmuk.org.uk)). This trend is also reflected in the frontline setting too.<br><br>ASCT is a treatment option for a small minority of patients, but would be used as consolidation after chemoimmunotherapy ie. As well as rather than instead of. The majority of patients would not be suitable for ASCT given the median age at diagnosis and |

| | |
|---|---|
| | the toxicity concerned. For the younger patients, this would potentially be considered as an option for those with more aggressive disease. |
| **Key issue 2:**<br>Patients with cardiovascular disease and those taking warfarin were excluded from the ASPEN trial | In clinical practice, the majority of patients in the real world would have been eligible for this clinical trial based on the cardiac inclusion/exclusion criteria. Only those with significant cardiac disease were excluded due to the known cardiac toxicity seen with other BTK inhibitors. I think we would feel comfortable in those with some dardiovascular disease still considering this treatment option once cardiac disease has been optimised.<br>We would switch patients off warfarin onto an alternative agent in those for whom we were considering this drug anyway and so do not think this is a significant issue in generalisability. |
| **Key issue 3:**<br>The evidence for treatment naïve patients is based on small numbers of patients and has limited generalisability | This area is such a huge unmet need, given that they cannot derive benefit from chemoimmunotherapy and so are left with the only treatment options being rituximab monotherapy or chlorambucil which we know has inferior outcomes and so whilst the numbers are small, we can tell that they do at least as well as those who are having this therapy in the R/R setting. |
| **Key issue 4:**<br>Survival data for zanubrutinib are immature | Agree with this, and there are no other sources of data that we could use to extrapolate, but I would be confident that modelling longer term data could be performed given the very similar outcomes seen in the ASPEN study in terms of PFS between ibrutinib and zanubrutinib that the extrapolation could be made from the longer term studies seen with ibrutinib. |
| **Key issue 5:**<br>The indirect comparisons with rituximab and bendamustine (BR) and dexamethasone, rituximab and cyclophosphamide (DRC) are unreliable | Whilst this may be unreliable, I believe there is no truly reliable comparator that can be made at present and thus this may represent the best option available. Given I have no background in modelling and indirect comparisons, I cannot comment on the validity of the comparisons made. |

| | |
|---|---|
| **Key issue 6:**<br>The choice of a partitioned survival model and its underlying assumptions | I am a clinician and thus have no expertise in the modelling used and which choice is better. From a clinical perspective this is the model that patients go through:<br><br>1. Patient has indication for therapy and starts treatment with zanubrutinib.<br>2. They achieve a response and continue on the drug until either<br>i. They get toxicity requiring cessation of therapy<br>ii. They progress and have an indication for further therapy<br>iii. A minority of patients will not respond and will need to go onto next therapy<br>iv. Some patients will die on treatment<br><br>In scenario 2i.- they may have achieved sufficient response that there will be a period of time when they are not taking the zanubrutinib but before they progress and then time to next treatment.<br><br>In scenario 2ii.- they may progress and have a period of time whilst they are still taking zanubrutinib “in the progressed state” before they require the next line of therapy. This is because stopping the zanubrutinib may lead to acceleration of the disease progression and similarly there may be some time before they require the next line of therapy even though they have progressed by IWWM definition. |
| **Key issue 7:**<br>The model does not include all comparators mentioned in the NICE scope | Please see response to issue 1 |
| **Key issue 8:** | No further comment |

| | |
|---|---|
| Ibrutinib should not be included as direct comparator and/or subsequent treatment option in the economic model | |
| **Key issue 9:** The partitioned survival analysis chosen by the company relies on estimates for progression-free survival (PFS) and overall survival (OS), secondary and exploratory endpoints respectively. Only a small number of PFS and OS events had occurred at the time of this appraisal. | Agree and no further comment |
| **Key issue 10:** Plausibility of OS hazards falling below background mortality hazards. | This is not clinically valid. |
| **Key issue 11:** The use of data from patients with *MYD88*$^{MUT}$ only. | The majority of patients with WM have mutation in *MYD88* so this would be applicatble for these patients. For the few who are wild type, whilst they were excluded from the main part of the trial due to the reports of possible lower efficacy of ibrutinib in this cohort of patients, they were included in a small substudy that has shown similar PFS rates as the mutated cohort albeit in small numbers and with short follow up.<br><br>advancesadv2020003010absf1.png (1280×926) (silverchair-cdn.com)<br>*Blood Adv* (2020) 4 (23): 6009–6018. |
| **Key issue 12:** Assumption of lifelong treatment effectiveness. | Bar giving someone “extra time” as this is an extra line of effective therapy that otherwise would not be available to them, I do not believe it would lead to lifelong treatment effectiveness. |
| **Key issue 13:** | This is a statistical “quirk” and would not be clinically realistic. |

| | |
|---|---|
| PFS utility higher than general UK population values. | |
| **Key issue 14:**<br>The value and standard error implemented for post-progression utility is not evidence-based. | No comment |
| **Key issue 15:**<br>Large discrepancy between the deterministic incremental cost effectiveness ratio (ICER) and the probabilistic ICER. | No comment |
| **Key issue 16:**<br>**Treatment effectiveness being analysed for the different comparisons separately**. | No comment |
| **Additional issue:**<br>Are there any important issues that have been missed in ERG report? | no |

## Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

Step-change improvement compared to current standards of care for treatment of WM

Well tolerated and uses less healthcare resource for monitoring than intravenous treatment

Toxicity profile manageable

Click or tap here to enter text.

Click or tap here to enter text.

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.

For more information about how we process your personal data please see our privacy notice.

NICE National Institute for Health and Care Excellence

## Patient expert statement

# [Waldenstrom's Macroglobulinaemia - Zanubrutinib ID1427]

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

**Information on completing this expert statement**

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

| **About you** | |
|---|---|
| 1.Your name | **Ronald Vincent Presswood** |
| 2. Are you (please tick all that apply): | ☒ a patient with the condition?<br>☐ a carer of a patient with the condition?<br>☐ a patient organisation employee or volunteer? |

| | |
|---|---|
| | ☐ other (please specify): |
| 3. Name of your nominating organisation | WMUK |
| 4. Did your nominating organisation submit a submission? | ☒ yes, they did in January 2021.<br>☐ no, they didn't<br>☐ I don't know |
| 5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission) | ☒ yes, I agree with it<br>☐ no, I disagree with it<br>☐ I agree with some of it, but disagree with some of it<br>☐ other (they didn't submit one, I don't know if they submitted one etc.) |

| | |
|---|---|
| 6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.) | ☐ yes |
| 7. How did you gather the information included in your statement? (please tick all that apply) | ☒ I have personal experience of the condition<br>☒ I have personal experience of the technology being appraised<br>☐ I have other relevant personal experience. Please specify what other experience:<br>☐ I am drawing on others' experiences. Please specify how this information was gathered: |
| **Living with the condition** | |
| 8. What is it like to live with the condition? What do carers experience when caring for someone with the condition? | Diagnosed with WM on 10 December 2003 I was on active monitoring until the end of 2014 during which time my quality of life had become extremely poor because walking was particularly difficult without having to rest at regular intervals due to breathlessness. Climbing the stairs at home presented similar problems as did performing other routine tasks such as washing my car, etc.<br>On 8 January 2015 my key blood results were HGB 72, WCC 4.1, Platelets 72, Neutrophils 1.8, RCC 2.24 and paraprotein 30 g/L.<br>I commenced BR chemotherapy on 12 January 2015 and had one cycle of Bendamustine (Days 1 & 2) which was aborted shortly afterwards due to intolerance.<br>On 19 January 2015 a CT Scan confirmed minor lymphadenopathy with an enlarged spleen of 13cm and multiple enlarged abdominal lymph nodules of 1cm across in the short axis. |

| | |
|---|---|
| | One week later I was hospitalised with neutropenia which was treated as neutropenic sepsis.<br><br>Following discharge on 3 February 2015 Pancytopenia set in and by 13 February 2015 my blood spectrum reached an all time low with HGB 63, WCC 0.9, Platelets 14, Neutrophils 0.3, RCC 1.98 and paraprotein 27 g/L.<br><br>After discharge I underwent a traumatic period of on-going hospital visits (21 days) variously for blood tests, specific treatment &/or other procedures of what I considered to be “salvage therapy” until 15 May 2015 before I was considered well enough to undergo further chemotherapy. This period included, 3 blood transfusions, drugs taken orally and intravenously, a CT scan, a BMB and an X-ray.<br><br>Notably on 6 March 2015 the BMB outcome stated: Immunophenotyping results indicate the majority of lymphocytes present in the bone marrow are of T-cell lineage. No increase in B-cells. The bone marrow trephine shows heavy disease load of an indolent mature B-cell lymphoma, consistent with LPL. Cellularity has increased greatly (~ 60%).<br><br>On 18 May 2015 I commenced six cycles of DRC therapy, supported by five days of self administered G-CSF, which ended in mid-September. Throughout most of this period a feeling of acute nausea was commonplace in spite of having unsuccessfully used Ondansetron and Omeprazole to alleviate the problem. By day 3 of every cycle I could have cheerfully disposed of the remaining cyclophosphamide tablets down the toilet because of severe abdominal pain.<br><br>My period of remission only lasted for only 15 months due a very poor partial response to treatment before the need for further chemotherapy became inevitable. At the time I was advised that the Haematologists were in a quandary of what to do next because they felt unable to provide an alternative regimen because of toxicity considerations and had decided that although it was not normal to repeat the same regimen they had determined to do so because the risk was considered to be lower.<br><br>DRC treatment recommenced on 20 December 2016 but was aborted after cycle 3 due to intolerance, because my blood results taken on 20 February 2017 were considered to be particularly disappointing, viz. HGB 108, WCC 3.4, Platelets 59, Neutrophils 1.50, RCC 2.93 and paraprotein 21 g/L.<br><br>In March 2017 my Haematologist advised that he wanted to establish whether I had developed another treatment related cancer, namely MDS – Myelodysplastic syndrome (myelodysplasia), which can sometimes be induced by chemotherapy and is known as secondary or treatment related MDS. |

| | |
|---|---|
| | Thankfully, however, the BMB confirmed that this was not the case.<br><br>Having recently viewed a podcast produced by IWMF about a new immunotherapy using a BTK inhibitor, namely Ibrutinib, that was producing some encouraging results, I enquired of him whether there were any similar clinical trials using Ibrutinib in the UK and he agreed to contact Dr Shirley D'Sa, Consultant Haematologist at University College London Hospitals (UCLH), to seek her advice and establish whether there were any clinical trials recruiting in the UK for treatment with Ibrutinib.<br><br>I had my initial consultation with Dr D'Sa on 23 May 2017, followed by subsequent appointments, treatments and screening before finally starting the BGB-3111-302 clinical trial at UCH on 29 December 2017. I am indeed very fortunate to have been able to have done so and without her intervention with the Sponsor it would never have been possible.<br><br>The alternative option to receive on-going treatment in oral form, such as with Zanubrutinib, by taking 2 x 80 mg capsules twice a day has been fantastic. In my case with virtually no serious debilitating side effects and certainly no pain and with a restored quality of life not that dissimilar to the one I enjoyed prior to diagnosis has exceeded my expectations.<br><br>My most recent CT scan on 27 February 2022 confirms that my spleen is normal at 11cm and that two of my three abdominal target nodes are also normal. The Consultant Radiologist's opinion states:<br><br>**Technically still a PR by trial criteria, although close to CR.**<br><br>Additionally, my on-going blood spectrums for FBC, LFT, U&E are entirely satisfactory and stable as is my paraprotein concentration at 12 g/L. Notably, Zanubrutinib is the only drug that has had a significant impact on lowering and maintaining my paraprotein concentration. (previous high 35 g/L)<br><br>In a sentence – Zanubrutinib changed my life virtually overnight and I can best put this into perspective by providing statistics of my therapeutic annual walking exercise, viz.<br><br>2015 42 miles, 2016 373 miles, 2017 534 miles, 2018 754 miles, 2019, 1,023 miles, 2020 1,243 miles, 2021 1,204 and 2022 also on target for 1,200+ miles.<br><br>Although 83 I can more than hold my own with men some 10 to 15 years younger. |

| | |
|---|---|
| **Current treatment of the condition in the NHS** | |
| 9. What do patients or carers think of current treatments and care available on the NHS? | Considering that WM is still considered to be an orphan decease, in my experience a specifically targeted immunotherapy treatment option, pain and stress free, is infinitely better than the existing standard chemotherapy options. |
| 10. Is there an unmet need for patients with this condition? | Yes, some patients are either intolerant to or do not respond to existing chemotherapy regimens and therefore need an effective alternative. |
| **Advantages of the technology** | |
| 11. What do patients or carers think are the advantages of the technology? | Patients are readily able to treat themselves at home by orally taking 2 x 80 mg capsules of Zanubrutinib TWICE a day, which is more convenient, less stressful and considerably less time consuming and importantly free from pain and/or discomfort. |
| **Disadvantages of the technology** | |
| 12. What do patients or carers think are the disadvantages of the technology? | In my experience during the 4 years 3 months I have been taking Zanubrutinib I have had no significant adverse effects. |
| **Patient population** | |
| 13. Are there any groups of patients who might benefit | Some WM patients will either not respond favourably to or be able to tolerate current chemotherapy regimens and will need to access this new immunotherapy technology. |

| | |
|---|---|
| more or less from the technology than others? If so, please describe them and explain why. | |
| **Equality** | |
| 14. Are there any potential equality issues that should be taken into account when considering this condition and the technology? | No – Zanubrutinib should be made available to all WM patients. |
| **Other issues** | |
| 15. Are there any other issues that you would like the committee to consider? | If the committee cannot approve this technology based on relevant ICER calculations as a standard first line treatment option, it is imperative that it is approved for use by those patients that are unable to derive benefit from current chemotherapy options. |

| **Topic-specific questions** | |
|---|---|
| 16. [To be added by technical team if required, after receiving the company submission. For example, if the company has deviated from the scope (particularly with respect to comparators) – check whether this is appropriate. Ask specific, targeted questions such as "Is comparator X [excluded from company submission] considered to be established clinical practice in the NHS for treating [condition Y]?"]<br><br>**if not delete highlighted rows and renumber below** | |

| **Key messages** |
| --- |
| 17. In up to 5 bullet points, please summarise the key messages of your statement:<br><br>• Initial BR chemotherapy, a gold standard, had to be aborted after 1 cycle of Bendamustine (Days 1 & 2) due to induced neutropenia which was treated as neutropenic sepsis. This resulted in an additional 21 out-patient hospital visits before I was well enough to undertake further necessary remedial therapy, resulting in an overall delay of almost 5 months.<br>• Alternative DRC chemotherapy, another gold standard, resulted in a poor partial response and gave only 15 months of remission.<br>• Repeat DRC chemotherapy, a reluctant but necessary choice because of toxicity considerations was aborted after 3 cycles due to a disappointing and unsatisfactory blood spectrum.<br>• Alternative pain and stress free oral treatment with Zanubrutinib has restored the quality of my life and I am almost in a similar situation as was the case at the time of WM diagnosis<br>• There is an inevitable group of WM patients that either do not respond well or at all to chemotherapy that need to be able to survive and enjoy a reasonable quality of life by being able to access this new drug. Additionally, this could prove to be a more appropriate and cost effective first line treatment option for most if not all WM patients |

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

NICE National Institute for Health and Care Excellence

# Clinical expert statement and technical engagement response form

## Zanubrutinib for treating Waldenstrom's macroglobulinaemia [ID1427]

Thank you for agreeing to comment on the evidence review group (ERG) report for this appraisal, and for providing your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature. The ERG report and stakeholder responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

## 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 for your views on key issues in the ERG report that are likely to be discussed by the committee. The key issues in the ERG report reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the ERG report [Section 1.1]. You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

A clinical perspective could help either:

- resolve any uncertainty that has been identified OR
- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.

In part 3 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 under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

**Please note, part 1** can be completed at any time. We advise that **part 2** is completed after the expert engagement teleconference (if you are attending or have attended). At this teleconference we will discuss some of the key issues, answer any specific questions you may have about the form, and explain the type of information the committee would find useful.

Deadline for comments by **5pm** on **Thursday 17th February**. 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 during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.**

Clinical expert statement

**Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.**

# Part 1: Treating Waldenstrom's macroglobulinaemia and current treatment options

**Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality**

| | |
|---|---|
| **1. Your name** | Shirley D'Sa |
| **2. Name of organisation** | Employer: UCLH NHS FT; Nominated by WMUK as clinical expert |
| **3. Job title or position** | Consultant Haematologist |
| **4. Are you (please tick all that apply)** | ☐ An employee or representative of a healthcare professional organisation that represents clinicians?<br>☒ A specialist in the treatment of people with Waldenstrom's macroglobulinaemia?<br>☒ A specialist in the clinical evidence base for Waldenstrom's macroglobulinaemia or technology?<br>☒ Other (please specify): Trustee of WMUK |
| **5. Do you wish to agree with your nominating organisation's submission?**<br>(We would encourage you to complete this form even if you agree with your nominating organisation's submission) | ☒ Yes, I agree with it<br>☐ No, I disagree with it<br>☐ I agree with some of it, but disagree with some of it<br>☐ 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.**<br>(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 Waldenstrom's macroglobulinaemia?** | To reduce the disease burden in symptomatic patients in order to improve well-being and quality of life and extend survival. |

| | |
|---|---|
| (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability) | In the setting of WM, 'symptomatic' can include lymphoma-related and/or IgM-related symptoms. |
| **9. What do you consider a clinically significant treatment response?**<br>(For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount) | A clinically significant response comprises a clinical response that is meaningful for the affected patient, in the context of their health situation and performance status. This varies from patient to patient.<br>In general, a significant response is one that leads to an improvement in health-related QOL- fewer symptoms due to the disease such as less fatigue, more stamina, improvements (where relevant) in symptoms of hyperviscosity, peripheral neuropathy, abrogation of weight loss, less shortness of breath and so forth.<br>There are internationally recognised response criteria which are followed in the clinical setting to measure response, but the categorical response may not mirror the clinical experience of the patient, as time to categorical response can be delayed in the WM setting (especially in regards to IgM responses). |
| **10. In your view, is there an unmet need for patients and healthcare professionals in Waldenstrom's macroglobulinaemia?** | The natural history of WM is to develop chemo/immunotherapy resistance over time and in response to use of such treatments, which form the traditional backbone of therapy for WM.<br>Whilst WM is highly responsive to chemoimmunotherapy at the outset, as time passes, the resistance that builds up leads to treatment options running out, and the sequential suppression of the patient's wellbeing and immune system.<br>The cumulative immunosuppression due to sequential therapies results in a reduced survival due to infections.<br>The key unmet need for WM patients is the availability of alternative treatment options that work in ways that are different to conventional chemoimmunotherapy. Targeted therapies such as BTK inhibitors offer the promise of meeting this need. |
| **11. How is Waldenstrom's macroglobulinaemia currently treated in the NHS?**<br>• Are any clinical guidelines used in the treatment of the condition, and if so, which? | At present, first line therapy is commenced for symptomatic patients, and comprises chemoimmunotherapy combinations, most often Dexamethasone, Rituximab, Cyclophosphamide (DRC) or Bendamustine and Rituximab (BR) usually for 6 cycles, and adjusted for patient frailty and bone marrow reserve. |

| | |
|---|---|
| • 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.)<br>• What impact would the technology have on the current pathway of care? | The choice between DRC and BR is based on physician choice, but there are guidelines regarding which regimen to consider over the other (BCSH guidelines; Pratt et al, BJH 2022) for clinical reasons. It is important to note that there are numerous IgM-related phenomena that may accompany WM and trigger therapy irrespective of the burden of lymphoma. Such phenomena may influence the selection of therapy.<br><br>Once first line treatment is complete, as long as there is a response, therapy is stopped and the patient undergoes active monitoring every 3-6 months until the disease progresses once more and then the disease is restaged and further therapy considered. At relapse, the pattern of disease is typically similar to that of presentation, however additional complications may arise that dictate therapy choice, such as Bing-Neel syndrome (CNS disease), high grade transformation to aggressive lymphoma, development of a cryoglobulin or AL amyloidosis.<br><br>At first relapse and beyond, the choice of therapy is determined by the initial or previous line of therapy, the quality of response that was achieved, the general condition of the patient and the goals of therapy. The pathway is better defined than it used to be, with greater uniformity of practice.<br><br>Since the availability of Ibrutinib on the CDF since 2017, this class of drug has become a lifeline for patients with WM and become a highly popular choice at all stages of relapse due to the oral rather than parenteral administration and generally good side effect profile compared to the greater intensity of cytotoxic drugs. Of course, Ibrutinib is taken continuously until it no longer works which is a difference compared to chemoimmunotherapy that is taken for a fixed period of time.<br><br>Given the more favourable adverse event profile of Zanubrutinib compared to Ibrutinib, this treatment is likely to be used in preference to Ibrutinib if both were available. |

| | |
|---|---|
| | The crucial feature of Zanubrutinib is to add to the therapeutic armamentarium, given that the natural history of WM is become progressively refractory to chemoimmunotherapy over time. |
| **12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?**<br>• How does healthcare resource use differ between the technology and current care?<br>• In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic)<br>• What investment is needed to introduce the technology? (for example, for facilities, equipment, or training) | Zanubrutinib would be used in a similar way to Ibrutinib is used in current clinical practice, having been available on the CDF since November 2017.<br><br>This technology is a step change in terms of current hospital care needed to administer chemoimmunotherapy. The frequency of visits to hospitals will be much lower (once stable on Zanu, I would expect patients to be seen every 3 months (which is not much more than the interval at which actively monitored patients are reviewed off therapy. In comparison, chemoimmunotherapy requires visits to a daycare unit every 3 to 4 weeks whilst the treatment lasts.<br>The setting for this technology is secondary and beyond care.<br>No investment will be needed as there is universal familiarity with BTKi in haematological practice. No additional facilities, equipment or training will be needed. If anything, the introduction of this technology will free up time for patients needing chemoimmunotherapy, and reduce waiting times for such treatments. |
| **13. Do you expect the technology to provide clinically meaningful benefits compared with current care?**<br>• Do you expect the technology to increase length of life more than current care?<br>• Do you expect the technology to increase health-related quality of life more than current care? | I believe Zanu will most definitely provide clinically meaningful benefits for patients who are becoming less responsive to chemoimmunotherapy and will contribute to an increase in the length of life compared to if it is not available.<br>Having spoken at length to patients receiving chemoimmunotherapy and BTKi, Hr-QOL will undoubtedly increase due to fewer invasive hospital visits, need for intravenous access and excellent tolerance of the treatment.<br>This is especially important in frailer patients who would struggle with chemoimmunotherapy. |

| | |
|---|---|
| **14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?** | Definitely a huge benefit for those patients who are unsuited to chemoimmunotherapy due to frailty.<br>Zanu does appear to be more effective in patients who do not have the MYD88 L265P mutation; though this is a small proportion of patients, it could make the difference to their clinical outcome. |
| **15. 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?**<br>(For example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed) | Easier due to oral administration.<br>Concomitant treatments such as viral prophylaxis or PCP prophylaxis are frequently used in the setting of chemoimmunotherapy.<br>No additional tests needed. |
| **16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?** | The same rules that dictate the use of chemoimmunotherapy would apply to Zanu. No additional testing. |
| **17. 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?**<br>• 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 think the main benefit that would not be captured by QALY calculations relate to the oral administration at home and the less harsh side effect profile compared to chemoimmunotherapy. |
| **18. 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?** | For sure- this the product of translational work based on recent understanding of the biology of the disease that allows specific pathway proteins to be targeted by a treatment compared to the relatively blunderbuss approach of chemoimmunotherapy. |

| | |
|---|---|
| • Is the technology a 'step-change' in the management of the condition?<br>• Does the use of the technology address any particular unmet need of the patient population? | The technology adds a new mechanism of action with which to tackle the WM disease- and this makes it a step-change in the management of the disease.<br>It offers a lifeline for chemorefractory patients and those who cannot tolerate chemoimmunotherapy. |
| **19. 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 first gen BTKi, Ibrutinib is well-tolerated by most patients compared to chemoimmunotherapy. Zanu has the added benefit of a superior adverse effect profile than Ibrutinib and this bodes extremely well for WM patients. |
| **20. Do the clinical trials on the technology reflect current UK clinical practice?**<br>• If not, how could the results be extrapolated to the UK setting?<br>• What, in your view, are the most important outcomes, and were they measured in the trials?<br>• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?<br>• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently? | Yes- in fact 'real-world' patients who are typically older and unselected compared to trials patients are likely to especially benefit from this technology compared to those who were eligible for trials.<br>The most important outcomes were response rates and the improved adverse event profile.<br>No new adverse events have come to light. |
| **21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?** | Zanu penetrates the CNS so would have utility in the setting of Bing Neel syndrome |
| **22. How do data on real-world experience compare with the trial data?** | There is not much real world experience that I am aware of. |
| **23. NICE considers whether there are any equalities issues at each stage of an appraisal. 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.** | No |

| | |
|---|---|
| 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.<br>Please state if you think this appraisal could<br>• exclude any people for which this treatment is or will be licensed but who are protected by the equality legislation<br>• lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population<br>• lead to recommendations that have an adverse impact on disabled people.<br>Please consider whether these issues are different from issues with current care and why.<br>More information on how NICE deals with equalities issues can be found in the NICE equality scheme.<br>Find more general information about the Equality Act and equalities issues here. | |

# Part 2: Technical engagement questions for clinical experts

We welcome your comments on the key issues below, but you may want to concentrate on issues that are in your field of expertise. If you think an issue that is important to clinicians or patients has been missed in the ERG report, please also advise on this in the space provided at the end of this section.

The text boxes will expand as you type. Your responses to the following issues will be considered by the committee and may be summarised and presented in slides at the appraisal committee meeting.

For information: the professional organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the ERG report. These will also be considered by the committee.

**Table 2 Issues arising from technical engagement**

| | |
|---|---|
| **Key issue 1:** The comparators are not in line with the NICE scope | Ibrutinib has become de facto part of the treatment landscape for WM in the UK so one could argue for its inclusion |
| **Key issue 2:** Patients with cardiovascular disease and those taking warfarin were excluded from the ASPEN trial | I do not expect this to be an issue, if we extrapolate from the Ibrutinib experience and work on the basis that Zanu has fewer off-target effects. |
| **Key issue 3:** The evidence for treatment naïve patients is based on small numbers of patients and has limited generalisability | Agreed |
| **Key issue 4:** Survival data for zanubrutinib are immature | As with many novel therapies, this is true. If we are to do such a treatment full justice then some real world experience would be needed in addition to the ongoing follow up of the ASPEN study. |

| | |
|---|---|
| **Key issue 5:**<br>The indirect comparisons with rituximab and bendamustine (BR) and dexamethasone, rituximab and cyclophosphamide (DRC) are unreliable | No. These are the comparators in real life. |
| **Key issue 6:**<br>The choice of a partitioned survival model and its underlying assumptions | I do not feel able to comment |
| **Key issue 7:**<br>The model does not include all comparators mentioned in the NICE scope | I do not feel able to comment |
| **Key issue 8:**<br>Ibrutinib should not be included as direct comparator and/or subsequent treatment option in the economic model | This is a tricky one. It has been available and widely used in real world practice for the past 4 years so in a way, leaving it out of the comparison seems flawed. |
| **Key issue 9:**<br>The partitioned survival analysis chosen by the company relies on estimates for progression-free survival (PFS) and overall survival (OS), secondary and exploratory endpoints respectively. Only a small number of PFS and OS events had occurred at the time of this appraisal. | I do not feel able to comment |
| **Key issue 10:**<br>Plausibility of OS hazards falling below background mortality hazards. | I do not feel able to comment |

| | |
|---|---|
| **Key issue 11:**<br>The use of data from patients with *MYD88*$^{MUT}$ only. | Most patients with WM have the MYD88 mutation |
| **Key issue 12:**<br>Assumption of lifelong treatment effectiveness. | I do not feel able to comment |
| **Key issue 13:**<br>PFS utility higher than general UK population values. | I do not feel able to comment |
| **Key issue 14:**<br>The value and standard error implemented for post-progression utility is not evidence-based. | I do not feel able to comment |
| **Key issue 15:**<br>Large discrepancy between the deterministic incremental cost effectiveness ratio (ICER) and the probabilistic ICER. | I do not feel able to comment |
| **Key issue 16:**<br>**Treatment effectiveness being analysed for the different comparisons separately**. | I do not feel able to comment |
| **Additional issue:**<br>Are there any important issues that have been missed in ERG report? | |

## Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

It is essential to offer BTKi to patients with WM, due to the step change in treatment this has offered.

Click or tap here to enter text.

Click or tap here to enter text.

Click or tap here to enter text.

Click or tap here to enter text.

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.

For more information about how we process your personal data please see our privacy notice.

NICE National Institute for Health and Care Excellence

# Technical engagement response form

## Zanubrutinib for treating Waldenström's macroglobulinaemia [ID1427]

As a stakeholder you have been invited to comment on the evidence review group (ERG) report for this appraisal.

Your comments and feedback on the key issues below are really valued. The ERG report and stakeholders' responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

## Information on completing this form

We are asking for your views on key issues in the ERG report that are likely to be discussed by the committee. The key issues in the ERG report reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the ERG report.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the ERG report that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this appraisal, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

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 under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

Deadline for comments by **5pm** on **Wednesday 23 February 2022**. 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 during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.**

**Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.**

NICE National Institute for Health and Care Excellence

# About you

### Table 1 About you

| | |
|---|---|
| **Your name** | |
| **Organisation name: stakeholder or respondent**<br>(if you are responding as an individual rather than a registered stakeholder, please leave blank) | Janssen-Cilag |
| **Disclosure**<br>Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry. | N/A |

# Key issues for engagement

**All**: Please use the table below to respond to the key issues raised in the ERG report.

### Table 2 Key issues

| Key issue | Does this response contain new evidence, data or analyses? | Response |
|---|---|---|
| **Key issue 1:**<br>The comparators are not in line with the NICE scope | No | Currently, ibrutinib is available via the CDF for relapsed/refractory (RR) Waldenström's macroglobulinaemia (WM) patients and as such should be used neither as a model comparator nor as a subsequent therapy.<br>Should ibrutinib exit the CDF while the zanubrutinib ID1427 appraisal is still ongoing, ibrutinib would become a comparator in the RR WM population only and a subsequent treatment in front-line chemoimmunotherapy-unsuitable patients. |
| **Key issue 2:**<br>Patients with cardiovascular disease and those taking warfarin were excluded from the ASPEN trial | Yes/No | N/A |
| **Key issue 3:**<br>The evidence for treatment naïve patients is based on small numbers of patients and has limited generalisability | No | It should be noted that ASPEN (Cohort 1) evidence for the treatment-naïve cohort is limited to the **chemoimmunotherapy-unsuitable subgroup**.<br>Treatment-naïve patients suitable for chemoimmunotherapy were excluded in this trial as reflected in the population described in the final Scope for this appraisal:<br>*"Adults with Waldenström's macroglobulinaemia:*<br>• *who have had at least 1 prior therapy, or*<br>• *whose disease is untreated, for whom chemoimmunotherapy is unsuitable".* |

| | | |
|---|---|---|
| | | When referring to this patient group, **it should be clarified that is the chemoimmunotherapy-unsuitable treatment-naïve patients to avoid confusion**.<br><br>As noted in Key issue 1, ibrutinib CDF recommendation (TA491) does not include these patients and if ibrutinib were to be recommended during the zanubrutinib ID1427 appraisal, it could not become a comparator for this subgroup. |
| **Key issue 4:**<br>Survival data for zanubrutinib are immature | Yes/No | N/A |
| **Key issue 5:**<br>The indirect comparisons with rituximab and bendamustine (BR) and dexamethasone, rituximab and cyclophosphamide (DRC) are unreliable | Yes/No | WM is a very rare type of lymphoma and there is a high unmet need for effective treatments in this disease.<br>The evidence on WM standard of care (SoC) prior to the ibrutinib CDF recommendation (TA491) was very scarce, and since its recommendation, ibrutinib has become the treatment of choice for RR WM patients. Some pragmatism is therefore required when assessing the evidence available for SoC considering this situation.<br>Nonetheless, there are limitations in the indirect evidence presented in this appraisal. The use of a mixed cohort of treatment-naïve and RR patients in the pairwise MAICs for zanubrutinib while DRC evidence (Dimopoulos 2007/Kastritis 2015) is in treatment-naive patients and BR evidence (Tedeschi 2015) is in RR patients opens a source of uncertainty to assess the relative treatment effect of zanubrutinib. |
| **Key issue 6:**<br>The choice of a partitioned survival model and its underlying assumptions | Yes/No | N/A |
| **Key issue 7:** | Yes/No | N/A |

| | | |
|---|---|---|
| The model does not include all comparators mentioned in the NICE scope | | |
| **Key issue 8:**<br>Ibrutinib should not be included as direct comparator and/or subsequent treatment option in the economic model | No | Currently, ibrutinib is available via the CDF for RR WM patients and as such should be used neither as a model comparator nor as a subsequent therapy.<br>Should ibrutinib exit the CDF while the zanubrutinib ID1427 appraisal is still ongoing, ibrutinib would become a comparator in the RR WM population only and a subsequent treatment in front-line chemoimmunotherapy-unsuitable patients. |
| **Key issue 9:**<br>The partitioned survival analysis chosen by the company relies on estimates for progression-free survival (PFS) and overall survival (OS), secondary and exploratory endpoints respectively. Only a small number of PFS and OS events had occurred at the time of this appraisal. | No | Model long-term extrapolations for ibrutinib and zanubrutinib key efficacy outcomes PFS and OS (see CS Figure B.3.9 p91 and Figure B.3.10 p92) seem clinically implausible given ASPEN 19m follow-up trial PFS ITT results (see ERG report Figure 3.1 p47), and ASPEN 32m follow-up trial OS ITT results (see ERG report Figure 3.3 p50) which all show a high degree of overlap between KM curves for ibrutinib and zanubrutinib.<br>Therefore the 0.82 total discounted QALY difference between the two arms suggested in CS CE results for pairwise comparison 1 (see CS Table B.3.32 p142), and which is driven by PFS in the model, is not clinically plausible.<br>Any QALY difference between ibrutinib and zanubrutinib is expected to come from the treatment initiation phase, reflecting different safety profiles, and should be smaller than 0.82, as beyond this phase, PFS/OS ASPEN trial data shows overlapping KM curves.<br>Extrapolations using more mature data-cuts are more appropriate for decision making. In the absence of more mature data-cuts, Janssen considers the CS **scenario** analysis described in CS Table B.3.34 (p154), in which a **HR=1 is being applied to ibrutinib OS/PFS/TTD after 30 months**, yielding a 0.27 total discounted QALY difference, **seems clinically more plausible**. |
| **Key issue 10:**<br>Plausibility of OS hazards falling below background mortality hazards. | Yes/No | N/A |

| | | |
|---|---|---|
| **Key issue 11:**<br>The use of data from patients with *MYD88*$^{MUT}$ only. | Yes/No | WM patients with *MYD88*$^{WT}$ have a poorer prognosis compared to those with *MYD88*$^{MUT}$, which is the population studied in ASPEN.<br>The exclusion of *MYD88*$^{WT}$ patients from ASPEN Cohort 1 does not affect the relative treatment effect of zanubrutinib (Arm A) versus ibrutinib (Arm B), as the two treatment arms are randomised in this phase 3 trial. However, patients with *MYD88*$^{WT}$ are not excluded from the DRC (Dimopoulos 2007/Kastritis 2015) and BR (Tedeschi 2015) studies used in the base-case MAICs, therefore introducing a potential bias around the relative treatment effect of zanubrutinib in the *MYD88*$^{MUT}$ patient population.<br>Ibrutinib, which is a BTKi, has demonstrated efficacy in the *MYD88*$^{WT}$ population randomised in the phase 3 trial iNNOVATE (Buske 2021). Zanubrutinib is also a BTKi and there is 18m follow-up evidence of zanubrutinib efficacy in ASPEN Cohort 2/Arm C (Dimopoulos 2020); it would be pertinent to obtain clinical opinion to extrapolate the relative efficacy trial results for the *MYD88*$^{MUT}$ population (Cohort 1) to *MYD88*$^{WT}$ patients, to confirm the extent to which zanubrutinib is expected to be efficacious in the *MYD88*$^{WT}$ patient population vs SoC including ibrutinib. |
| **Key issue 12:**<br>Assumption of lifelong treatment effectiveness. | Yes/No | N/A |
| **Key issue 13:**<br>PFS utility higher than general UK population values. | Yes/No | N/A |
| **Key issue 14:**<br>The value and standard error implemented for post-progression utility is not evidence-based. | Yes/No | N/A |
| **Key issue 15:**<br>Large discrepancy between the deterministic incremental cost | Yes/No | N/A |

| | | |
|---|---|---|
| effectiveness ratio (ICER) and the probabilistic ICER. | | |
| **Key issue 16:**<br>**Treatment effectiveness being analysed for the different comparisons** separately. | Yes/No | N/A |


in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

# Zanubrutinib for Waldenström's macroglobulinaemia [ID1427]

## ADDENDUM: Critique of the company's response to Technical Engagement


| | |
|---|---|
| **Produced by** | Kleijnen Systematic Reviews Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University |
| **Authors** | Rob Riemsma, Reviews Manager, Kleijnen Systematic Reviews Ltd, UK<br>Ben Wijnen, Health Economist, Maastricht UMC, the Netherlands<br>Sabine Grimm, Health Economist, Maastricht UMC<br>Jeremy Howick, Systematic Reviewer, KSR Ltd<br>Nigel Armstrong, Health Economist, KSR Ltd<br>Willem Witlox, Health Economist, Maastricht UMC<br>Thomas Otten, Health Economist, Maastricht UMC<br>Kate Misso, Information Specialist, KSR Ltd<br>Manuela Joore, Health Economist, Maastricht UMC<br>Jos Kleijnen, Director, KSR Ltd, Professor of Systematic Reviews in Health Care, Maastricht University |
| **Correspondence to** | Rob Riemsma, Kleijnen Systematic Reviews Ltd<br>Unit 6, Escrick Business Park<br>Riccall Road, Escrick<br>York, YO19 6FD<br>United Kingdom |
| **Date completed** | 03/03/2022 |

**Company's response to technical engagement**

The purpose of this addendum is to provide a critique of the new evidence submitted by the company as part of their response to the technical engagement (TE) report.

In their response to technical engagement, the company submitted responses to the key issues raised in the Technical Report written by the National Institute for Health and Care Excellence (NICE) technical team, and some additional evidence relevant to these issues.

***Key issue 1: The comparators are not in line with the NICE scope***

As stated in the Evidence Review Group (ERG) report: fludarabine and rituximab (FR), fludarabine, cyclophosphamide and rituximab (FCR), and cladribine and rituximab (Clad-R) have not been included as comparators due to lack of data according to the company. Autologous stem cell transplantation (ASCT) has not been included in any of the literature searches reported in the company submission (CS). Ibrutinib has been included as a comparator. However, NICE explicitly excluded ibrutinib as a comparator.

In their response to TE, the company state *"while the Company acknowledge that this approach does not include all treatments within the National Institute for Health and Care Excellence (NICE) scope, we consider that the included comparators reflect standard of care for the vast majority of patients with WM in the UK"*. In addition, the company *"acknowledge the ERG decision to remove ibrutinib from the model in line with the NICE guidance for appraisals"*. However, the company does request that the evolving ibrutinib appraisal is monitored by the NICE team in relation to this appraisal.

Furthermore, the company point out that rituximab and bendamustine (BR) and dexamethasone, rituximab and cyclophosphamide (DRC) represent established clinical practice in the United Kingdom (UK) and that European and British guidelines, combined with registry data highlight the lack of use of remaining scoped comparators.

Dr. El-Sharkawi pointed out in her clinical expert statement that the comparators used in the CS *"are in scope with what is used most frequently in the UK and what has been advised by the latest BSH guidelines (BJ Haem 2022 epub). In these guidelines we have stated that purine analgoues such as fludarabine whilst efficacious, are not recommended due to toxicity concerns such as long term risk of MDS and AML"*.

***Key issue 2: Patients with cardiovascular disease and those taking warfarin***

The company states that patients taking warfarin are not within the target population of this appraisal.

The company also highlights that the exclusion of patients taking warfarin in the ASPEN trial is in line with the European licensed population for zanubrutinib: *"Warfarin or other vitamin K antagonists should not be administered concomitantly with BRUKINSA"*. As such, these patients would not be eligible for treatment with zanubrutinib in UK clinical practice and hence do not fall within the target population of this appraisal according to the company.

The clinical expert points out that *"in clinical practice, the majority of patients in the real world would have been eligible for this clinical trial based on the cardiac inclusion/exclusion criteria. Only those with significant cardiac disease were excluded due to the known cardiac toxicity seen with other BTK inhibitors"*.

***Key issue 3: The evidence for treatment naïve patients is based on small numbers of patients and has limited generalisability***

The company acknowledge that there are only a small number of treatment naïve patients within the ASPEN trial, but do not agree that it limits the generalisability of the ASPEN trial data to the UK because historically, treatment naïve patients have a better prognosis than patients with relapsed/refractory (R/R) Waldenström's macroglobulinaemia (WM).

The clinical expert also points out that *"whilst the numbers are small, we can tell that they do at least as well as those who are having this therapy in the R/R setting"*.

***Key issue 4: Survival data for zanubrutinib are immature***

The company acknowledges that survival data from the ASPEN are immature.

In their response to TE, the company provides extended long-term follow-up from Study 118E for ibrutinib in patients with R/R W/M (follow-up 59-months versus median follow-up 37-months previously). Therefore, no additional data for zanubrutinib have been provided.

To address uncertainty in survival data, the company digitised Kaplan-Meier (KM) plots obtained from Study 118E and long-term survival was extrapolated as demonstrated in Figures 2 and 3 in Appendix C of the company's response to TE. This analysis indicates, according to the company, that long-term survival as a result of treatment with ibrutinib should be expected, with mean extrapolated undiscounted survival ranging from 18.40 to 18.88 years (considering all-cause mortality, Table 8, Appendix C of the company's response to TE). In addition, the company states that *"whilst the OS data for zanubrutinib is immature, it is comparable to the long-term ibrutinib OS data"*.

The clinical expert states that she *"would be confident that modelling longer term data could be performed given the very similar outcomes seen in the ASPEN study in terms of PFS between ibrutinib and zanubrutinib that the extrapolation could be made from the longer term studies seen with ibrutinib"*.

***Key issue 5: The indirect comparisons with BR and DRC are unreliable***

The ERG pointed out in the original ERG report that the matching-adjusted indirect comparisons (MAICs) with BR and DRC are unreliable for the following reasons:

- Only progression-free survival (PFS), overall survival (OS), and adverse events (AEs) were considered as outcomes in the MAIC.
- These survival data for zanubrutinib are immature.
- There is a substantial risk of bias. The CS listed a range of baseline patient variables considered to be potential prognostic factors or effect modifiers and would therefore likely cause bias in a MAIC if the included studies had differences in these variables. As no study presented the requisite summary data to match on all variables, no MAIC matched on all these variables.
- In addition to the potential prognostic factors or effect modifiers listed in the CS, other variables are also to cause bias and were not matched for in the MAICs, including socio-economic status, year of study, location of study, general health of patients.
- Additionally, the definitions of outcomes were not always consistent between studies, and the interventions were administered differently in each study.
- Finally, it is unclear to what extent the MAICs are relevant to a contemporary National Health Service (NHS) population, given differences in baseline variables between the studies in the

MAICs (to which the patients in ASPEN were matched) and the patients with WM in UK clinical practice.

In the response to TE, the company performed an additional indirect treatment comparison utilising the simulated treatment comparison (STC) methodology. STC's were performed to indirectly compare zanubrutinib with BR and with DRC separately. Results from these STCs are very favourable for Zanubritinib: compared with BR, zanubrutinib was associated with statistically significantly improved PFS (hazard ratio [HR] [redacted]) and statistically significantly improved OS (HR [redacted]); compared with DRC, zanubrutinib was associated with statistically significantly improved PFS (HR [redacted]) and statistically significantly improved OS (HR [redacted]).

However, as with the MAICs, only PFS and OS were considered as outcomes in the STCs, and these survival data for zanubrutinib are immature. In addition, as the company points out, not all prognostic factors and effect modifiers could be considered in each STC since there were cases where no events occurred in patients with certain baseline characteristics. Therefore, the STCs are also unreliable.

***Key issue 6: The choice of a partitioned survival model and its underlying assumptions***

In response to TE, the company stated that the development of a state-transition model (STM) would be unnecessarily complex and would increase uncertainty (referring to technical appraisal (TA) 4910). While the STM has limitations on its own, as mentioned by the company, the NICE Decision Support Unit (DSU) Technical Support Document (TSD) 19 recommends presenting a STM alongside the PSM to assist in verifying the plausibility of the partitioned survival model (PSM) extrapolations. This may be particularly important given the data immaturity.

***Key issue 7: The model does not include all comparators mentioned in the NICE scope***

See response to key issue 1.

***Key issue 8: Ibrutinib should not be included as direct comparator and/or subsequent treatment option in the economic model***

The ERG acknowledges the difficulties arising from removing ibrutinib from the model. However, as mentioned in the ERG report, according to NICE's position statement on treatments currently in the Cancer Drug Fund (CDF), ibrutinib should not be included as direct comparator and/or subsequent treatment option in the economic model. The company is correct to state that it was not possible for the ERG to exclude the benefits of subsequent ibrutinib use (i.e. survival benefit) and as a consequence only the costs of ibrutinib subsequent treatment following progression on BR or DRC could be removed from the model (i.e. this would result in a higher incremental cost effectiveness ratio (ICER) for Zanubrutinib). During TE, the company has provided exploratory scenario analyses in which the post-progression survival across the BR and DRC treatments was reduced. These scenarios do not appear to be based on empirical data and lacked methodological explanations. In the revised base case, the company has included ibrutinib as subsequent treatment, i.e. included the treatment costs.

***Key issue 9: The partitioned survival analysis chosen by the company relies on estimates for PFS and OS, secondary and exploratory endpoints respectively. Only a small number of PFS and OS events had occurred at the time of this appraisal.***

No additional evidence or comments were provided by the company.

***Key issue 10: Plausibility of OS hazards falling below background mortality hazards.***

The company states that *"for a patient who was diagnosed at approximately 70 years"* (the majority of patients in the UK are diagnosed between 60-70 years, aligning with the baseline mean age in ASPEN) it would be clinically reasonable for this patient to achieve a normal life expectancy. The company acknowledges that for patients diagnosed at a younger age, *"achieving a normal life expectancy may be less likely"*. As a consequence, the company has fitted flexible regression models to the KM data, i.e. spline models with 1,2, and 3 knots. In the revised base-case, the company has opted for the flexible Odds k=1 model for OS extrapolation of zanubrutinib (matched to DRC) and DRC. However, the ERG was presented with insufficient information (i.e. details regarding assessment of the NICE DSU TSD 14 and 21 criteria) to adequately appraise the choice of these models. Hence, it is unclear if the estimation of parametric survival models is fully consistent with reported guidance from NICE DSU TSD 14 and 21 on (flexible methods for) survival analyses.

***Key issue 11: The use of data from patients with MYD88MUT only***

The company responded that *"across both zanubrutinib arms in Cohort 1 and Cohort 2, the ASPEN clinical trial included approximately an 80%:20% split of $MYD88^{MUT}$:$MYD88^{WT}$ patients"*. As stated in the ERG report, this is likely not to be reflective of UK clinical practice (i.e. 90% of $MYD88^{MUT}$ and 5-10% of $MYD88^{WT}$). The company states in the TE response that *"a weighted analysis would rebalance the ASPEN data to include slightly fewer $MYD88^{WT}$ patients"*, and *"hence would improve the clinical outcomes of the pooled patient population"*. The ERG emphasizes that such an analysis would also require information regarding the mix of mutations in the comparator arm. However, in line with the statements of the clinical experts, the impact of this assumption is likely to be minor.

***Key issue 12: Assumption of lifelong treatment effectiveness***

The company states that *"the ERG's decision to implement an arbitrary 5-year treatment waning within their base case is not evidence based, and instead relies on past appraisals in different populations which consider less efficacious treatment options"*. While the company is correct to state that a treatment waning assumption based on past appraisals in different population may be suboptimal, it is likewise not ideal to assume a lifelong treatment effect. Moreover, the responses of the clinical expert indicated that a lifelong treatment effect was unlikely: *"Bar giving someone "extra time" as this is an extra line of effective therapy that otherwise would not be available to them, I do not believe it would lead to lifelong treatment effectiveness"*.

***Key issue 13: PFS utility higher than general UK population values***

No new evidence was provided by the company. Moreover, the clinical expert indicated that *"this is a statistical "quirk" and would not be clinically realistic"*.

***Key issue 14: The value and standard error implemented for post-progression utility is not evidence-based***

The company has included the ERG's preferred assumption for the post-progression utility value within their revised base case.

***Key issue 15: Large discrepancy between the deterministic ICER and the probabilistic ICER***

The company states that *"the observed discrepancy between the deterministic and probabilistic results was driven by large variation of the survival curves across treatment arms"* and has hence revised the

probabilistic sensitivity analysis (PSA) code. While results between the deterministic and probabilistic analyses are now similar, the ERG would have liked to see convergence plots (as mentioned in the ERG report). Moreover, for the ERG report, the ERG implemented convergence plots in the model which demonstrated relatively stable results after approximately 2,000 runs. However, the resulting ICERs were structurally higher compared to the deterministic ICERs.

***Key issue 16: Treatment effectiveness being analysed for the different comparisons separately.***

In the response, *"the company acknowledges the concern from NICE and the ERG on the presentation of pairwise comparisons for zanubrutinib versus BR and versus DRC"* and has presented a revised base-case in which *"the cost-effectiveness results of zanubrutinib versus standard of care (consisting of 49% BR and 51% DRC based on the UK 2021 Rory Morrison Registry report) have been weighted to produce an overall ICER of zanubrutinib versus standard of care"*. However, this pooled analysis was not based on individual-patient data and therefore suffers from the same limitations as the original base-case analysis, i.e. differences in the populations between the BR and DRC MAIC in several characteristics.

**Updates to the company's base case following technical engagement**

See below an overview of the company's adjustments along with ERG comments.

**Table 1. Updates to the company's base case following technical engagement**

| # | Company update | Changes made | ERG comment |
|---|---|---|---|
| 1 | Key issue 1 - Pairwise comparisons of zanubrutinib versus BR, and versus DRC | Estimation of a weighted ICER to reflect standard of care in line with Ibrutinib NICE appraisal TA491 | Pooled analysis was not based on individual-patient data and therefore suffers from the same limitations as the original base-case analysis (i.e. differences in the populations between the BR and DRC MAIC in several characteristics). |
| 5 | Key issue 5 - The indirect comparisons with BR and DRC are unreliable | Addition of STC analyses vs. BR and DRC for PFS and OS endpoints. | Both the original MAIC and the STC are subject to bias. The ERG prefers to stick with the original MAIC. |
| 8 | Key issue 8 - Ibrutinib should not be included as direct comparator and/or subsequent treatment option in the economic model | Included treatment costs of subsequent Ibrutinib use. | This is not in line with the position statement issued by NICE. |
| 10/11 | Key issue 10 (mention as 11 in the company's TE response) - Plausibility of OS hazards falling below background mortality hazards. | Addition of flexible survival analyses for PFS and OS for zanubrutinib (matched BR), BR, zanubrutinib (matched DRC) and DRC treatment arms. | The ERG was presented with insufficient information (i.e. details regarding assessment of the NICE DSU TSD 14 and 21 criteria) to adequately appraise the choice of these models. Hence, it is unclear if the estimation of parametric survival models is fully consistent with reported guidance from NICE DSU TSD 14 and 21 on (flexible methods for) survival analyses. |

| # | Company update | Changes made | ERG comment |
|---|---|---|---|
| | | | Hence, the ERG prefers to stick with the original parametric distributions. |
| 15 | Key issue 15 - Large discrepancy between the deterministic ICER and the probabilistic ICER. | Restriction on dependent Gamma OS treatment effect covariate from varying to values greater than 0.<br><br>Variation of the Weibull OS independent BR curve programmed using the "Norm.Inv" function to prevent extreme variation of the scale and shape parameters which lead to almost an vertical OS curves for BR. | While results between the deterministic and probabilistic analyses are now similar, the ERG would have liked to see convergence plots (as mentioned in the ERG report). |

BR = rituximab and bendamustine; DRC = dexamethasone, rituximab and cyclophosphamide; DSU = Decision Support Unit; ERG = Evidence Review Group; ICER = incremental cost effectivenss ratio; MAIC = matching-adjusted indirect comparison; NICE = National Institute for Health and Care Excellence; OS = overall survival; PFS = progression-free survival; STC = simulated treatment comparison; TA = technology appraisal; TE = technical engagement; TSD = Technical Support Document

**Conclusion**

Based on the new evidence submitted by the company as part of their response to the TE report, the ERG did not perform any additional analyses. None of the implemented changes require a reconsideration of the ERG base case (see Table 1). Although the ERG acknowledges the fact that excluding the costs of subsequent ibrutinib use from the model while (implicitly) maintaining subsequent treatment benefits of ibrutinib may increase the ICER of zanubrutinib compared to BR and DRC, it was not possible for the ERG to remove this bias from the model. It is uncertain to what extent this may bias the ICER.

| | | |
|---|---|---|
| Company base case ICER (deterministic) | | £20,054 |
| **Company base case assumption** | **ERG preference** | **Assumption impact on company base case ICER (£)** |
| Ibrutinib costs and benefits included as **subsequent treatment option** | Ibrutinib costs **excluded**, subsequent treatment benefits of ibrutinib **not included in RCT data** | +£23,971 |
| No **treatment effect cut-off** | Assume treatment effect 5yr cut-off | -£10,973 |
| Use of **STC** instead of MAIC | MAIC | +£6,785 |
| **No treatment effect cut-off + ibrutinib costs & benefits included** | Assume 5-year cut-off and exclude ibrutinib costs | +£34,108 |
| **ERG exploratory base case (weighted ICER)** | | £78,383 |

**Table 1.1: Deterministic ERG base-case**

| Technologies | Total costs (£) | Total QALYs | Incremental costs (£) | Incremental QALYs | ICER (£/QALY) |
|---|---|---|---|---|---|
| **ERG base-case** | | | | | |
| Zanbrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | [redacted] | [redacted] | | | |
| Zanbrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | [redacted] | [redacted] | | | |
| Zanubrutinib | [redacted] | [redacted] | [redacted] | [redacted] | £78,383 |
| Soc | [redacted] | [redacted] | | | |
| **Company's corrected base-case (CS base case without ibrutinib costs & MAIC results)** | | | | | |
| Zanbrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | [redacted] | [redacted] | | | |
| Zanbrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | [redacted] | [redacted] | | | |
| Zanubrutinib | [redacted] | [redacted] | [redacted] | [redacted] | £53,210 |
| Soc | [redacted] | [redacted] | | | |

**Table 1.3: Probabilistic ERG base-case**

| Technologies | Total costs (£) | Total QALYs | Incremental costs (£) | Incremental QALYs | ICER (£/QALY) |
|---|---|---|---|---|---|
| Zanbrutinib (match BR) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| BR | [redacted] | [redacted] | | | |
| Zanbrutinib (match DRC) | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| DRC | [redacted] | [redacted] | | | |
| Zanubrutinib | [redacted] | [redacted] | [redacted] | [redacted] | £86,675 |
| Soc | [redacted] | [redacted] | | | |