NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (Review of TA397) [ID1591]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (Review of TA397) [ID1591]

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 GlaxoSmithKline
2. **Clarification questions and company responses**
3. **Patient group, professional group and NHS organisation submission** from:
   a. Lupus UK
   b. British Society of Rheumatology
   c. Renal Association
4. **Evidence Review Group report** prepared by Kleijnen Systematic Reviews
5. **Evidence Review Group – factual accuracy check**
6. **Technical engagement response** from GlaxoSmithKline
   a. Main response
   b. Company addendum
7. **Technical engagement responses from experts:**
   a. Chris Edwards, Consultant Rheumatologist – clinical expert, nominated by GlaxoSmithKline
   b. Peter Lanyon, Consultant Rheumatologist – clinical expert, nominated by NHS England Specialised Rheumatology CRG
   c. Jane Robinson, Patient and Public Voice representative – patient expert, nominated by NHS England Specialised Rheumatology CRG
8. **Technical engagement response from consultees and commentators:**
   a. Lupus UK
   b. British Society of Rheumatology
   c. Renal Association
9. **Evidence Review Group critique of company response to technical engagement** prepared by Kleijnen Systematic Reviews
   a. Main critique
   b. Addendum incorporating revised PAS for IV formulation

10. **Company additional clarification** from GlaxoSmithKline

11. **Evidence Review Group critique of additional clarification** from Kleijnen Systematic Reviews

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

Belimumab for treating active autoantibody-positive systemic lupus erythematosus [ID1591]

Document B

Company evidence submission

November 2020

File name	Version	Contains confidential information	Date
ID1591_Belimumab_Document B_22042021 [ACIC redacted]	**2.0**	**Yes**	**22/04/2021**

Contents

Tables and figures

Abbreviations

ACEi	Angiotensin-converting enzyme inhibitor
ACR	American College of Rheumatology
AE	Adverse event
AESI	Adverse events of special interest
ANCOVA	Analysis of covariance
ARB	Angiotensin II receptor blocker
AWR	Access with Research
BILAG	British Isles Lupus Assessment Group
BLyS	B lymphocyte stimulator
CD	Cluster of differentiation
CI	Confidence interval
CNS	Central nervous system
COVID	Coronavirus disease
C-SSRS	Colombia Suicide Severity Rating Scale
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded DNA
EQ-5D	EuroQol-5 dimensions
ERG	Evidence Review Group
FACIT	Functional Assessment of Chronic Illness Therapy
GSK	GlaxoSmithKline
HDA	High disease activity
HR	Hazard ratio
HRQoL	Health-related quality of life
IFN	Interferon
ITC	Indirect treatment comparison
ITT	Intention-to-treat
IV	Intravenous
IVIG	Intravenous immunoglobulin
LTE	Long-term extension
MCID	Minimal clinically important difference
MCS	Mental Component Score
NA	Not applicable
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NMSC	Non-melanoma skin cancer
NSAID	Non-steroidal anti-inflammatory drug
OR	Odds ratio

PCS	Physical Component Score
PD	Pharmacodynamics
PGA	Physician's Global Assessment
PK	Pharmacokinetics
PS	Propensity score
PSM	Propensity score-matching
QoL	Quality of life
RCT	Randomised controlled trial
RWE	Real-world evidence
SAE	Serious adverse event
SC	Subcutaneous
SD	Standard deviation
SDI	SLICC/ACR Damage Index
SE	Standard error
SELENA	Safety of Estrogen in Lupus Erythematosus National Assessment
SF-36	Short Form-36
SFI	SLE Flare Index
SLE	Systemic Lupus Erythematosus
SLEDAI	Systemic Lupus Erythematosus Disease Activity Index
SLICC	Systemic Lupus International Collaborating Clinics
SLR	Systematic literature review
SRI-4	SLE responder index-4
ST	Standard therapy
TLC	Toronto Lupus Cohort
TNF	Tumour necrosis factor
UK	United Kingdom
US	United States

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

B.1.1 *Decision problem*

The submission focuses on part of the technology's marketing authorisation. The proposed target population is narrower than the marketing authorisation and includes only a subgroup of patients with high disease activity defined by both clinical and serological markers.

Following TA397 (2016), belimumab was recommended in patients with evidence of high clinical (SELENA-SLEDAI score ≥10) and serological (low complement AND positive anti-double-stranded deoxyribonucleic acid [dsDNA]) disease activity; hereafter referred to as high disease activity subgroup-1 (HDA-1). This subgroup is where greater clinical benefit of belimumab can be expected, as patients fulfilling these criteria, experienced an increased belimumab treatment effect compared with the overall population of the pivotal BLISS trials.

As part of the managed access agreement following TA397, it was proposed to utilise the UK British Isles Lupus Assessment Group - Biologics Registry (BILAG-BR) for up to five years to generate real-world data for belimumab as prescribed in UK clinical practice. Since 2016, data collected in the BILAG-BR as part of the Managed Access Agreement (MAA) revealed that the number of patients receiving belimumab in England was substantially smaller than anticipated. One of the primary reasons identified for this was that the agreed target population (HDA-1) was too restrictive and that patients will often experience levels of high disease activity but only have one of the two defined serological biomarkers. Furthermore, patients who have both biomarkers at the time of diagnosis and are managed with current standard therapies, may subsequently experience normalisation of one of the two serological biomarkers but continue to have high disease activity clinically due to a suboptimal treatment response. Additionally, some patients with high disease activity may have an underlying complement deficiency and therefore access to belimumab would be

unattainable with the current criteria. These are important considerations when defining the most clinically appropriate criteria.

Therefore, to more accurately reflect the patient characteristics of a high disease activity subgroup encountered in clinical practice and to better address the unmet need in systemic lupus erythematosus (SLE) patients, we propose an alternative target population defined as ***patients with a SELENA-SLEDAI score ≥10 AND at least one of the following serological features: low complement OR positive anti-dsDNA***; hereafter referred to as high disease activity subgroup-2 (HDA-2). This definition combines routinely used objective laboratory measures with a clinical measure of disease activity. GSK believes that adopting this new target population will allow more patients with SLE to derive benefit from treatment with belimumab whilst still maintaining a cost-effective use of limited NHS resources.

This submission is generally consistent with the final NICE scope and the NICE reference case; all differences are summarised in Table 1.

Table 1. The decision problem

	Final scope issued by NICE/reference case	**Decision problem addressed in the company submission**	**Rationale if different from the final NICE scope**
Population	People aged 5 years or more with active, autoantibody-positive SLE with a high degree of disease activity despite standard therapy.	Phase 3 Trial Population: Patients with active autoantibody-positive SLE as enrolled in belimumab pivotal trials. High Disease Activity Subgroup-1 (HDA-1): Patients with a SELENA SLEDAI score ≥10 AND low complement AND positive anti-dsDNA (current NICE guidance population; TA397) High Disease Activity Subgroup-2 (HDA-2): patients with a SELENA-SLEDAI score ≥10 AND at least one of the following serological features: low complement OR positive anti-dsDNA – *The Base case*	Mindful of NHS resources, the proposed population of interest to this decision problem is a subgroup of the phase 3 trial population which applies the additional criteria of evidence for high serological (low complement AND positive anti-dsDNA) and clinical (SELENA-SLEDAI score of ≥10) disease activity. This subgroup experienced an additional treatment benefit of belimumab, resulting in the HDA-1 population becoming the recommended population within TA397. Following TA937, data collected as part of the managed access agreement since 2016 through the British Isles Lupus Assessment Group (BILAG) Biologics Registry (BR) has revealed that the number of patients receiving belimumab in England was substantially smaller than anticipated. This suggests that the HDA-1 population was too restrictive when applied in clinical practice and, to better address the unmet need in SLE and more accurately reflect patients with high disease activity, we propose belimumab be considered in the HDA-2 population defined as 'patients with a SELENA-SLEDAI score ≥10 AND at least one of the following serological features: low complement OR positive anti-dsDNA'. To support the adoption of the HDA-2 subgroup, it is our proposed base case for the economic modelling. GSK presents the results of PLUTO, the paediatric trial of IV belimumab compared with placebo within an appendix of the submission. The paediatric population recruited in PLUTO is limited (due to the rarity of childhood SLE) and the study was not statistically powered to show a difference between treatment groups. The economic evaluation will not specifically address a paediatric population; all inputted data pertains to an adult population. We assume that the

			resultant NICE guidance would apply to a paediatric population under the NHS England Commissioning policy for adolescents and paediatrics.
Intervention	Belimumab as an add-on to standard therapy.	As per the NICE scope. Please note that this submission refers to the previously appraised IV formulation and introduces a new subcutaneous (SC) formulation in the form of a pre-filled pen	SC formulation has been developed as an additional formulation to the currently available IV formulation, to offer physicians and patients a choice of treatment modalities based on the individual's needs, supporting increased access to treatment and adherence. It also reduces the burden on NHS resources as regular clinic time is not required for administration.
Comparator(s)	• Standard therapy alone. For people in whom it is considered appropriate: • Rituximab plus standard therapy • Cyclophosphamide plus standard therapy.	Evidence from clinical trials is available versus standard therapy alone; this is presented in this submission. Rituximab Although GSK acknowledges that rituximab would be used in patients eligible for belimumab if belimumab were not made available in the future, we have not conducted a formal indirect comparison versus rituximab. Cyclophosphamide is not included as a comparator.	Rituximab: With the lack of positive RCT data, and limited robust published observational data for rituximab, particularly in terms of long-term follow-up data, no attempt has been made to conduct a formal indirect comparison between rituximab and belimumab. The data provided for rituximab (Appendix P) from the BILAG-BR demonstrates the difficulty in assessment - how patients are managed on rituximab. Although a comparison of the two medicines is provided in Appendix P, these results should be interpreted with caution due to the observational nature of the study. Other statistical techniques, such as a matching adjusted indirect comparison, were not possible, due to the small sample size, particularly for belimumab. Considering rituximab as a comparator is not straightforward. Although rituximab could be used in patients eligible for belimumab if belimumab were not made available in the future, the recently published NHS England commissioning policy for rituximab in the treatment of SLE states that belimumab should be considered prior to rituximab in the treatment pathway. Data for rituximab collected from the BILAG-BR are presented in Appendix P to this submission for completeness.

			Cyclophosphamide: Used to treat patients with severe lupus. It is largely reserved for the treatment of lupus nephritis or CNS lupus, both of which are outside of the current marketing authorisation for belimumab. Therefore, cyclophosphamide plus standard therapy is not a relevant comparator for this appraisal. In addition and as stated by clinical experts in Section 4.3 of TA397adverse effects associated with long-term exposure to cyclophosphamide (bladder cancer, bone marrow suppression, haematologic malignancies, infections, myelodysplasia, and infertility[1])severely limits the use of cyclophosphamide in patients with SLE, who are more often women of childbearing age.
Outcomes	The outcome measures to be considered include: • disease activity • rate and duration of response • rate and duration of remission • incidence and severity of flares • impact on disease manifestations • incidence of long-term complications and/or organ damage • corticosteroid use • rate and duration of corticosteroid-free remission • mortality • health-related quality of life • adverse effects of treatment.	As per the scope, except for the rate and duration of remission.	The rate and duration of remission were therefore not considered to be suitable outcomes in this submission.
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed	As per the NICE reference case.	No deviation from NICE scope; however, only the adult SLE population was modelled as described above. The economic analysis used a lifetime horizon and captured relevant direct health effects and costs.

	in terms of incremental cost per QALY. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective. The availability of any commercial arrangements for the intervention, comparator and subsequent treatment technologies will be taken into account. The availability of any managed access arrangement for the intervention will be taken into account.		

B.1.2 *Description of the technology being appraised*

The description of belimumab IV and SC formulations is described in Table 2. Summaries of product characteristics for both the IV and SC formulations are provided in Appendix C.

Table 2. Technology being appraised

UK approved name and brand name	Belimumab, Benlysta®
Mechanism of action	Belimumab is a human IgG1λ monoclonal antibody that binds to soluble human B-lymphocyte stimulator (BLyS; also known as B cell activating factor) and inhibits its biological activity[2]. BLyS inhibits B cell apoptosis and stimulates the differentiation of B cells into immunoglobulin-producing plasma cells. Overexpression of BLyS by transgenic mice results in autoimmune-like disease[3]. Furthermore, BLyS is overexpressed in patients with systemic lupus erythematosus (SLE) and other autoimmune diseases[4, 5]. In patients with SLE followed for 2 years, BLyS levels correlated with changes in disease activity and with elevated anti-dsDNA antibody titres; worsening disease activity was predicted by rises in serum BLyS concentrations. Inhibition of BLyS by belimumab promotes apoptosis in autoreactive B cells[3].
Marketing authorisation/CE mark status	<u>Intravenous (IV) formulation</u> Benlysta® 120 mg powder for concentrate for solution for infusion. Benlysta® 400 mg powder for concentrate for solution for infusion. Marketing authorisation was granted by the European Commission on 13 July 2011. <u>Subcutaneous (SC) formulation</u> Benlysta® 200 mg solution for injection in pre-filled pen. Benlysta® 200 mg solution for injection in pre-filled syringe. A type II variation to the original marketing authorisation was approved by the European Commission in November 2017, introducing the SC formulation in Europe. SC formulation was temporarily made available in the UK from June 2020 until Dec 2020 to support existing patients on IV belimumab during the COVID-19 pandemic. An extension to the temporary supply period is subject to further discussion with the MHRA.

Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	Belimumab is indicated as add-on therapy in patients aged 5 years and older with active, autoantibody-positive systemic lupus erythematosus (SLE) with a high degree of disease activity (e.g., positive anti-dsDNA and low complement) despite standard therapy[6].
Method of administration and dosage	IV formulation[6]: The recommended dose regimen is 10 mg/kg on Days 0, 14 and 28, and at 4-week intervals thereafter. Premedication including an antihistamine, with or without an antipyretic, may be administered before the infusion of belimumab. The infusions should be administered by a qualified healthcare professional trained to give infusion therapy. SC formulation[6]: The recommended dose is 200 mg once weekly, administered subcutaneously. Dosing is not based on weight. The recommended injection sites are the abdomen or thigh. When injecting in the same region, patients should be advised to use a different injection site each week.
Additional tests or investigations	No additional tests or investigations are needed for selection of patients eligible for belimumab treatment other than those currently used routinely in clinical practice. The patient's condition should be evaluated continuously and discontinuation of treatment with belimumab should be considered if there is no improvement in disease control after 6 months of treatment[6].
List price and average cost of a course of treatment	List price: Benlysta® IV: £405 for the 400mg vial and £121.50 for the 120mg vial. For a patient with an average weight of 70kg, this equates to an annual price of £10,003.50 based on 13 infusions per year. Benlysta® SC, price for the 4-pack 200mg pen [redacted] Patient Access Scheme (PAS) price: Benlysta® IV: [redacted] Benlysta® SC: [redacted]
Patient access scheme (if applicable)	A simple discount patient access scheme is being offered with this medicine.

EPARs for belimumab IV and SC can be found in Appendix C.

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

B.1.3.1 The Health Condition - Systemic lupus erythematosus

Systemic lupus erythematosus (SLE) is a chronic autoimmune, multi-system disease with varying manifestations characterised by an unpredictable clinical course, autoantibody production, abnormal B lymphocyte function and chronic inflammation leading to high morbidity and mortality rate[7].

B.1.3.2 Epidemiology

In the UK, SLE affected nearly 1 in 1000 of the population with a female predominance of 9:1 (female: male ratio)[8]. It typically affects women of child-bearing age between the ages of 20 and 60 years[8], with the peak incidence between ages 40–49 years; considerably younger than the peak in men (60–69 years)[8]. Accordingly, it affects women in the 'prime' of life; during reproductive and working years. It is also more common in people of African-Caribbean and South Asian descent[8-10].

Whilst standardised mortality ratios (SMRs) in SLE have improved in the past 3-4 decades; nonetheless, mortality remains high with a 10% mortality over 20 years and a mean age of death of 53.7 years[11]. Around one in three patients in the UK develop lupus nephritis which can lead to end-stage renal failure (ESRF)[7]. A patient in whom lupus is diagnosed at 20 years of age still has a 1 in 6 chance of dying by 35 years of age, most often from the complications of lupus or infection[12].

B.1.3.3 Presentation and diagnosis of SLE

Diagnosis of SLE can be extremely challenging due to the complexity and heterogeneity of the condition. There are no definitive tests for diagnosing SLE and this is further complicated by the variation in presentation and the extent and severity of which clinical signs and symptoms can occur in any organ system; SLE can mimic other diseases at presentation and until a correct diagnosis is reached. In addition, patients can also have a combination of one or more rheumatological conditions

which adds further complexity and delays in diagnosis. Patients can often be referred to a number of different specialties within secondary care depending on their initial presentation, e.g. rheumatologist for joint pain, dermatologist for skin rash etc., prior to a diagnosis of SLE being made. A UK survey demonstrated a mean (SD) time to diagnosis from the first symptom of 6.4 (9.5) years, with 47% initially being given a different diagnosis prior to lupus[13].

The diagnosis of SLE is widely based on a set of clinical and laboratory criteria developed by the American College of Rheumatology (ACR) in 1982 and revised in 1997[14, 15]. In order for a diagnosis SLE to be established, four of 11 clinical and laboratory criteria must be met.[16] Other sets of classification criteria include the 2012 Systemic Lupus International Collaborating Clinics (SLICC) criteria[17] and the new ACR/EULAR joint criteria[18].

B.1.3.4 Burden of SLE

SLE is a relapsing and remitting condition with disease activity fluctuating between periods of exacerbation (flares) and relative quiescence.

SLE can affect multiple organs and systems including musculoskeletal, renal, central nervous system (CNS), cardiovascular, pulmonary, haematological, ophthalmic, muco-cutaneous and gastrointestinal systems, giving rise to a wide range of clinical manifestations and serological features such as decreased levels of complement and increased levels of autoantibodies[19, 20].

Patients typically present with symptoms involving the skin and joints, of which pain and fatigue are amongst the most debilitating symptoms interfering with daily life, domestic and professional activities, and social and sexual lives[21]. Factors that contribute to fatigue in SLE patients include depression, pain, poor sleep quality and physical fitness, perceived social support, potential side effects of medications and possibly disease activity[22]. Facial scarring and hair loss (alopecia) as a result of skin involvement can leave permanent physical and psychological scars. Inflammation of joints can result in pain, and impaired physical function, with a large proportion of those with SLE unable to remain in paid employment[23, 24]. Lupus inevitably forces a

patient to relinquish control of their lives, nullifying their ability to maintain normalcy or predictability.

In addition to the persistent risk of disease flares, long-term active SLE may cause irreversible organ damage [25,26, 27], leading to organ dysfunction (e.g. kidney failure, neurocognitive abnormalities and cardiovascular complications) and increased morbidity[28]. More than one-half of SLE patients develop permanent organ system damage, which progresses steadily over time[29]. Disease activity scores correlate significantly with organ damage in SLE patients with long-term disease activity (>10 years)[30].

Damage may result from previous disease activity leading to organ failure or from medications[28]. Therapy, especially long-term high-dose glucocorticoid treatment, can contribute to myopathy, osteoporosis, hypertension, diabetes, atherosclerotic vascular disease, infections, and death[29]. In a European observational study in patients with SLE, after 10 years of disease duration 72% of patients were receiving ongoing treatment with corticosteroids; the cumulative impact of both disease and choice of treatment likely had an impact on the total sum of end organ damage.[30] The accrual of organ damage either related to SLE itself or to SLE treatment should also be evaluated.

In addition to the autoimmune-mediated disease consequences of lupus, patients with SLE appear to be at high risk for other disease and therapy related morbidity, including infections, especially of the respiratory and urinary systems[31, 32]. atherosclerosis, vascular disease and coronary artery disease[33-35]; and haematological and solid tumours[36-38], as well as increased risk for mortality[39, 40]. SLE is also associated with significant maternal and foetal morbidity, including spontaneous abortion, pre-eclampsia, intrauterine growth restriction, foetal death and pre-term delivery[41].

Several studies have confirmed that patients with SLE have reduced health-related quality of life (HRQoL) compared with healthy individuals[23]. It impacts on all aspects of HRQoL, including physical and mental health, vitality, pain, social and emotional

functioning and activities of daily living. A high prevalence of disability has been associated with SLE and the number of patients leaving work due to disability increases with disease duration[24].

B.1.3.5 Clinical pathway of care and proposed positioning of belimumab

To improve long-term patient outcomes, the overarching aim of treatment should be the remission of disease symptoms and signs, the prevention of flares, the prevention of organ damage accrual, the minimisation of drug side effects, and improvement in patients' quality of life[42]. More specifically preventing flares and maintaining symptoms with the lowest possible dose of glucocorticoids.

Standard therapy (ST) includes, either alone or in combination, the use of antimalarials (e.g. hydroxychloroquine), non-steroidal anti-inflammatory drugs (NSAIDs), corticosteroids and immunosuppressants[42, 43]; many of which are unlicensed for SLE. In some patients, however, these treatments fail to adequately control their disease and therefore lead to increased corticosteroid use or use of unlicensed treatments i.e. rituximab, if patients have been assessed as not eligible for belimumab, or clinical trials[44]. Importantly, until but also since the approval of belimumab, there had been little therapeutic innovation in the field of SLE, with no new treatments developed and licensed for several decades.

The British Society for Rheumatology (BSR) guideline for the management of SLE in adults (2017)[43] is NICE accredited. However, the most recent guideline is provided by the European League Against Rheumatism (EULAR), in which recommendations for the management of SLE were updated in 2019[42]. The EULAR guidelines propose that belimumab should be considered in patients with non-renal SLE that is inadequately controlled (i.e., there is ongoing disease activity or frequent flares) on first-line treatments (typically hydroxychloroquine and corticosteroids +/- immunosuppressants), and an inability to taper daily corticosteroids doses i.e. ≤7.5 mg/day. The proposed position of belimumab within the SLE treatment pathway, taking into account EULAR and BSR guidelines, is presented in Figure 1.

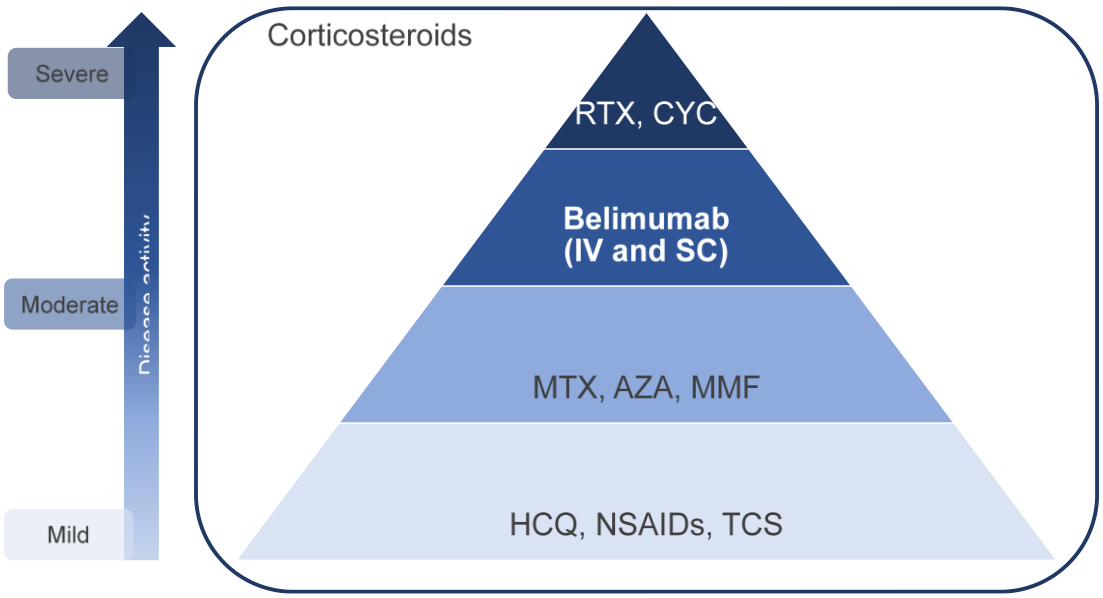


Figure 1. Proposed positioning of belimumab within the clinical pathway of care for SLE

AZA, azathioprine; CYC, cyclophosphamide; HCQ, hydroxychloroquine; IV: intravenous; MMF, mycophenolate mofetil; NSAIDs: non-steroidal anti-inflammatory drugs; MTX, methotrexate; RTX, rituximab; SC: subcutaneous; TCS: topical corticosteroids

B.1.4 *Equality considerations*

No equality considerations have been identified.

B.2 Clinical effectiveness

Belimumab has been previously assessed by NICE (TA397) and received a positive recommendation based on a managed access agreement as an add-on treatment option for patients with clinically active SLE (a Safety of Estrogen in Lupus Erythematosus National Assessment-Systemic Lupus Erythematosus Disease Activity Index [SELENA-SLEDAI] score ≥10) and high serological disease activity defined as low complement **AND** anti-double-stranded deoxyribonucleic acid (dsDNA) antibodies (June 2016). As part of the managed access agreement, it was proposed to utilise the UK British Isles Lupus Assessment Group Biologics Registry (BILAG-BR) over three to five years to generate real-world data for belimumab as prescribed in UK clinical practice. This submission provides additional information compared with TA397, across four key areas:

1. Present an update on the new evidence collected since the previous submission (2011).
2. Introduce a subcutaneous (SC) formulation of belimumab as an alternative option to the previously assessed intravenous (IV) formulation.
3. Introduce a new high disease activity subgroup (HDA-2, defined by SELENA-SLEDAI score ≥10 AND at least one of the following serological features: low complement OR positive anti-dsDNA), that is more clinically applicable and better reflects a subgroup of patients with HDA compared to the current NICE-approved target population, and in which treatment with belimumab is still likely to be more beneficial than in the Intent-to-Treat (ITT) population as defined in the pivotal trials.
4. Fulfil the obligation of the original market access agreement by presenting the evidence collected in the BILAG-BR.

Given the paucity of long-term and real-world data at the time of TA397, several uncertainties related to the use of belimumab within NHS England were identified during the appraisal process. The additional data presented in this submission addresses these uncertainties where possible, as outlined in Table 13.

B.2.1 *Identification and selection of relevant studies*

See Appendix D for full details of the process and methods used to identify and select the clinical evidence relevant to the technology being appraised.

B.2.1.1 Systematic literature review

A systematic literature review (SLR) previously conducted to support the 2011 NICE submission was updated to capture all evidence relating to the efficacy, safety, and tolerability of belimumab and appropriate comparators relevant to this submission. The previous SLR captured studies published from 1970 to August 2010, and the update covered literature from February 2010 to January 2020.

The search update in January 2020 yielded 1,376 unique records. Of those, 227 abstracts were accepted for further review at full text and 34 publications were

included. Additionally, five conference abstracts were included (four unique studies and one related publication; one RCT and four non-RCTs). Including both peer-reviewed publications and grey literature, the search update yielded 39 new publications (18 on RCTs and 21 on non-RCTs), representing 26 unique studies (10 RCTs and 16 non-RCTs). A summary of included publications and unique studies from both the SLR conducted in 2010 and the 2020 SLR update is provided in Table 3, with further details provided in Appendix D.

Table 3. Study yield by systematic literature review and updates

	Total Publications	Total Unique Studies	Total Publications (RCTs)	Total Unique Studies (RCTs)	Total Publications (non-RCTs)	Total Unique Studies (non-RCTs)
2010 SLR	45	39	43	38	2	1
2020 SLR	39	26	18	10	21	16
Total as of 2020	84	65	61	48	23	17

RCT: randomised controlled trial; SLR: systematic literature review.

In total, 61 publications of 48 unique RCTs were included. Nineteen studies were of low risk of bias according to the NICE checklist (January 2009) for quality assessment of RCTs; no studies were judged to have a high risk of bias. A total of 16 unique non-RCT studies across 21 publications were included in the 2020 update, of which five were open-label extensions of RCTs, 10 were cohort studies, and one was cross-sectional. Further results are shared in Appendix D.

B.2.2 *List of relevant clinical effectiveness evidence*

Standard therapy (ST) treatments for SLE include belimumab alone, antimalarials, immunosuppressants and/or steroids. Clinical evidence presented in this submission compares belimumab added onto to ST treatments for SLE. In addition, real-world data of belimumab is also presented.

For other comparators listed in the scope, i.e. rituximab and cyclophosphamide, no formal indirect comparisons are presented within this submission. Direct RCT data comparing rituximab to belimumab does not exist, and the justification for not

performing an indirect comparison is presented in Section B.2.9. However as the BILAG-BR has collected data on rituximab, an exploratory multilevel mixed effects regression modelling technique was undertaken for completeness (with appropriate cautions to its interpretation) to compare data for three comparable cohorts (belimumab, rituximab and non-biologics treatment (see BILAG-BR Appendix P).

Cyclophosphamide is now rarely used due to toxicity, and largely reserved for the treatment of severe or refractory disease, such as lupus nephritis. Indeed, the EULAR recommendations for the management of SLE state that "*Cyclophosphamide can be considered in organ-threatening disease (especially renal, cardiopulmonary or neuropsychiatric) and only as rescue therapy in refractory non-major organ manifestations. Due to its gonadotoxic effects, cyclophosphamide should be used with caution in women and men of fertile age*"[42]. Severe active lupus nephritis and CNS lupus are outside of the proposed target population for belimumab; therefore, cyclophosphamide plus standard therapy is not a relevant comparator. In addition, adverse effects associated with long-term exposure to cyclophosphamide including bladder cancer, bone marrow suppression, haematologic malignancies, infections, myelodysplasia, and infertility (SLE predominantly affects women of childbearing age)[1] limit the appropriateness of IV cyclophosphamide as a suitable comparator for belimumab. Of note, the Final Appraisal Determination (FAD) for TA397 notes that the Committee was aware that cyclophosphamide was also included as a comparator in the scope for the appraisal, but acknowledged GSK's justification that it was largely used for lupus nephritis, which was a different population to the one included in the trials of belimumab and covered by the current marketing authorisation for belimumab.

The following RCTs which evaluate the use of belimumab in addition to ST treatments are described in this submission:

- BLISS-SC: pivotal trial for the SC formulation of belimumab. The introduction of the SC formulation is further supported by a study of the SC autoinjector, which includes bridging study data on transitioning from IV to SC belimumab

based on a pharmacokinetic study, and an indirect treatment comparison (ITC) of the two formulations (Table 4).

- BLISS-52 and BLISS-76: pivotal trials for the IV formulation of belimumab. These trials have been described in the previous NICE submission (TA397), therefore pooled data from BLISS-52 and BLISS-76 is presented as a reminder of the results, with full study population data available in Appendix L. Whilst these studies present results of two belimumab doses (1 mg/kg and 10 mg/kg), only the 10 mg/kg dose will be presented in this submission. The current submission builds on these data by providing results from two open-label long-term extension (LTE) studies of BLISS-52 and BLISS-76, not previously available, which provide evidence on the long-term efficacy and safety of belimumab (see Table 5 and Table 6).

For the aforementioned RCTs, the following populations are considered:

- Full study population, as enrolled in the trial. For BLISS-SC, this is presented in Section B.2.6. Pooled data across the Phase 3 trials of IV belimumab, BLISS-52 and BLISS-76, are shown in Appendix L.

- HDA-1 (SELENA-SLEDAI score ≥10 AND low complement AND positive anti-double-stranded deoxyribonucleic acid [ds-DNA]) – current NICE guidance population, presented in Section B.2.7.

- HDA-2 (SELENA-SLEDAI score ≥10 AND at least one of the following serological features: low complement OR positive anti-dsDNA), presented in Section B.2.7. This population represents the base case for the economic evaluation described in Section B.3.2.1.

Additional relevant information provided in this submission:

- HDA population data: results from BLISS-SC[45] and pooled BLISS-52[46] and BLISS-76[47] trials for two HDA populations, the NICE-approved HDA-1 population and the new HDA-2 population, are provided in Section B.2.7.

- Systemic Lupus International Collaborating Clinics/American College of Rheumatology (SLICC/ACR) Damage Index (SDI) Indirect Cohort propensity score-matched (PSM) comparative analysis[48], to assess long-term organ damage in patients with SLE treated with belimumab. Further details are provided in Section B.2.6.

In addition, data for the full populations of BLISS-SC LTE, BLISS-52/76 non-US and BLISS-76 US LTEs are presented in Section B.2.6, and provide supportive evidence on long-term safety and efficacy of belimumab. Please note that these studies were non-randomised, open-label extension studies that primarily provided data on safety and tolerability of belimumab, as well as on long-term organ damage accrual (Table 7, Table 8, and Table 9.)

Table 4. Clinical effectiveness evidence – BLISS-SC[45]

Study	"A Phase 3, Multi-Center, Randomized, Double-Blind, Placebo- Controlled, 52-Week Study to Evaluate the Efficacy and Safety of Belimumab (HGS1006) Administered Subcutaneously (SC) to Subjects with Systemic Lupus Erythematosus (SLE) – Double-Blind Endpoint Analysis".				
Study design	Phase 3, multicentre, international, randomised, double-blind, placebo-controlled, 52-week study.				
Population	Patients with a clinical diagnosis of SLE according to the ACR criteria and clinically active SLE disease, defined as a SELENA-SLEDAI disease activity score of ≥8 at screening. Patients with severe lupus kidney disease, severe active lupus nephritis or CNS lupus were excluded.				
Intervention(s)	Belimumab 200 mg SC once weekly plus ST				
Comparator(s)	Placebo plus ST				
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	Evidence on the effectiveness of belimumab SC versus placebo, both added to ST (HDA-1 and HDA-2 populations)				
Reported outcomes specified in the decision problem	• Disease activity: Change in PGA and SELENA-SLEDAI score; BILAG scores; SDI-4 response and its components • Rate and duration of response: SRI-4 response by visit, and at Week 52 (primary efficacy endpoint) • Incidence and severity of flares: Time to SFI flare/severe flare and rate of SFI flare/severe flare per 100 subject years • Incidence of long-term complications and/or organ damage: SELENA-SLEDAI and BILAG scores by visit; SDI change at Week 52 • Corticosteroid use: Mean/median changes in steroid dose by visit; percent of patients whose average prednisone use reduced by ≥25% to ≤7.5 mg/day • Mortality: not assessed as an outcome, although included in safety reporting • HRQoL: FACIT-Fatigue Scale at Week 52 and by visit • Adverse effects of treatment: monitoring of adverse events, clinical laboratory tests, vital signs, physical examinations, and immunogenicity.				
All other reported outcomes	NA				

ACR: American College of Rheumatology; BILAG: British Isles Lupus Assessment Group; CNS: central nervous system; FACIT: Functional Assessment of Chronic Illness Therapy; HRQoL: health-related quality of life; NA: not applicable; PGA: Physician's Global Assessment; SC: subcutaneous; SDI: Systemic Lupus International Collaborating Clinics (SLICC)/ACR Damage Index; SELENA: Safety of Estrogen in Lupus Erythematosus National Assessment; SFI: SLE Flare Index; SLE: systemic lupus erythematosus; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index; SRI-4: SLE responder index-4; ST: standard therapy.

Table 5. Clinical effectiveness evidence – BLISS-52[46]

Study	"A Phase 3, Multi-Center, Randomized, Double-Blind, Placebo-Controlled, 52-Week Study to Evaluate the Efficacy and Safety of Belimumab (HGS1006, LymphoStat-B™), a Fully Human Monoclonal Anti-BLyS Antibody, in Subjects with Systemic Lupus Erythematosus (SLE)".				
Study design	Phase 3, multicentre, randomised, double-blind, placebo-controlled,52-week study				
Population	Patients with a clinical diagnosis of SLE according to the ACR criteria and clinically active SLE disease, defined as a SELENA-SLEDAI disease activity score of ≥6 at screening. Patients with severe active lupus nephritis or CNS lupus were excluded. Note that this submission presents pooled data across BLISS-52 and BLISS-76 (Table 6).				
Intervention(s)	Belimumab 1mg/kg IV plus ST or Belimumab 10 mg/kg (licensed dose) IV on days 0, 14, 28 and every 28 days thereafter, plus ST				
Comparator(s)	IV placebo plus ST				
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	Evidence on the effectiveness of belimumab IV versus placebo, both added to ST; note that BLISS-52 and BLISS-76 data are pooled (HDA-1 and HDA-2 populations)				
Reported outcomes specified in the decision problem	• Disease activity: Change in PGA and SELENA-SLEDAI score • Rate and duration of response: SRI-4 response by visit, and at Week 52 (primary efficacy endpoint) • Incidence and severity of flares: time to first flare and first severe flare, number of flares and severe flares per subject-year • Incidence of long-term complications and/or organ damage: Change in SELENA-SLEDAI, BILAG scores, and change in SDI at Week 52 • Corticosteroid use: Percent of patients whose average prednisone use reduced by ≥25% to ≤7.5 mg/day • Mortality: not assessed as an outcome, although included in safety reporting • HRQoL: FACIT-Fatigue Scale, SF-36, and EQ-5D • Adverse effects of treatment: monitoring of adverse events, clinical laboratory tests, vital signs, and immunogenicity				
All other reported outcomes	NA				

ACR: American College of Rheumatology; BILAG: British Isles Lupus Assessment Group; BLyS: B lymphocyte stimulator; CNS: central nervous system; EQ-5D: EuroQol 5-dimensions; FACIT: Functional Assessment of Chronic Illness Therapy; HRQoL: health-related quality of life; IV: intravenous; NA: not applicable; PGA: Physician's Global Assessment; SDI: Systemic Lupus International Collaborating Clinics (SLICC)/ACR Damage Index; SELENA: Safety of Estrogen in Lupus Erythematosus National Assessment; SF-36: Short Form 36; SFI: SLE Flare Index; SLE: systemic lupus erythematosus; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index; SRI-4: SLE responder index-4; ST: standard therapy.

Table 6. Clinical effectiveness evidence – BLISS-76[47]

Study	"A Phase 3, Multi-Center, Randomized, Double-Blind, Placebo-Controlled, 76-Week Study to Evaluate the Efficacy and Safety of Belimumab (HGS1006, LymphoStat-B™), a Fully Human Monoclonal Anti-BLyS Antibody, in Subjects with Systemic Lupus Erythematosus (SLE)".				
Study design	Phase 3, multicentre, randomised, double-blind, placebo-controlled, 76-week study.				
Population	Patients with a clinical diagnosis of SLE according to the ACR criteria and clinically active SLE disease, defined as a SELENA-SLEDAI disease activity score of ≥6 at screening. Patients with severe active lupus nephritis or CNS lupus were excluded. Note that this submission presents pooled data across BLISS-52 (Table 5) and BLISS-76.				
Intervention(s)	Belimumab 1mg/kg IV plus ST or Belimumab 10 mg/kg IV (licensed dose) on days 0, 14, 28 and every 28 days thereafter plus ST				
Comparator(s)	IV placebo plus ST				
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	Evidence on the effectiveness of belimumab IV versus placebo, both added to ST; note that BLISS-52 and BLISS-76 data are pooled (HDA-1 and HDA-2 populations).				
Reported outcomes specified in the decision problem	As per BLISS-52, plus SRI-4 response rate at Week 76.				
All other reported outcomes	NA				

ACR: American College of Rheumatology; CNS: central nervous system; IV: intravenous; NA: not applicable; SELENA: Safety of Estrogen in Lupus Erythematosus National Assessment; SLE: systemic lupus erythematosus; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index; SRI-4: SLE responder index-4; ST: standard therapy.

Table 7. Clinical effectiveness evidence – BLISS-SC LTE[49]

Study	"A Phase 3, Multi-Center, Randomized, Double-Blind, Placebo-Controlled, 52-Week Study to Evaluate the Efficacy and Safety of Belimumab (HGS1006) Administered Subcutaneously (SC) to Subjects with Systemic Lupus Erythematosus (SLE) - **Open-label phase**".				
Study design	Although the LTE to BLISS-SC is described separately in this submission, it was a multicentre, open-label, 6-month extension phase of BLISS-SC, defined within the same trial protocol as the double-blind phase described in Table 4.				
Population	Patients were eligible to participate in this open-label phase of BLISS-SC if they completed the Week 52 visit of the double-blind phase and were scheduled to receive the first belimumab dose in the extension phase approximately 1 week after the last study treatment dose in the double blind phase of the trial.				
Intervention(s)	Belimumab 200 mg SC once weekly plus ST				
Comparator(s)	None (patients continued to receive ST)				
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	BLISS SC-LTE not included in the model due to limited follow up period. An alternative extension study with longer follow-up is utilised (integrated analysis of Phase 2 and Phase 3 IV LTE studies) to estimate the year 2 onwards discontinuation rate (for HDA-1 and HDA-2 populations).				
Reported outcomes specified in the decision problem	As per BLISS-SC (Table 4); however, efficacy data were only collected at the end of the LTE phase (Week 24) or at the exit visit (1–4 weeks after the last belimumab dose) for those discontinuing the study early.				
All other reported outcomes	NA				

LTE: long-term extension; NA: not applicable; SC: subcutaneous; ST: standard therapy.

Table 8. Clinical effectiveness evidence – BLISS 52/76 non-US LTE[50]

Study	"A Multi-Center, Continuation Trial of Belimumab (HGS1006, LymphoStat-B), a Fully Human Monoclonal Anti-BLyS Antibody, in Subjects with Systemic Lupus Erythematosus (SLE) who Completed the Phase 3 Protocol HGS1006-C1056 or HGS1006-C1057".				
Study design	Multicentre continuation trial of belimumab in SLE patients who completed the Phase 3 BLISS-52 or BLISS-76 trial.				
Population	Non-US patients who completed either BLISS-76 through the Week 72 visit or BLISS-52 through the Week 48 visit were eligible. Patients with significant non SLE-related conditions, or those who, in the Phase 3 trials, had experienced an adverse event that would put them at an undue risk, or developed other conditions that made them unsuitable for the study were excluded from the trial.				
Intervention(s)	Belimumab 1mg/kg IV plus ST (prior to protocol amendment 01 only) or Belimumab 10 mg/kg IV (licensed dose) every 28 days plus ST				
Comparator(s)	None (patients continued to receive ST)				
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	BLISS SC-LTE not included in the model due to limited follow up period. An alternative extension study with longer follow-up is utilised (integrated analysis of Phase 2 and Phase 3 IV LTE studies) to estimate the year 2 onwards discontinuation rate (for HDA-1 and HDA-2 populations).				
Reported outcomes specified in the decision problem	• Efficacy/ Safety: SDI • Safety: adverse event monitoring and laboratory tests				
All other reported outcomes	NA				

IV: intravenous; LTE: long-term extension; NA: not applicable; SDI: Systemic Lupus International Collaborating Clinics (SLICC)/ACR Damage Index; SLE: systemic lupus erythematosus; ST: standard therapy; US: United States.

Table 9. Clinical effectiveness evidence – BLISS 76 US LTE[51]

Study	"A Multi-Center, Continuation Trial of Belimumab (HGS1006, LymphoStat-B), a Fully Human Monoclonal Anti-BLyS Antibody, in Subjects with Systemic Lupus Erythematosus (SLE) Who Completed the Phase 3 Protocol HGS1006-C1056 in the United States".				
Study design	Multicentre continuation trial of belimumab in SLE patients who completed the Phase 3 BLISS-76 trial in the US.				
Population	Patients who completed BLISS-76 through Week 72 were eligible. Patients with significant non SLE-related conditions, or those who, in the Phase 3 trial, had experienced an adverse event that would put them at an undue risk, or developed other conditions that made them unsuitable for the study were excluded from the trial.				
Intervention(s)	Belimumab 1mg/kg IV plus ST (prior to protocol amendment 01 only) or Belimumab 10 mg/kg IV (licensed dose) every 28 days plus ST				
Comparator(s)	None (patients continued to receive ST)				
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	Used as key evidence for the primary analysis in the SLICC/ACR Damage Index (SDI) Indirect Cohort Study[48], to assess long-term organ damage in patients with SLE treated with belimumab.				
Reported outcomes specified in the decision problem	• SRI-4 response rate • Efficacy/Safety: adverse event monitoring, laboratory tests, and SDI				
All other reported outcomes	NA				
ACR: American College of Rheumatology;IV: intravenous; LTE: long term extension; NA: not applicable; SDI: (SLICC)/ACR Damage Index; SLE: systemic lupus erythematosus; SLICC: Systemic Lupus International Collaborating Clinics; SRI-4: SLE responder index-4; ST: standard therapy; US: United States.					

Non-RCT evidence, supplementing the RCT data, includes two studies:

- BILAG-BR– as part of the Managed Access Agreement (MAA) following the previous NICE appraisal (TA397), we agreed to collect UK real-world data from this registry over a 3-5-year period for belimumab (Table 10).
- OBSErve – a series of ongoing, real-world, retrospective, observational studies conducted in the US, Germany, Spain, Canada, Argentina, and Switzerland to evaluate the use of belimumab in real-world clinical practice (Table 11).

Table 10. Clinical effectiveness evidence – BILAG-BR[52, 53]

Study	BILAG Biologics Prospective Cohort: The Use of Novel Biological Therapies in the Treatment of Systemic Lupus Erythematosus (SLE).				
Study design	Independent, investigator-led prospective cohort study consisting of two cohorts of patients, all of whom were treated by their consultant according to clinical need and the consultant's decision in their usual clinical setting.				
Population	Patients treated with biological therapies (commencing treatment for SLE with a biological agent within the previous 12 months) were recruited along with a control group (commencing treatment for SLE with a non-biological, immunosuppressive agent, with similar disease characteristics but exposed only to non-biological systemic therapies. Please note that for the BENLYSTA Sub-Study, which includes belimumab-treated patients, this registry provides data on the currently NICE-approved HDA-1 population and is presented in Section B.2.7.				
Intervention(s)	Any biologic therapy (data on belimumab-treated patients are presented herein) plus ST				
Comparator(s)	Non-biologic therapy, rituximab plus ST				
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	Limited data points were available as input for the economic analysis. Note that the IV model includes the BILAG-BR-captured weight distribution for belimumab patients which is used to calculate the number of vials of belimumab required (based on 10 mg/Kg dosing).				
Reported outcomes specified in the decision problem	• Disease activity: SLEDAI-2K; BILAG index; SDI • HRQoL: LupusQoL; SF-36v2; EQ-5D • Steroid use • Adverse effects of treatment: Serious adverse events; adverse events of special interest.				
All other reported outcomes	• Study population characteristics • Hospitalisations due to SLE in each year of follow-up • Time to treatment discontinuation.				

BILAG: British Isles Lupus Assessment Group; BR: Biologics Registry; HDA: high disease activity; HRQoL: health-related quality of life; LTE: long-term extension; NICE: National Institute for Health and Care Excellence; QoL: quality of life; SDI: Systemic Lupus International Collaborating Clinics (SLICC)/ACR Damage Index; SLE: systemic lupus erythematosus; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index.

Table 11. Clinical effectiveness evidence – OBSErve registry[54]

Study	OBSErve registry – Evaluation of use of belimumab in clinical practice settings.				
Study design	A series of non-randomised, single-arm, retrospective, observational studies over 2 years in the US[55] and in Argentina[56], and over 6 months in Germany[57], in Spain[58], in Canada[59], and in Switzerland[60].				
Population	Adults (aged ≥18 years) with a clinical diagnosis of SLE who had initiated IV belimumab as part of their usual SLE care ≥6 months prior to enrolment and for whom reasons for belimumab initiation could be identified.				
Intervention(s)	Belimumab IV as part of usual clinical care, plus ST.				
Comparator(s)	None.				
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	Alternative studies, providing longer follow-up and/or an RCT setting are available to inform the model.				
Reported outcomes specified in the decision problem	NA				
All other reported outcomes	• Physician-assessed overall clinical response • Description of patient characteristics • Treatment patterns • Patient and treatment characteristics associated with an overall clinical response				

IV: intravenous; NA: not applicable; RCT: randomised controlled trial; SLE: systemic lupus erythematosus; ST: standard therapy; US: United States.

In addition, a summary of data from a number of studies and analyses is provided in the Appendices:

- PLUTO: is a paediatric trial of IV belimumab compared with placebo (both in addition to ST)[61], with an ongoing long-term follow-up phase (see Table 52). Further details of the PLUTO trial are available in Appendix O.

- LBSL02 Phase 2 belimumab trial[62] (further details in Appendix L) and its LTE[63] (further details in Appendix M) provided data on up to 13 years of belimumab exposure.

- Pooled BLISS-52[46] and BLISS-76[47] data: the pooled results pertaining to the full study populations were presented as part of the previous NICE submission, and are provided in Appendix L for completeness.

- BASE post-marketing safety study[64] assessed mortality and adverse events of special interest in SLE patients over 52 weeks. Further details are provided in Appendix F, with data on steroid use presented in Appendix O.

- Treatment holiday study (NCT02119156)[65], a post-marketing commitment study that investigated the effects of belimumab treatment holiday and reintroduction, and treatment discontinuation. Further details are provided in Appendix O.

- Two studies in key ethnic populations: the post-marketing EMBRACE[66] trial (people of black race), and the pivotal trial NCT01345253 (SLE patients in North-East Asia): Further details for these are provided in Appendix O. While two LTEs to the North-East Asia trial have been conducted, one in Japan and Korea and the other in China, the results of these LTEs are not presented within this submission or the Appendices due to the lack of generalisability to the UK population. The results can, however, be provided upon request.

A summary of all evidence presented in the main body and appendices to this submission is provided in Table 12. In response to the NICE appraisal of belimumab

(TA397), extensive additional evidence has been generated which addresses the key areas of uncertainty, outlined in Table 13.

Trials and observational studies not included in the economic model provide supportive, long-term and/or real-world evidence on the efficacy of belimumab. This additional evidence is relevant to the decision problem to facilitate informed decision making.

Table 12. Summary of presented evidence

Clinical trials				
Trial name	**Description**	**Population (Total/HDA-1/HDA-2)**	**Included in the previous NICE submission? (Yes/No)**	**Location in the current submission**
BLISS-52 and BLISS-76	Pivotal trials of IV belimumab	Total (pooled across BLISS-52 and BLISS-76)	Yes	Appendix L
		HDA-1 (pooled across BLISS-52 and BLISS-76)	Yes	Document B Section 2.7
		HDA-2 (pooled across BLISS-52 and BLISS-76)	No	Document B Section 2.7
BLISS 76 US LTE	LTE study of US patients previously enrolled in BLISS-76	Total	No	Document B Section 2.6 with further details in Appendix M
BLISS-52/76 non-US LTE	LTE study of non-US patients previously enrolled in BLISS-52 or BLISS-76	Total	No	Document B Section 2.6 with further details in Appendix M
BLISS-SC	Pivotal trial of SC belimumab	Total	No	Document B Section 2.6
		HDA-1	No	Document B Section 2.7
		HDA-2	No	Document B Section 2.7
BLISS SC LTE	Open-label extension for patients previously enrolled in BLISS-SC	Total	No	Document B Section 2.6 with further details in Appendix M
LBSL02 Phase 2 trial	Initial evidence on safety and efficacy of belimumab	Total	Yes	Appendix L
LBSL02 LTE	Data on long-term (up to 13 years) experience with belimumab	Total	Partially (further data have become available with additional follow-up)	Appendix M
BASE (post-marketing)	Safety study capturing mortality and adverse events of special interest	Total	No	Document B Section 2.10, with steroid use briefly described in Section B 2.6. Methodology in Appendix F
Treatment holiday study (NCT02119156, post-marketing)	A study investigating the effect of treatment holiday on belimumab efficacy	Total	No	Appendix O

Clinical trials				
Trial name	**Description**	**Population (Total/HDA-1/HDA-2)**	**Included in the previous NICE submission? (Yes/No)**	**Location in the current submission**
EMBRACE (post-marketing)	Placebo-controlled trial of belimumab in people of black race	Total	No	Appendix O
NCT01345253	Placebo-controlled trial of belimumab in people from North-East Asia	Total	No	Appendix O
Real-world evidence				
BILAG-BR	UK-based registry of biologic therapy (including belimumab) for SLE	HDA-1 (belimumab data only)	No	Document B Section 2.7 and Appendix P
OBSErve	A multi-country Evaluation Of use of Belimumab in clinical practice Settings	Total	No	Document B Section 2.6
Post-hoc analyses				
SLICC (ACR)/SDI Indirect Cohort Comparison Study (206347)[48]	A PSM comparative analysis between BLISS-76 US LTE and the Toronto Lupus Cohort to assess long-term organ damage in patients treated with belimumab	Total	No	Document B, Section 2.6 and Section B.3.3.6

ACR: American College of Rheumatology; BILAG: British Isles Lupus Assessment Group; BR: Biologics Registry; HDA: high disease activity; IV: intravenous; LTE: long-term extension; NICE: National Institute for Health and Care Excellence; PSM: propensity score-matching; SC: subcutaneous; SDI: SLICC/ACR Damage Index; SLE: systemic lupus erythematosus; SLICC: Systemic Lupus International Collaborating Clinics; UK: United Kingdom; US: United States.

Table 13. New evidence addressing key areas of uncertainty as previously identified by NICE

Area of uncertainty	Studies addressing the area of uncertainty	Key results pertaining to the area of uncertainty	Section of the submission where the results are displayed
1. Treatment benefit across the full range of disease manifestations	The submission presents a large body of trial and real-world evidence collected since the previous NICE appraisal of belimumab. This is presented for the broad population enrolled in belimumab studies, as well as in the HDA-1 and HDA-2 populations. Data are derived from: • BLISS-52 and BLISS-76 (pooled) • BLISS SC In addition, long-term data on treatment benefit across the broad population of patients with active SLE: • LBSL02 Phase 2 trial and its LTE • BLISS-52/76 non-US LTE • BLISS-76 US LTE • BLISS-SC LTE Trial data are supplemented with real-world evidence from: • BILAG-BR (HDA-1 population only) • OBSErve registry series and with the SLICC/SDI Indirect Cohort Comparison Study[48]	• Primary endpoint – SRI-4 response • SRI-4 response components: o 4-point reduction in SELENA-SLEDAI o No worsening (<0.3 point increase) in PGA o No BILAG flare (no new BILAG A organ domain score or 2 new BILAG B organ domain scores) • SFI: o Time to flare o Flares per patient-year o Severe flares per patient-year • Long-term SRI-4 and SFI data	• Section B.2.6: o BLISS-SC o BLISS-76 US LTE o BLISS-SC LTE • Appendix L: o Pooled BLISS-52 and BLISS-76 data o LBSL02 Phase 2 trial • Appendix M: o LBSL02 LTE • Section B.2.7: o BILAG-BR o BLISS-SC HDA-1, HDA-2 populations o BLISS 52/76 HDA-1, HDA-2 populations
2. Development of organ damage	Information on long-term reduction of organ damage in a broad population of belimumab-treated patients: • BLISS-SC LTE • BLISS-52/76 non-US LTE	• SDI: o Mean change from baseline in SDI o Percent of patients with any SDI worsening	• Section B.2.6: o BLISS-52/76 non-US LTE o BLISS-76 US LTE, including the SLICC/SDI Indirect Cohort Comparison Study [48]

Area of uncertainty	Studies addressing the area of uncertainty	Key results pertaining to the area of uncertainty	Section of the submission where the results are displayed

	• BLISS-76 US LTE, including the SLICC/SDI Indirect Cohort Comparison Study [48]	• Other measures of organ damage
		  ○ Organ improvement/worsening by SELENA-SLEDAI
		  ○ Organ improvement/worsening by BILAG
		  ○ Renal flares
			• Section B.2.7
			  ○ BLISS-SC HDA-1, HDA-2 populations
			  ○ BLISS 52/76 HDA-1, HDA-2 populations

3. Extent and impact of the steroid sparing effect	Reduction in steroid use has been assessed in:
	• BLISS-SC
	• BLISS-52
	• BLISS-76
	• BASE
	• OBSErve
	
	This is supplemented with longer-term data from BLISS-SC LTE, BLISS-76 US LTE, and real-world data from the BILAG-BR (HDA-1 population only)
		• Percentage of patients whose average prednisone dose had been reduced by ≥25% from baseline to ≤7.5 mg/day during Weeks 40–52
		• Change in steroid dose over time:
		  ○ Mean/median changes in steroid dose (mg/day), by visit
		• Percentage of patients with daily prednisone dose reduced or increased
			• Section B.2.6:
			  ○ BLISS-SC
			  ○ BLISS-SC LTE
			  ○ BLISS-76 US LTE
			  ○ OBSErve
			• Appendix O:
			  ○ BASE
			• Available upon request:
			  ○ Pooled BLISS-52 and BLISS-76 data
			• Section B.2.7
			  ○ BILAG-BR
			  ○ BLISS-SC HDA-1, HDA-2 populations
			  ○ BLISS 52/76 HDA-1, HDA-2 populations

4. Impact of belimumab on QoL	• BLISS-52
	• BLISS-76
	• BLISS-SC
	• BLISS-76 US LTE
	• BLISS-SC LTE
	• BILAG-BR (HDA-1 population only)
		• FACIT Fatigue Score
		• Mean change in SF-36 Health Survey PCS score
		• EQ-5D
			• Section B.2.6:
			  ○ BLISS-SC
			  ○ BLISS-76 US LTE
			  ○ BLISS-SC LTE
			• Available upon request:
			  ○ Pooled BLISS-52 and BLISS-76 data
			• Section B.2.7:
			  ○ BILAG-BR
			  ○ BLISS-SC HDA-1, HDA-2 populations
			  ○ BLISS 52/76 HDA-1, HDA-2 populations

5. Length of treatment/extent of	• LTE studies: LBSL02, BLISS-SC, BLISS-52/76 non-US, BLISS-76	• Treatment discontinuation rates
		• Reason for discontinuation
			• Available upon request:
			  ○ BLISS-SC LTE
			  ○ BLISS-52/76 non-US LTE

Area of uncertainty	Studies addressing the area of uncertainty	Key results pertaining to the area of uncertainty	Section of the submission where the results are displayed
discontinuations over time	US LTEs, and integrated analysis • RWE studies: OBSErve, BILAG-BR (HDA-1 population only)		○ BLISS-76 US LTEs ○ Integrated analysis of LTE studies • Appendix N: ○ OBSErve • Appendix M: ○ LBSL02 LTE • Section B.2.7: ○ BILAG-BR
6. Type of standard of care in UK clinical practice	• BILAG-BR (HDA-1 population only)	• Treatments used in the control (non-biologic) group in the BILAG-BR. Concomitant medications used in the BILAG-BR	• Appendix P: ○ BILAG-BR

BILAG: British Isles Lupus Assessment Group; BR: Biologics Registry; FACIT: Functional Assessment of Chronic Illness Therapy; HDA: high disease activity; LTE: long-term extension; NICE: National Institute for Health and Care Excellence; PCS: Physical Component Score; PGA: Physician's Global Assessment; PSM: propensity score-matching; QoL: quality of life; SC: subcutaneous; SDI: Systemic Lupus International Collaborating Clinics (SLICC)/ACR Damage Index; SELENA: Safety of Estrogen in Lupus Erythematosus National Assessment; SFI: SLE Flare Index; SLE: systemic lupus erythematosus; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index; SRI-4: SLE responder index-4; US: United States.

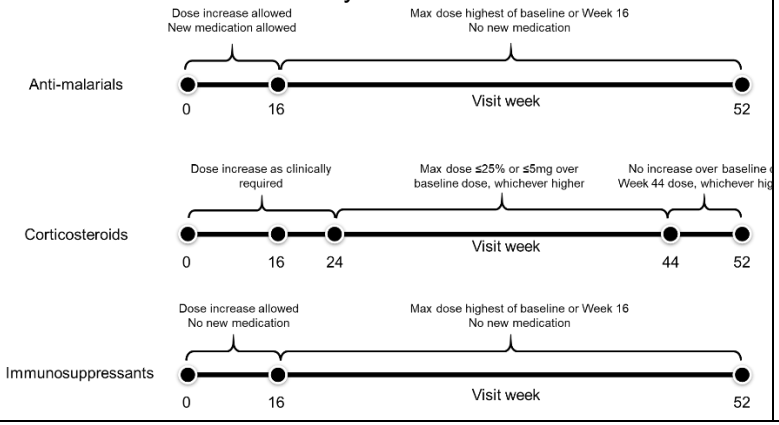
B.2.3 *Summary of methodology of the relevant clinical effectiveness evidence*

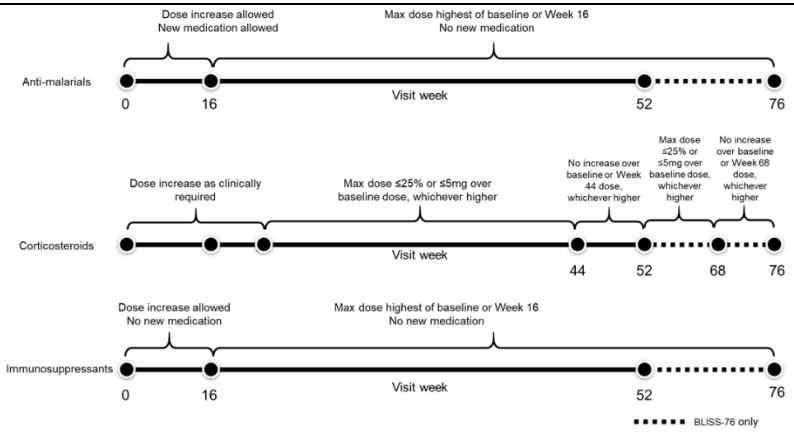
B.2.3.1 Pivotal RCTs of belimumab

Belimumab as an add-on to ST was investigated in three RCTs. BLISS-52 and BLISS-76 investigated the IV formulation of belimumab, presented to NICE as part of TA397, while BLISS-SC, the pivotal trial for the SC formulation, is newly introduced herein. The methodology of these three trials is summarised in Table 14, while comprehensive details are provided in Appendix L.

Table 14. Comparative summary of the methodology of pivotal belimumab RCTs

Trial acronym	BLISS-SC[45, 67]	BLISS-52[46] and BLISS-76[47]
Trial design	Phase 3, multicentre, international, randomised, double-blind, placebo-controlled, 52-week study.	Phase 3, randomised, multicentre, international, double-blind, placebo-controlled, parallel-group study. BLISS-52 was 52 weeks in duration, while BLISS-76 was 76 weeks in duration.
Eligibility criteria for participants	Patients with a clinical diagnosis of SLE according to the ACR criteria and clinically active SLE disease, defined as a SELENA-SLEDAI disease activity score of ≥8 at screening. Patients with severe lupus kidney disease, severe active lupus nephritis or CNS lupus were excluded.	Patients with a clinical diagnosis of SLE according to the ACR criteria and clinically active SLE disease, defined as a SELENA-SLEDAI disease activity score of ≥6 at screening. Patients with severe active lupus nephritis or CNS lupus were excluded.
Settings and locations where the data were collected	30 countries in North America, Central America, South America, Western Europe, and. Eastern Europe. The trial was run in the UK (6 patients enrolled in 3 centres). Locations were hospital settings, academic institutions (i.e. University hospitals), medical centres, rheumatology departments.	BLISS-52: 13 countries in Latin America, Asia-Pacific and eastern Europe BLISS-76: 19 countries in North America and Europe (including the UK). Locations were hospital settings, academic institutions (i.e. University hospitals), medical centres, rheumatology departments.
Intervention	Belimumab 200 mg administered by SC injection on Day 0 and then weekly (i.e., every 7 days ± 1 day) through 51 weeks, plus ST (N=556).	BLISS-52: Belimumab 1 mg/kg (N=288) or belimumab 10 mg/kg (N=290) administered by IV infusion on Days 0, 14, 28 and every 28 days thereafter plus ST BLISS-76: belimumab 1 mg/kg (N=271) or belimumab 10 mg/kg (N=273) administered by IV infusion on Days 0, 14, 28 and every 28 days thereafter plus ST.
Comparator	Placebo administered by SC injection on Day 0 and then weekly (i.e., every 7 days ± 1 day) through 51 weeks, plus ST (N=280).	BLISS-52: Placebo (N=287) administered by IV infusion on Days 0, 14, 28 and every 28 days thereafter plus ST BLISS-76: Placebo (N=275) administered by IV infusion. on Days 0, 14, 28 and every 28 days thereafter plus ST.

Trial acronym	BLISS-SC[45, 67]	BLISS-52[46] and BLISS-76[47]
Permitted and disallowed concomitant medications	Permitted medications: 1) Anti-malarials 2) Corticosteroids 3) Other immunosuppressive/immunomodulatory agents 4) NSAIDs and aspirin Disallowed medications: 1) Other investigational agents (biologic or non-biologic). 2) Co-enrolment into another study of a different investigational agent, or that could interfere with the conduct of the BLISS-SC study protocol. 3) Anti-TNF therapy within 90 days of Day 0. 4) Other biologics. 5) IVIG. 6) IV cyclophosphamide within 90 days of Day 0. 7) Plasmapheresis. 8) Live vaccines The following were specific exclusion factors: treatment with any B cell targeted therapy at any time, receipt of abatacept or a biologic investigational agent other than B cell targeted therapy within 364 days of Day 0. Progressive restrictions were placed on concomitant medication use over the course of the study:  	Permitted medications: 1) Anti-malarials 2) Corticosteroids 3) Other immunosuppressive/immunomodulatory agents 4) NSAIDs and aspirin 5) Statins 6) ACEis or ARBs Disallowed medications: 1) Other investigational agents (biologic or non-biologic). 2) Co-enrolment into another study of a different investigational agent, or that could interfere with the conduct of the BLISS-52/76 study protocol. 3) Anti-TNF therapy within 90 days of Day 0. 4) Other biologics. 5) IVIG. 6) IV cyclophosphamide within 180 days of Day 0 7) Plasmapheresis. The following were specific exclusion factors: treatment with any B cell targeted therapy at any time, receipt of abatacept or a biologic investigational agent other than B cell targeted therapy within 364 days of Day 0. Progressive restrictions were placed on concomitant medication use over the course of the studies:

Trial acronym	BLISS-SC[45, 67]	BLISS-52[46] and BLISS-76[47]
		
Primary outcome	The primary efficacy endpoint was SRI-4 response rate at Week 52. SRI-4 response was defined as: • ≥4-point reduction from baseline in SELENA-SLEDAI score, AND: • No worsening (increase of <0.30 points from baseline) in PGA, AND: • No new BILAG A organ domain score or 2 new BILAG B organ domain scores compared with baseline, at the time of assessment (i.e., at Week 52).	
Other outcomes used in the economic model/specified in the scope	• Disease activity: Change in PGA and SELENA-SLEDAI score • Rate and duration of response: SRI-4 response by visit, and at Week 52 (primary efficacy endpoint) • Incidence and severity of flares: Time to SFI flare/severe flare and rate of SFI flare/severe flare per 100 subject years • Incidence of long-term complications and/or organ damage: SELENA-SLEDAI and BILAG, scores by visit; SDI change at Week 52 • Corticosteroid use: Mean/median changes in steroid dose by visit; percent of patients whose average prednisone use reduced by ≥25% to ≤7.5 mg/day • Mortality: not assessed as an outcome, although included in safety reporting • HRQoL: FACIT-Fatigue Scale at Week 52 and by visit.	• Disease activity: Change in PGA and SELENA-SLEDAI score • Rate and duration of response: SRI-4 response by visit, and at Week 52 (primary efficacy endpoint) • Incidence and severity of flares: Time to SFI flare, Time to first flare, number and rate of flares • Incidence of long-term complications and/or organ damage: Change in SELENA-SLEDAI, BILAG scores, and change in SDI at Week 52 • Corticosteroid use: Percent of patients whose average prednisone use reduced by ≥25% to ≤7.5 mg/day • Mortality: not assessed as an outcome, although included in safety reporting • HRQoL: FACIT-Fatigue Scale, SF-36, and EQ-5D at Week 52.

Trial acronym	BLISS-SC[45, 67]	BLISS-52[46] and BLISS-76[47]
	Adverse effects of treatment: monitoring of adverse events, clinical laboratory tests, vital signs, physical examinations, and immunogenicity.	Adverse effects of treatment: monitoring of adverse events, clinical laboratory tests, vital signs, physical examinations, and immunogenicity • For BLISS-76, SRI-4 response rate at Week 76.
Pre-planned subgroups	Pre-planned subgroup analyses for the primary endpoint (SRI-4) were performed in the following subgroups: • Baseline SELENA-SLEDAI score (≤9 vs ≥10). • Race (White, American Indian, Asian, and Black). • Baseline anti-dsDNA (≥30 IU/mL vs <30 IU/mL). • Baseline prednisone dose level (≤7.5 mg/day vs >7.5 mg/day). • Baseline complement levels (C3 and/or C4 low vs other). • Baseline complement and anti-dsDNA (C3 and/or C4 low AND anti-dsDNA ≥30 vs other). • Region (US/Canada, Europe/Australia/Israel, Asia/Americas excluding US and Canada).	For both BLISS-52 and BLISS-76, pre-planned subgroup analyses for the primary endpoint (SRI-4) were performed in the following subgroups: • Baseline SELENA-SLEDAI score (≤9 vs ≥10). • Race (African descent or indigenous-American descent vs other). • Baseline anti-dsDNA (≥30 IU/mL vs <30 IU/mL). • Baseline prednisone dose level (≤7.5 mg/day vs >7.5 mg/day). • Baseline proteinuria level (<2 g/24-hour vs ≥2 g/24 hour equivalent). • Baseline C3 levels (normal/high vs low). • Baseline C4 levels (normal/high vs low). In addition, BLISS-76 included a pre-planned subgroup analysis by country region (North America vs Central and South America vs Europe). As there were no patients in South America, the regions analysed were US/Canada, Americas excluding US/Canada, and Europe, divided into Western and Eastern Europe.
Key post-hoc subgroups	Key post-hoc subgroups in which treatment with belimumab is likely to provide particular benefit, described in Section B.2.7: • HDA-1 population (belimumab: N=186, placebo: N=78) • HDA-2 population (belimumab: N=296, placebo: N=141)	Key post-hoc subgroups in which treatment with belimumab is likely to provide particular benefit, described in Section B.2.7: • Pooled HDA-1 population (belimumab: N=193, placebo: N=203) • Pooled HDA-2 population (belimumab: N=262, placebo: N=270)

ACEi: Angiotensin-converting enzyme inhibitor; ACR: American College of Rheumatology; ARB: angiotensin II receptor blocker; BILAG: British Isles Lupus Assessment Group; CNS: central nervous system; dsDNA: double-stranded deoxyribonucleic acid; FACIT: Functional Assessment of Chronic Illness Therapy; HDA: high disease activity; HRQoL: health-related quality of life; IV: intravenous; IVIG: intravenous immunoglobulin; NSAID: non-steroidal anti-inflammatory drug; PGA: Physician's Global Assessment; SC: subcutaneous; SDI: Systemic Lupus International Collaborating Clinics (SLICC)/ACR Damage Index; SELENA: Safety of Estrogen in Lupus Erythematosus National Assessment; SFI: SLE Flare Index; SLE: Systemic Lupus Erythematosus; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index; SRI-4: SLE responder index-4; TNF: tumour necrosis factor; UK: United Kingdom; US: United States.

B.2.3.1.1 Baseline characteristics of patients included in pivotal RCTs of belimumab

This section briefly characterises the populations included in BLISS-SC, BLISS-52 and BLISS-76 (Table 15). Please note that whilst a 1 mg/kg IV belimumab dose was assessed in the Phase 3 BLISS-52 and BLISS-76 studies, we will only present results for the 10 mg/kg belimumab dose in this submission, as this is the dose approved for Marketing Authorisation.

Table 15. Baseline characteristics of participants included in pivotal belimumab trials

	BLISS-SC[45]		**Pooled BLISS-52 and BLISS-76 data[46, 47]**	
	Belimumab 200 mg SC N=556	**Placebo N=280**	**Belimumab 10 mg/kg IV N=563**	**Placebo N=562**
Demographics				
Female, N (%)	521 (93.7)	268 (95.7)	539 (95.7)	522 (92.9)
Age (years), mean (SD)	38.1 (12.10)	39.6 (12.61)	37.9 (11.3)	38.1 (12.0)
≤45 years, N (%)	403 (72.5)	193 (68.9)	414 (73.5)	414 (73.7)
Race, N (%)				
White	336 (60.4)	166 (59.3)	260 (46.2)	270 (48.0)
Asian	119 (21.4)	63 (22.5)	127 (22.6)	116 (20.6)
African American/African Heritage	56 (10.1)	30 (10.7)	50 (8.9)	50 (8.9)
American Indian or Alaska Native	43 (7.7)	21 (7.5)	126 (22.4)	125 (22.2)
Native Hawaiian or Other Pacific Islander	2 (0.4)	0	0	1 (0.2)
Multiracial[a]	6 (1.1)	3 (1.1)	4 (0.7)	3 (0.5)
Ethnicity: Hispanic or Latino origin, N (%)	160 (28.8)	80 (28.6)	192 (34.1)	198 (35.2)
Disease characteristics				
SLE disease duration (years), mean (SD)	6.4 (6.60)	6.8 (6.83)	6.08 (6.42)	6.66 (6.48)
BILAG organ domain involvement, N (%)				
At least 1A or 2B	388 (69.8)	210 (75.0)	332 (59.0)	353 (62.8)
At least 1A	87 (15.6)	51 (18.2)	78 (13.9)	89 (15.8)
At least 1B	499 (89.7)	258 (92.1)	509 (90.4)	517 (92.0)
No A or B	29 (5.2)	13 (4.6)	54 (9.6)	45 (8.0)
SELENA-SLEDAI category, N (%)				

	BLISS-SC[45]		Pooled BLISS-52 and BLISS-76 data[46, 47]	
	Belimumab 200 mg SC N=556	**Placebo N=280**	**Belimumab 10 mg/kg IV N=563**	**Placebo N=562**
0–3	4 (0.7)	0 (0.0)	11 (2.0)	4 (0.7)
4-9			256 (45.5)	259 (46.1)
≤9	200 (36.0)	112 (40.0)		
10–11	161 (29.0)	74 (26.4)	137 (24.3)	137 (24.4)
≥10	352 (63.3)	168 (60.0)	296 (52.6)	299 (53.2)
≥12	191 (34.4)	94 (33.6)	159 (28.2)	162 (28.8)
SELENA-SLEDAI score, mean (SD)	10.5 (3.19)	10.3 (3.04)	9.75 (3.77)	9.75 (3.79)
SFI, N (%)				
At least 1 flare	92 (16.5)	57 (20.4)	115 (20.4)	139 (24.7)
At least 1 severe flare	8 (1.4)	4 (1.4)		
Severe flare			8 (1.4)	4 (0.7)
PGA Category, N (%)				
0–1	40 (7.2)	19 (6.8)		
<1			83 (14.7)	76 (13.5)
1–<2			387 (68.7)	391 (69.6)
≥2			93 (16.5)	95 (16.9)
>1–2.5	507 (91.2)	255 (91.1)		
>2.5	7 (1.3)	5 (1.8)		
Missing	2 (0.4)	1 (0.4)		
PGA, N Mean (SD)	554 1.6 (0.43)	279 1.5 (0.45)		
SDI score, mean (SD)	0.6 (0.99)	0.7 (1.17)	0.74 (1.21)	0.77 (1.23)
SDI score =0, N (%)			338 (60.0)	327 (58.2)
SDI score =1, N (%)			122 (21.7)	136 (24.2)
SDI score ≥2, N (%)			103 (18.3)	99 (17.6)
Proteinuria category (g/24 h), N (%)				
≥2	19 (3.4)	20 (7.1)	34 (6.0)	32 (5.7)
Proteinuria level (g/24 h), mean (SD)	0.4 (0.71)	0.4 (0.84)	0.48 (0.83)	0.50 (1.00)
Medication usage				
Average daily prednisone dose, mg/day, N (%)				
>0–≤7.5	146 (26.3)	73 (26.1)	154 (27.4)	170 (30.2)
>7.5	335 (60.3)	168 (60.0)	324 (57.6)	318 (56.6)
Average daily prednisone dose (mg/day), mean (SD)	10.8 (8.21)	11.2 (9.09)	10.9 (9.1)	10.7 (8.5)
Number (%) of patients taking:				
Steroids	481 (86.5)	241 (86.1)	478 (84.9)	488 (86.8)
Antimalarials	391 (70.3)	189 (67.5)	353 (62.7)	381 (67.8)

	BLISS-SC[45]		Pooled BLISS-52 and BLISS-76 data[46, 47]	
	Belimumab 200 mg SC N=556	Placebo N=280	Belimumab 10 mg/kg IV N=563	Placebo N=562
Immunosuppressants	244 (43.9)	137 (48.9)	271 (48.1)	276 (49.1)
Aspirin	94 (16.9)	45 (16.1)		
Aspirin >1000mg/day			0	1 (0.2)
NSAIDs	124 (22.3)	72 (25.7)	159 (28.2)	178 (31.7)

[a]Patients who checked more than 1 race category are counted under individual race category according to the minority rule as well as the multiracial category.
Greyed boxes indicate that the category was not measured within the trial.
BILAG: British Isles Lupus Assessment Group; ITT: intention-to-treat; IV: intravenous; NSAID: non-steroidal anti-inflammatory drug; PGA: Physician's Global Assessment; SC: subcutaneous; SD: standard deviation; SDI: Systemic Lupus International Collaborating Clinics (SLICC)/ACR Damage Index; SELENA: Safety of Estrogen in Lupus Erythematosus National Assessment; SFI: SLE Flare Index; SLE: Systemic Lupus Erythematosus; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index.

B.2.3.2 Long-term extensions of pivotal RCTs

The long-term safety and efficacy of belimumab as an add-on to ST treatment was investigated in three LTE studies. BLISS-SC LTE continued to monitor the safety and efficacy of the SC formulation of belimumab in patients who had participated in the BLISS SC trial, whilst BLISS-52/76 non-US LTE and BLISS-76 US LTE monitored the safety and efficacy of the IV formulation in patients who had participated in BLISS-52 and BLISS-76 outside the US; and BLISS-76 within the US, respectively. The methodology of these three trials is summarised in Table 16, while comprehensive details are provided in Appendix M.

Table 16. Comparative summary of the methodology of belimumab LTEs

Trial acronym	BLISS-SC LTE[49]	BLISS-52/76 non-US LTE[50]	BLISS-76 US LTE[51]
Trial design	6-month open-label extension phase to the BLISS-SC pivotal Phase 3 trial.	Multicentre, continuation trial of belimumab IV in SLE patients who completed the Phase 3 BLISS-52 or BLISS-76 trials[a].	Multicentre continuation trial of belimumab IV in SLE patients who completed the Phase 3 BLISS-76 study in the US.
Eligibility criteria for participants	Completion of the double-blind phase of the BLISS-SC trial.	Non-US patients who had completed the Phase 3 BLISS-52 or BLISS-76 trials.	US patients who had completed the Week 72 visit of the Phase 3 BLISS-76 trial.
Settings and locations where the data were collected	24.9% of patients from the US, 24.2% from Eastern Europe, 21.6% from Asia, 21.1% from the Americas (excluding the US and Canada), and 8.2% from Western Europe.	Patients from 28 countries in Americas, excluding US and Canada (43.0%), Asia (28.6%), Eastern Europe (12.4%), Canada (1.5%) and Western Europe/Australia/Israel (14.6%).	Patients from the US (100%).
Intervention	Belimumab 200 mg administered by SC injection weekly for 6 months (N=662) plus ST.	Belimumab IV 10 mg/kg every 28 days[b] plus ST. Patients could continue receiving belimumab treatment in this trial until it became commercially available in their country (N=735).	Belimumab IV 10 mg/kg every 28 days[b] plus ST (N=268).
Comparator	None.	None.	None.
Permitted and disallowed concomitant medications	Permitted medications: The investigator could adjust concurrent medications (add, eliminate, change dose level/frequency) as clinically appropriate (see BLISS SC Permitted medications, Table 14). Disallowed medications: As per the double-blind phase of the BLISS-SC study (Table 14).	Permitted medications: The investigator could adjust concurrent medications (add, eliminate, change dose level/frequency) as clinically appropriate (see BLISS 52 and BLISS 76 Permitted medications Table 14). Disallowed medications 1) Other investigational agents or participation in another study 2) Anti-TNF therapy 3) Other biologics 4) IV cyclophosphamide.	Permitted medications: The investigator could adjust concurrent medications (add, eliminate, change dose level/frequency) as clinically appropriate (see BLISS 52 and BLISS 76 Permitted medications Table 14). Disallowed medications 1) Other investigational agents or participation in another study 2) Anti-TNF therapy 3) Other biologics 4) IV cyclophosphamide.
Efficacy outcomes	Efficacy data were collected at the end of the LTE phase (Week 24). The primary endpoint was the SRI-4	The protocol-specified efficacy endpoint in this study was SDI, to assess irreversible organ damage as a measure of disease	The primary efficacy endpoint was the SRI-4 response rate at each belimumab visit (see Primary Efficacy Endpoint, Table 14) for

Trial acronym	BLISS-SC LTE[49]	BLISS-52/76 non-US LTE[50]	BLISS-76 US LTE[51]
	response rate at Week 52, for which data was collected during the double-blind phase (see Primary Efficacy Endpoint, Table 14). Other efficacy endpoints were as per the double-blind phase.	activity. SDI can also be considered a safety endpoint.	definition).Other efficacy assessments included: SELENA-SLEDAI, BILAG, PGA, SFI, and prednisone use.
Other outcomes used in the economic model/specified in the scope	As per BLISS-SC (Table 14).	• Safety: adverse event monitoring and laboratory tests.	• Efficacy/Safety: adverse event monitoring, laboratory tests, and SDI.
Pre-planned subgroups	Response with belimumab 200 mg SC was evaluated by race classification of Black versus Other for patients in the open-label phase of the study.	Pre-defined subgroups were: • Baseline SELENA-SLEDAI score (≤9 and ≥10) • Age (<65 years and ≥65 years) • Sex (Male and Female) • Race (White, American Indian or Alaska Native, Asian, Black or African American) • Race Stratification (Black and Other).	Pre-defined subgroups were: • Baseline SELENA-SLEDAI score (≤9 and ≥10) • Age (<65 years and ≥65 years) • Sex (Male and Female) • Race (White, American Indian or Alaska Native, Asian, Black or African American) • Race Stratification (Black and Other).

[a]5 patients from Mexico who were still on treatment with belimumab SC in BEL112232 (NCT00732940) at the time this study was terminated were permitted to enrol, allowing them to continue treatment with belimumab. [b]Patients who received 1 mg/kg belimumab IV in the parent studies received the same dose in the LTE study until marketing approval was obtained for 10 mg/kg belimumab IV, at which time their dose was increased to 10 mg/kg.

BILAG: British Isles Lupus Assessment Group; IV: intravenous; LTE: long-term extension; PGA: Physician's Global Assessment; SC: subcutaneous; SDI: Systemic Lupus International Collaborating Clinics (SLICC)/ACR Damage Index; SELENA: Safety of Estrogen in Lupus Erythematosus National Assessment; SFI: SLE Flare Index; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index; SRI-4: SLE responder index-4; ST: standard therapy; TNF: Tumour necrosis factor; US: United States.

B.2.3.2.1 Baseline characteristics of patients included in LTE studies of belimumab

Baseline characteristics of patients included in LTE studies of belimumab are presented in Table 17. Note that for all three LTE studies, baseline was defined as the last available value prior to the initiation of treatment with belimumab[49-51]. Therefore, it occurred at different time points for patients who were randomised to placebo compared to patients who were randomised to belimumab in the parent study [49-51]. Parent study baseline was used for patients originally randomised to belimumab, while the last available value from the parent study was used for patients originally randomised to placebo[49-51].

Table 17. Baseline characteristics of participants included in LTEs of pivotal belimumab trials

	BLISS-SC LTE[49]	**BLISS-52/76 non-US LTE[50]**	**BLISS-76 US LTE[51]**
	Belimumab 200 mg SC N=662	**Belimumab 10 mg/kg IV N=735**	**Belimumab 10 mg/kg IV N=268**
Female, N (%)	626 (94.6)	695 (94.6)	250 (93.3)
Age (years), mean (SD)	38.7 (11.86)	37.2 (11.17)	42.8 (11.33)
≤45 years, N (%)	473 (71.5)	560 (76.2)	162 (60.4)
Race, N (%)			
White/Caucasian	403 (60.9)	278 (37.8)	186 (69.4)
Asian	147 (22.2)	214 (29.1)	13 (4.9)
African American/African Heritage	56 (8.5)	18 (2.4)	57 (21.3)
American Indian or Alaska Native	55 (8.3)	225 (30.6)	8 (3.0)
Native Hawaiian or Other Pacific Islander	1 (0.2)	0	0
Multiracial	8 (1.2)	2 (0.3)	4 (1.5)
Ethnicity, N (%)			
Hispanic or Latino	194 (29.3)	315 (42.9)	52 (19.4)
SLE disease duration (years), mean (SD)	6.7 (6.58)	6.3 (5.99)	7.7 (6.77)
BILAG organ domain involvement, N (%)			
At least 1A or 2B	356 (53.8)	324 (44.1)	137 (51.1)
At least 1A	74 (11.2)	107 (14.6)	20 (7.5)
At least 1B	528 (79.8)	531 (72.2)	204 (76.1)
No A or B	105 (15.9)	171 (23.3)	56 (20.9)

	BLISS-SC LTE[49]	BLISS-52/76 non-US LTE[50]	BLISS-76 US LTE[51]
	Belimumab 200 mg SC N=662	**Belimumab 10 mg/kg IV N=735**	**Belimumab 10 mg/kg IV N=268**
SELENA-SLEDAI category, N (%)			
0–3	57 (8.6)[a]		
≤9	340 (51.4)	446 (60.7)	188 (70.1)
10–11	147 (22.2)		
≥10		284 (38.6)	80 (29.9)
≥12	175 (26.4)		
Missing	0	5 (0.7)	0
SELENA-SLEDAI score, mean (SD)	9.0 (4.03)	8.3 (4.29)	7.8 (3.86)
SFI, N (%)			
At least 1 flare	72 (10.9)	107 (14.6)	65 (24.3)
At least 1 severe flare	6 (0.9)	4 (0.5)	2 (0.7)
PGA Category, N (%)			
0–1	151 (22.8)	231 (31.4)	79 (29.5)
>1–2.5	507 (76.6)	499 (67.9)	188 (70.1)
>2.5	3 (0.5)	5 (0.7)	1 (0.4)
Missing	1	0	0
PGA, mean (SD)	1.32 (0.597) [note N=661]	1.19 (0.60)	1.2 (0.60)
SDI score, mean (SD)	0.6 (1.06)	0.6 (1.02)	1.2 (1.51)
Proteinuria category (g/24 h), N (%)			
≥2	20 (3.0)	42 (5.7)	7 (2.6)
Proteinuria level (g/24 h), mean (SD)	0.36 (0.718) [note N=658]	0.5 (0.94)	0.3 (0.56)

[a]2 patients in the belimumab 200 mg to belimumab 200 mg group and 55 patients in the placebo to belimumab 200 mg group had a baseline SELENA-SLEDAI score <4.
Greyed boxes indicate that the category was not measured within the trial.
BILAG: British Isles Lupus Assessment Group; IV: intravenous; LTE: long-term extension; PGA: Physician's Global Assessment; SC: subcutaneous; SD: standard deviation; SDI: Systemic Lupus International Collaborating Clinics (SLICC)/ACR Damage Index; SELENA: Safety of Estrogen in Lupus Erythematosus National Assessment; SFI: SLE Flare Index; SLE: systemic lupus erythematosus; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index; US: United States.

B.2.3.3 SLICC(ACR)/SDI Indirect Cohort Study

A limitation of the LTE studies is the lack of comparator arm which precluded a direct comparison of belimumab plus ST with ST alone in these studies. Consequently, the question of long-term relative efficacy required further investigation. A propensity score matching (PSM) analysis was conducted which matched patients treated with belimumab (plus ST) in the BLISS-76 US LTE study (primary analysis) with patients

from an external SLE cohort treated with ST, to enable a long-term comparative analysis of belimumab versus ST[48]. An SLR was performed to identify cohorts, registries or other databases that supported SLE. The objective was to identify a comparison cohort with population characteristics similar to the BLISS trial population with adequate sample size with complete clinical data and at least five years follow-up. In total 92 cohorts were identified of which 21 cohorts had at least 400 patients and from which data was extracted. Evaluation criteria included cohort size, ethnicity, age, duration of SLE, severity of disease activity, extent of organ damage, follow-up and scope of data collection and data availability. The Toronto Lupus Cohort (TLC)[26, 35] was selected as the preferred source of ST data for this post-hoc longitudinal PS-matched study, based on the size of the cohort, the extent of organ damage among the patients and the severity of their disease activity within the cohort[48]. Moreover the scales for disease activity, organ damage progression and health-related quality of life were compatible with those from the BLISS studies.

A SLR was used to identify publications that reported predictors of SLE organ damage and progression. Key predictors found in the literature were reviewed by a clinical expert and limited to those available in both the BLISS LTE studies and the TLC. This generated a list of 14 predictors, which correlated to 17 operationalised variables used in the primary PSM analysis of the BLISS US LTE/TLC datasets[48].

The primary objective was to compare organ damage progression (SDI score) from baseline to Year 5 in patients treated with belimumab (plus ST) or ST alone, using PS-matched data from the BLISS-76 US LTE study and the TLC external cohort. Secondary objectives included comparing the time to organ damage progression and the magnitude of damage accrual. The time to organ damage progression analysis included all patients with >1 year of follow-up and excluded TLC patients with ≥15 years of follow-up[48]. Further methodology information is presented in Appendix M.[48]

B.2.3.4 Key real-world evidence

The long-term safety and efficacy of belimumab in the real-world setting was investigated in two RWE studies. The BILAG-BR Benlysta Sub-Study was an

observational cohort study of UK patients who were prescribed belimumab in clinical practice in accordance with the NICE guidance for belimumab (TA397), and managed in their usual clinical setting, and the OBSErve (evaluation Of use of Belimumab in clinical practice SEttings) registry series was a patient-level meta-analysis of retrospective multicentre observational cohort studies, conducted in 6 countries: Argentina, Canada, Germany, Spain, Switzerland, and the USA.

The methodology of these two RWE studies is summarised in Table 18 with further details provided in Appendix N.

The full report from the BILAG-BR Benlysta Sub-Study is available in Appendix P and provides additional methodological details of the analyses performed to address study objectives, as well as the full results of this study. Baseline characteristics of patients included in OBSErve and BILAG-BR Benlysta Sub-Study are listed in Table 19.

Table 18. Comparative summary of the methodology of belimumab RWE studies

Trial name	**BILAG-BR: BENLYSTA Sub Study[53]**	**OBSErve registry series[68]**
Trial design	Prospective, multicentre, non-randomised, observational, registry study of patients prescribed belimumab in clinical practice in accordance with the NICE guidance (TA397).	Patient-level meta-analysis of retrospective, multicentre, observational, exploratory cohort studies.
Trial duration	Patient characteristics, confounders and disease severity were recorded at baseline before treatment was instigated. On treatment, patients were followed up after 3, 6 and 12 months during the first year and every 12 months thereafter to collect data on health outcomes including disease severity and quality of life.	2 years in Argentina and US, 6 months in Canada, Germany, Spain and Switzerland.
Eligibility criteria for participants	Patients aged ≥5 years commencing treatment with belimumab IV for their SLE at the clinical decision of their treating consultant.	Adults (≥18 years, with a clinical diagnosis of SLE) who had initiated belimumab IV as part of their usual SLE care ≥6 months prior to enrolment and for whom reasons for belimumab initiation could be identified.
Settings and locations where the data were collected	Data presented from hospitals throughout England	Clinics in 6 countries: Argentina, Canada, Germany, Spain, Switzerland, and the US. 830 patients were included in the pooled analysis.
Intervention	Benlysta cohort: Defined as patients who are anti-dsDNA positive and have either a low complement 3 or 4 level (defined by the centre's own criteria for each biomarker) and SLEDAI-2K score ≥10 prior to starting belimumab IV.	Non-interventional study. Belimumab IV was prescribed by the treating physician as part of usual care.
Comparator	Comparator cohorts: • Rituximab cohort: Defined as patients who are anti-dsDNA positive and have either a low complement 3 or 4 level (defined by the centre's own criteria for each biomarker) and	No comparator.

Trial name	BILAG-BR: BENLYSTA Sub Study[53]	OBSErve registry series[68]
	SLEDAI-2K score ≥10 prior to starting rituximab. In addition, excludes patients with a BILAG A in either renal or CNS domains. Only rituximab patients recruited since the NHS England interim rituximab SLE policy was published in October 2013 were included in the rituximab cohort. • Non-Biologic Therapies Cohort: Defined as patients who are anti-dsDNA positive and have either a low complement 3 or 4 level (defined by the centre's own criteria for each biomarker) and SLEDAI-2K score ≥10 prior to starting non-biologic therapy. In addition, excludes patients with a BILAG A in either renal or CNS domains.	
Permitted and disallowed concomitant medications	Patients were treated according to clinical need and according to the physician's decision in their usual clinical setting.	
Efficacy outcomes	• Overview of Disease Activity Analyses: SLEDAI-2K; BILAG index; SDI • HRQoL analyses: LupusQoL; SF-36v2; EQ-5D • Steroid Use • Time to Withdrawal • SLE Manifestations: hospitalisations due to SLE • Safety Analyses: serious adverse events; adverse events of special interest.	The primary endpoint was physician-assessed overall clinical response to belimumab therapy at 6 months. Secondary objectives were to explore: • Demographic, clinical and treatment characteristics • Changes in steroid use • Physician-assessed clinical response in patients with HDA • Change in SLEDAI score • Reasons for belimumab initiation and discontinuation.
Other outcomes used in the economic model/specified in the scope	NA	NA
Pre-planned subgroups	None	

BILAG: British Isles Lupus Assessment Group; CNS: central nervous system; dsDNA: double-stranded deoxyribonucleic acid; HDA: high disease activity; HRQoL: health-related quality of life; IV: intravenous; NA: not applicable; NICE: National Institute for Health and Care Excellence; SDI: Systemic Lupus International Collaborating Clinics (SLICC)/ACR Damage Index; SLE: systemic lupus erythematosus; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index; UK: United Kingdom; US: United States.

Table 19. Baseline characteristics from the BILAG-BR and OBSErve studies

	BILAG-BR (belimumab-treated patients only) N=86*	OBSErve (pooled analysis of data) N=830
Female, N (%)	[redacted]	741 (89.3)
Age (years), mean (SD)	[redacted]	41.9 (12.6)
Weight (kg) mean (SD)	[redacted]	NR
Missing (n)	[redacted]	NR
Race, N (%)		
White/Caucasian	[redacted]	540 (65.1)
Asian	[redacted]	36 (4.3)
African American/African Origin/West Indian/Black	[redacted]	134 (16.1)
American Indian/Native American	[redacted]	5 (0.6)
Hispanic	[redacted]	91 (11.0)
Mixed	[redacted]	17 (2.1)
Other	[redacted]	7 (0.8)
Missing (n)	[redacted]	-
SLE disease duration (years), n/N (%)		
≤5 years	[redacted]	377/828 (45.5)
≥6 years	[redacted]	451/828 (54.5)
SLE disease duration (years), mean (SD)	[redacted]	NR
Missing	[redacted]	-
SLE severity at baseline, n/N (%)		
Mild	[redacted]	58/822 (7.1)
Moderate	[redacted]	593/822 (72.1)
Severe	[redacted]	171/822 (20.8)
Low complement and/or high dsDNA, n/N (%)	[redacted]	681/822 (82.8)
SELENA-SLEDAI at baseline, n/N (%)		
<10	[redacted]	138/345 (40.0)
≥10	[redacted]	207/345 (60.0)
Occupational status, n/N (%)		
Full-time	[redacted]	NR
Part-time	[redacted]	NR
Sickness/Disability	[redacted]	NR
Unemployed/student	[redacted]	NR
Retired	[redacted]	NR
Missing	[redacted]	NR

**at baseline of any treatment round. See Section B.2.7.4.1 for further details.*

***only patients meeting HDA-1 population criteria were included in the analysis*

BILAG: British Isles Lupus Assessment Group; BR: Biologics Registry; SD: standard deviation; SELENA: Safety of Estrogen in Lupus Erythematosus National Assessment; SLE: systemic lupus erythematosus; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index.

In comparison with the OBSErve study series, the population included in the BILAG-registry was more uniform in disease severity, as all patients met the HDA-1 population criteria of low complement, positive anti-dsDNA and SELENA-SLEDAI score ≥10. The BILAG-BR registry also included a larger proportion of Asian and Black patients compared with the pooled OBSErve studies, and a higher proportion of females. These differences are likely to reflect the OBSErve study series being conducted across several countries, with the BILAG-BR Benlysta Sub-Study only including patients eligible for belimumab treatment in England. Nonetheless, the BILAG-BR Benlysta Sub-Study included patients presenting with a wide range of SLE manifestations (defined using ACR criteria, see Table 26 in Appendix P) and commonly encountered comorbidities (most frequently hypertension, depression, and asthma, see Table 21 in Appendix P), increasing generalisability of the study.

B.2.3.5 Additional relevant studies

B.2.3.5.1 Phase 2B open-label, single-arm, repeat-dose study to evaluate the reliability of the SC autoinjector

While pre-filled syringes were used to deliver belimumab and placebo in the BLISS-SC trial, a single-use disposable autoinjector pen device was developed to maximise the safety and effectiveness of SC belimumab self-injections in routine clinical practice[69]. An open-label, single-arm, multi-dose Phase 2B study was conducted to assess the suitability of the autoinjector for self-administration of belimumab by patients with SLE (primary objective)[69]. Secondary objectives were to assess the use of the autoinjector inside and outside of the clinic setting[69]. Other objectives were to evaluate any injection failures related to use or device performance, to evaluate the safety and tolerability of belimumab administered via the autoinjector, and to characterise the change in belimumab trough concentrations when switching from IV to SC administration. Additional methodological details are provided in Appendix O.

B.2.3.5.2 Indirect treatment comparison between SC and IV belimumab formulations

An indirect treatment comparison (ITC) was performed to compare the efficacy of SC and IV belimumab formulations in patients with autoantibody-positive SLE with HDA to aid decision-making for physicians/patients considering a switch from IV to SC belimumab[70]. Patients were included in the analysis if they met one of the two following criteria: (i) low complement (C3 or C4) AND anti-ds DNA positive, or (ii) low complement (C3 or C4) OR SELENA-SLEDAI score ≥10. Data on 10 mg/kg IV belimumab were derived from the BLISS-52 (see Table 5), BLISS-76 (see Table 6) and North-East Asia studies (see Appendix O), while data on belimumab 200 mg SC were derived from the BLISS-SC trial[70] (see Table 4). See Appendix O for further methodological details on the ITC.

B.2.3.5.3 Phase 2 safety and efficacy study and its LTE: LBSL02

A Phase 2 study assessed the safety, tolerability, biologic activity, and efficacy of belimumab in combination ST in patients with SLE[62]. Patients with a SELENA-SLEDAI score ≥4 were randomly assigned to receive belimumab IV (1, 4, or 10 mg/kg) or placebo in a 52-week study. Co-primary endpoints were the percent change in the SELENA-SLEDAI score at week 24 and the time to first SLE flare[62]. The long-term safety and efficacy of belimumab IV was assessed over 13 years in patients who had completed the Phase 2 study[63]. An integrated analysis of this Phase 2 study and Phase 3 IV LTE studies (BLISS-52 and BLISS-76) were used to estimate the year 2 onwards discontinuation rate in the economic models (see section B.3.3.4.2). Further methodological details for the Phase 2 study are provided in Appendix L, and for its long-term extension in Appendix M.

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

Randomised controlled trials: statistical methodology of the BLISS-SC, BLISS-52 and BLISS-76 RCTs is summarised in Table 20. Additional statistical methods for the BLISS-SC study, are presented in Appendix L.

Long-term extension trials: statistical methodology of the BLISS-SC LTE, BLISS-52/76 non-US LTE and BLISS-76 US LTE is summarised in Table 21. Additional statistical methods, including the SLICC/SDI Indirect Cohort Comparison Study[48] are presented in Appendix M.

Real-world evidence studies: statistical consideration around key real-world evidence is summarised in Table 22. Additional details are provided in Appendix N.

Table 20. Summary of statistical analyses in the pivotal trials of belimumab[46, 47, 71]

Trial acronym	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
BLISS-SC[71]	Demonstrate superiority of belimumab 200 mg SC over placebo when comparing the SRI-4 response at Week 52.	The proportion of patients achieving a treatment response at Week 52 was compared between belimumab and placebo using a logistic regression model. The independent variables in the model included treatment groups, baseline SELENA-SLEDAI score, complement level and race. The analysed population was the same as BLISS-52 and BLISS-76, i.e. patients who were randomised and received ≥1 dose of study treatment. For the primary and the major secondary efficacy endpoints, a step-down sequential testing procedure was used to control the overall type 1 error rate. With this procedure, the primary and two major secondary endpoints were evaluated for statistical significance (2-sided alpha=0.05) based on a pre-specified sequence for interpretation: (1) SRI-4 response rate at Week 52, (2) time to first severe SLE flare, and (3) percentage of patients with average prednisone dose that has been reduced by ≥25% from baseline to ≤7.5 mg/day during Weeks 40 through 52.	The study aimed to randomise and treat approximately 816 patients, with a target of at least 544 patients in the belimumab arm and 272 patients in the placebo arm. This sample size provided at least 90% power at a 5% level of significance to detect a minimum of an evidence-based 12% absolute improvement in the response rate for the belimumab group relative to the placebo group at Week 52.	Similar across BLISS-52/76/SC: For the SRI-4 endpoint and its components, any patient who was classified as a treatment failure was considered a non-responder for the primary efficacy analysis and the supportive analyses of the primary efficacy endpoint. A treatment failure was defined as any patient who: withdrew from the study prior to Week 52 and had no visit within ±28 days of Week 52, and/or received a protocol-prohibited medication or a dose of allowable (but protocol-restricted) medication that resulted in treatment failure designation prior to Week 52.
BLISS-52 and BLISS-76	Demonstrate superiority of belimumab 10 mg/kg IV over placebo when comparing the SRI-4	The percentage of patients achieving a response at Week 52 was compared between belimumab 10 mg/kg and placebo using a logistic regression model. The independent variables in the model included treatment groups, baseline SELENA-	Both BLISS-52 and BLISS-76 studies aimed to randomise approximately 810 patients (per study), with a target of at least 270 patients per treatment group	

Trial acronym	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
	response rate at Week 52.	SLEDAI score (≤9 vs ≥10), baseline proteinuria level (<2 g/24 hour vs ≥2 g/24 hour equivalent) and race (African descent or indigenous-American descent versus other). The population analysed was defined as for BLISS-SC.	(per study). This sample size provided at least 90% power at a 5% level of significance to detect a minimum of a 14% absolute improvement in the response rate in the 10 mg/kg belimumab group relative to the placebo group at Week 52.	
IV: intravenous; SELENA: Safety of Estrogen in Lupus Erythematosus National Assessment; SC: subcutaneous; SLE: systemic lupus erythematosus; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index; SRI-4: SLE responder index-4.				

Table 21. Summary of statistical analyses in LTEs to key belimumab RCTs[49,50,51]

Trial acronym	Parent study	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
BLISS-SC LTE[49]	BLISS SC	As these trials were open-label, continuation studies, no formal statistical hypothesis testing was performed. All analyses were exploratory in nature and were summarised using descriptive statistics.	Similar across trials: unless otherwise stated, continuous variables were summarised with mean, median, SD, 25th and 75th percentiles, minimum and maximum. Categorical variables were summarised with frequency counts and percentages.	Enrolment was voluntary and dependent on completion of the parent study, thus no sample size calculations were performed. Analyses were conducted using descriptive statistics, and no power calculations were required.	Analyses were performed on the ITT population. The ITT analysis was performed according to the treatment that a patient was randomised to receive, regardless of the actual treatment received.
BLISS-52/76 non-US LTE[50]	BLISS-52 or BLISS-76[a]				Analyses were performed on the modified ITT population, defined as all patients who were enrolled and treated with at least one dose of belimumab in the continuation study.
BLISS-76 US LTE[51]	BLISS-76				
[a]In addition, 5 patients from BEL112232 (NCT00732940) entered into the non-US LTE study per Mexico National Amendment 01. These patients, participating in the only Mexican site in BEL112232, originally received belimumab SC. *ITT: intention-to-treat; LTE: long-term extension; RCT: randomised controlled trial; SC: subcutaneous; SD: standard deviation.*					

Table 22. Summary of statistical analyses in key real-world studies[53,68]

Trial acronym	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
BILAG-BR[53]	Due to the observational nature of the data, no formal hypothesis was tested.	The following descriptive statistics were used to summarise the characteristics of the data: continuous variables: mean, SD, median, 25th and 75th percentiles, minimum and maximum; categorical variables: frequency counts and percentages.	The data followed a multilevel longitudinal structure, where patients were recruited from multiple centres across the UK and followed up over time, thus no sample size calculations were performed, and no power calculations were required.	The baseline assumption was that patients would receive the treatment regime required to manage their disease. The primary analysis therefore followed a similar protocol to an ITT analysis where treatment assigned at baseline regardless of dose, frequency or adherence was investigated. Two sensitivity analyses were conducted 1. Only the first round of treatment for study participants were included. 2. Only included participants who continued the same treatment follow up period. i.e. excluding patients who have switched treatments.
OBSErve registry series[68]	No formal statistical hypothesis was tested. The study was exploratory in nature.	Continuous data were statistically summarised using means, corresponding 95% confidence intervals, SD, minimum, median, maximum, range. Categorical data were summarised using counts and frequencies. A logistic regression model was employed to assess physician-assessed overall clinical response to belimumab treatment. The overall clinical response was compared between subgroups of patients exposed to 1 vs ≥2 immunosuppressants prior to belimumab initiation using the Wilcoxon rank sum test. Primary and secondary endpoints were selected	For each individual study, the sample size was based on the feasibility. The pooled meta-analysis included data on 830 patients who were exposed to belimumab in clinical practice settings.	Patient data was collected from the individual studies in a manner that they could not be identified, directly or through identifiers linked to the patients. Given the small number of patients with available data and in only some studies, characteristics of patients discontinuing treatment with belimumab and time to treatment discontinuation in the

Trial acronym	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
		based on consistent reporting across individual OBSErve studies.		pooled OBSErve cohort were not analysed.
BILAG: British Isles Lupus Assessment Group; ITT: intention-to-treat; SD: standard deviation; UK: United Kingdom.				

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

Quality assessment of belimumab RCTs is available in Table 23. The quality assessment of the non-randomised LTE studies and the RWE studies, both using the Downs and Black checklist[72], is available in Appendix D.

Table 23. Quality assessment results for pivotal belimumab RCTs

Trial acronym	BLISS-SC	BLISS-52	BLISS-76
Was randomisation carried out appropriately?	Yes	Yes	Yes
Was the concealment of treatment allocation adequate?	Yes	Yes	Yes
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	Yes	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes	Yes	Yes
Were there any unexpected imbalances in drop-outs between groups?	No	No	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No	No
Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Yes	Yes
ITT: intention-to-treat; RCT: randomised controlled trial; SC: subcutaneous.			

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

B.2.6.1 Introduction to presented evidence

We will only present results for the 10 mg/kg belimumab dose examined in the Phase 3 BLISS-52 and BLISS-76, as this is the dose submitted for Marketing Authorisation. When discussing the results from the Phase 3 studies, the belimumab group refers to belimumab plus ST and the placebo group refers to placebo plus ST.

All three (two IV: BLISS-52, BLISS-76; one SC: BLISS SC) Phase 3 trials were positive. The primary endpoint, SRI-4 response at Week 52 was met in all the Phase 3 trials. These studies provided clear evidence for the efficacy of belimumab, as measured by reductions in disease activity assessed using the SRI-4. Reductions in the risk of severe flare were also observed, alongside improvements in several other disease activity indices and QoL.

Key results for the total population of BLISS-SC are presented herein, while pooled data for the total population of BLISS-52 and BLISS-76 (previously presented to NICE in TA397) are presented in Appendix L.

Long-term data on the outcomes of treatment with belimumab, based on three LTE studies: one of SC belimumab (BLISS-SC LTE) and two of IV belimumab (BLISS-52/76 non-US LTE and BLISS-76 US LTE), are presented in this section. Where available, data from the aforementioned RCTs and their non-randomised LTEs are supplemented with real-world evidence from the OBSErve study series.

Please note that that this section describes results in the total population as enrolled in the trials and real-world studies described. Clinical trial and real-world data pertaining to HDA sub-populations, including the current NICE-approved HDA-1 population, are presented in Section B.2.7.

B.2.6.2 BLISS-SC

BLISS-SC was a 52-week, randomised, double-blind, placebo-controlled Phase 3 trial to assess the safety and efficacy of belimumab SC in patients with moderate to severe SLE[45]. The primary endpoint was SRI-4 response rate at Week 52[45]. Details of the trial methodology are provided in Appendix L. The results of this trial were published in 2017[73].

B.2.6.2.1 BLISS-SC: Primary efficacy endpoint (SRI-4 response rate at Week 52) and its components

SRI-4 response at Week 52 was defined as a ≥4-point reduction from baseline in SELENA-SLEDAI score, no worsening (increase of <0.30 points from baseline) in PGA, and no new BILAG A organ domain score or 2 new BILAG B organ domain scores compared with baseline[45].

The BLISS-SC trial met its primary endpoint. In the ITT population, the percentage of responders at Week 52 was higher in the belimumab group than the placebo group (61.4% versus 48.4%)[45]. The odds of being an SRI-4 responder at Week 52 were significantly higher for patients in the belimumab group compared with the placebo

group (Table 24)[45]. Maintained SRI-4 response was also achieved more promptly with belimumab than placebo (235 days versus 338 days; p=0.0001)[45].

Table 24. BLISS-SC: Primary efficacy endpoint (SRI-4) at Week 52 (ITT population)

	Placebo N=280	**Belimumab 200 mg SC N=556**
Response (primary efficacy analysis), N/N (%)	135/279 (48.4)	340/554 (61.4)
Observed difference vs placebo (%)	-	12.98
OR (95% CI) vs. placebo	-	1.68 (1.25–2.25)
p-value	-	0.0006
Note that 1 patient in the placebo group and 2 in the belimumab group who had no baseline PGA assessment were excluded from the analysis of SRI-4 and its components. *CI: confidence interval; OR: odds ratio; PGA: Physician's Global Assessment; SRI-4: SLE responder index-4.*		

The significant improvement associated with belimumab was observed consistently across the individual components of the primary endpoint (Table 25). Sensitivity analyses of the primary endpoint, including SRI5–8 and the use of SLEDAI 2K score were generally consistent with the results of the primary analysis[45]. The response rates observed in the pre-specified subgroups were also generally consistent with those observed in the overall population[45].

Table 25. BLISS-SC: Components of SRI-4 response rate at Week 52 (ITT population)[45]

Primary endpoint component	**Placebo N=280**	**Belimumab 200 mg SC N=556**
4-point reduction in SELENA-SLEDAI, N/N (%)	137/279 (49.1)	345/554 (62.3)
OR (95% CI) vs. placebo	-	1.69 (1.26, 2.27)
p-value	-	0.0005
No worsening in PGA, N/N (%)	203/279 (72.8)	450/554 (81.2)
OR (95% CI) vs. placebo	-	1.61 (1.15, 2.27)
p-value	-	0.0061
No new 1A/2B BILAG domain scores, N/N (%)	207/279 (74.2)	448/554 (80.9)
OR (95% CI) vs. placebo	-	1.46 (1.04, 2.07)
p-value	-	0.0305
Note that 1 patient in the placebo group and 2 in the belimumab group who had no baseline PGA assessment were excluded from the analysis of SRI-4 and its components. *BILAG: British Isles Lupus Assessment Group; CI: confidence interval; ITT: intention-to-treat; OR: odds ratio; PGA: Physician's Global Assessment; SELENA: Safety of Estrogen in Lupus Erythematosus National Assessment; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index; SRI-4, SLE responder index-4.*		

The significant improvement in SRI-4 response with belimumab was evident at all visits from Week 16 to Week 52, observed across all components of SRI-4 response (Figure 2)[45]

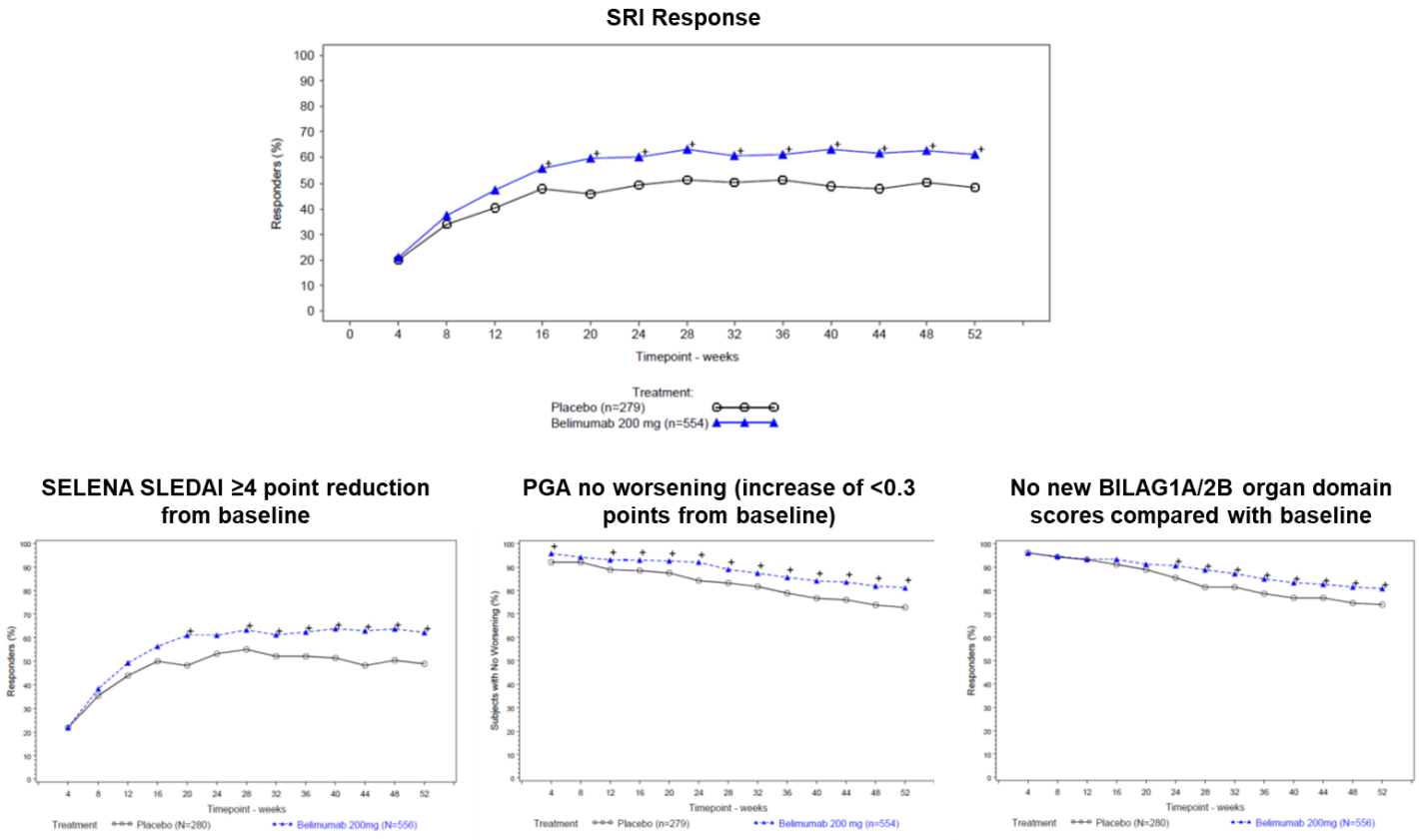


Figure 2. BLISS-SC: SRI-4 response and its components by visit (ITT population)[45]

BILAG: British Isles Lupus Assessment Group; CI: confidence interval; ITT: intention-to-treat; PGA: Physician's Global Assessment; SELENA: Safety of Estrogen in Lupus Erythematosus National Assessment; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index; SRI-4: SLE responder index-4.

B.2.6.2.2 BLISS-SC: SELENA-SLEDAI score

The SELENA-SLEDAI measures the presence or absence of SLE signs, symptoms or laboratory anomalies. Complete elimination of symptoms is required to indicate a change in disease activity. Therefore, a reduction in SELENA-SLEDAI score is clinically important because it represents resolution of individual manifestations of the patient's disease activity or normalisation of serology. Therefore, a 4 point-reduction in SELENA-SLEDAI score, is a clear demonstration of clinical benefit.

The mean change from baseline SELENA-SLEDAI score at Week 52 was significantly greater for belimumab than placebo, but the percent change from baseline was not statistically different across groups (Table 26)[45].

Table 26. BLISS-SC: Mean percent change and change in SELENA-SLEDAI scores from baseline at Week 52

	Placebo N=280	**Belimumab 200 mg SC N=556**
Mean change from baseline (±SE)	–3.55 (0.31)	–4.39 (0.26)
p-value		0.0069
Mean % change (±SE)	–33.22 (3.68)	–39.96 (3.07)
p-value		0.0660

All statistics are from an analysis of covariance (ANCOVA) model comparing belimumab and placebo with covariates for treatment group, baseline SELENA-SLEDAI score (≤9 vs ≥10), baseline complement levels (low C3 and/or C4 vs. no low C3 or C4) and race (Black vs. other).
SC: subcutaneous; SE: standard error; SELENA: Safety of Estrogen in Lupus Erythematosus National Assessment; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index.

B.2.6.2.3 BLISS-SC: PGA

Over time, patients in both placebo and belimumab groups showed reductions from baseline in PGA (i.e. improving overall condition). Add-on treatment with belimumab was associated with a significant reduction in PGA compared with ST alone at Week 52 (Table 27)[45].

Table 27. BLISS-SC: PGA mean percent change and change from baseline at Week 52

	Placebo N=280	Belimumab 200 mg SC N=556
Mean change from baseline (±SE)	–0.61 (0.04)	–0.77 (0.04)
p-value		0.0003
Mean % change (±SE)	–35.10 (2.91)	–47.87 (2.44)
p-value		<0.0001
All statistics are from an analysis of covariance (ANCOVA) model comparing belimumab and placebo with covariates for treatment group, baseline SELENA-SLEDAI score (≤9 vs. ≥10), baseline complement levels (low C3 and/or C4 vs. no low C3 or C4) and race (Black vs. other). *SC, subcutaneous; PGA: Physician's Global Assessment; SE: standard error; SELENA: Safety of Estrogen in Lupus Erythematosus National Assessment; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index.*		

B.2.6.2.4 BLISS-SC: Steroid use

The percentage of patients whose average prednisone dose was reduced by ≥25% from baseline to ≤7.5 mg/day during Weeks 40–52 in patients receiving >7.5 mg/day at baseline (N=503, 60.2% of the study population), was a major secondary endpoint of the BLISS-SC trial[45]. Given there was no tapering regime in the study protocol, and the double-blind design may have caused a hesitancy to reduce steroid dosage, a greater proportion of patients receiving belimumab were able to reduce their prednisone dose by ≥25% from baseline to ≤7.5 mg/day during Weeks 40–52. (18.2% vs 11.9%). Whilst the odds ratio did not reach statistical significance (Table 28)[45], clinical experts consulted during this submission process concurred that every 1 mg reduction in steroid dose significantly reduces the risk of long-term sequelae from steroid usage.

Table 28. BLISS-SC: Prednisone reduction by ≥25% from baseline to ≤7.5 mg/day during Weeks 40–52[45]

	Placebo N=280	Belimumab 200 mg SC N=556
Reduction in prednisone[a], N/N (%)	20/168 (11.9)	61/335 (18.2)
Observed difference vs. placebo		6.30
OR (95% CI) vs. placebo		1.65 (0.95, 2.84)
p-value		0.0732
[a]Includes only patients with baseline prednisone dose >7.5 mg/day. *CI: confidence interval; OR: odds ratio; SC: subcutaneous.*		

Furthermore, a lower percentage of patients in the belimumab group (range: 3.8–8.1%) than in the placebo group (range: 2.9–13.9%) required any increase in prednisone dose over the 52-week study period (Figure 3)[45]. At Week 52, a significantly greater proportion of patients in the placebo group than the belimumab group had any increase in prednisone dose (13.2% vs 8.1%, respectively; p=0.0117)[45].

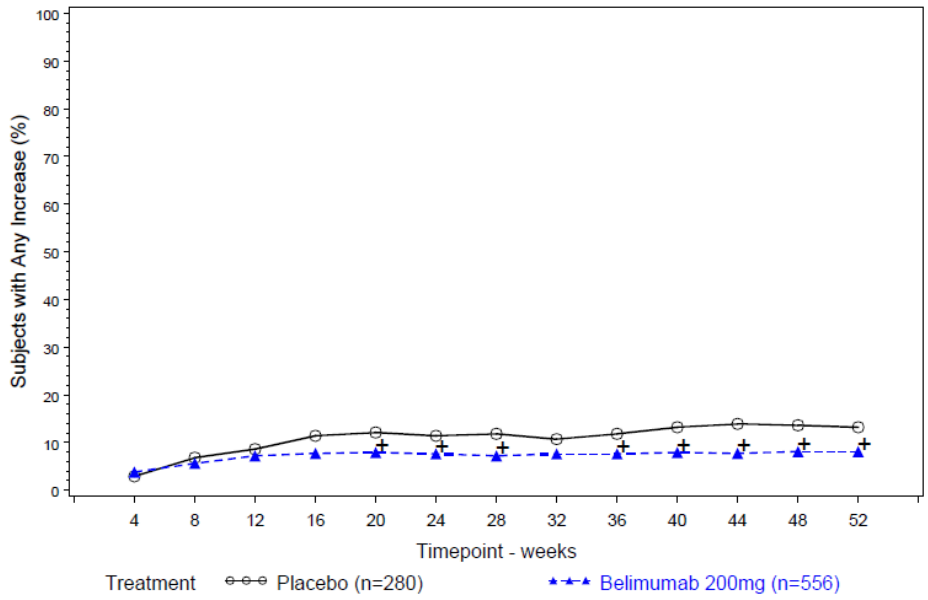


Figure 3. BLISS-SC: Any increase in prednisone dose from baseline by visit (ITT population)[45]
ITT: intention-to-treat; SC: subcutaneous.

B.2.6.2.5 BLISS-SC: SLE Flare Index

Time to first severe SFI flare was a major secondary endpoint of the BLISS-SC trial[45]. Add-on treatment with belimumab was associated with a 49% lower risk of experiencing a severe SFI flare than placebo (HR: 0.51; 95% CI: 0.35–0.74; p=0.0004)[45]. While a total of 51 (18.2%) patients in the placebo group and 59 (10.6%) patients in the belimumab group experienced a severe flare, the median time to first severe flare was delayed in the belimumab group compared with the placebo group (171.0 days vs. 118.0 days)[45]. The probability of experiencing a first severe flare over the 52-week study is presented in Figure 4A.

The risk of experiencing any SFI flare (mild/moderate or severe) was also significantly reduced with add-on belimumab treatment. Patients in the belimumab group had a 22% lower risk of a first SFI flare over 52 weeks than patients in the placebo group (HR 0.78; 95% CI: 0.65, 0.93; p=0.0061)[45]. Among the 192 (68.6%) patients in the placebo group and 337 (60.6%) patients in the belimumab group who experienced an SFI flare, median time to first flare was prolonged in the belimumab group compared with the placebo group (190 days vs 141 days)[45]. The probability of experiencing a first SFI flare (mild/moderate or severe) over the 52 weeks of the study is presented in Figure 4B.

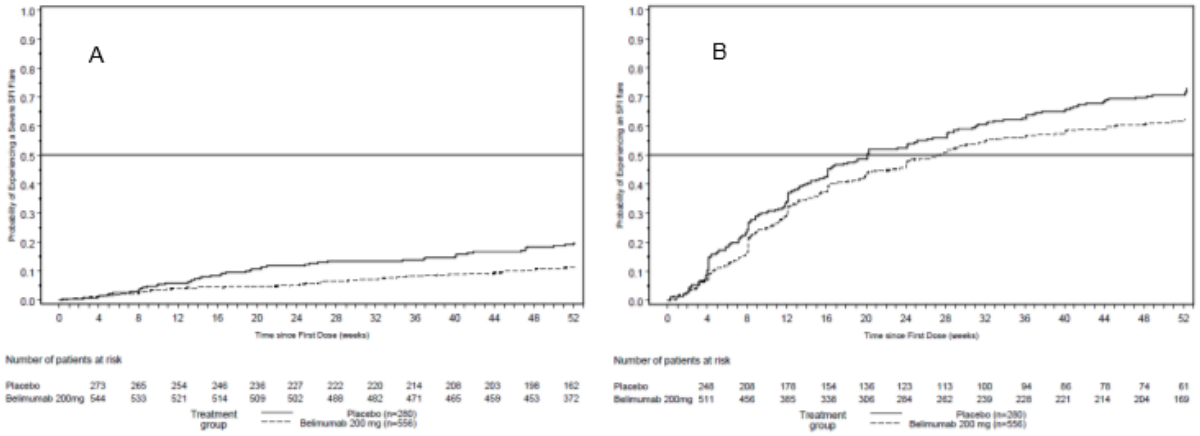


Figure 4. BLISS-SC: Time to first severe SFI flare over 52 weeks (A) and Time to first SFI flare over 52 weeks (B) (ITT population)[45]

ITT: intention-to-treat; SC: subcutaneous; SFI: SLE Flare Index.

B.2.6.2.6 BLISS-SC: Organ damage (SDI)

SDI scores were similar between belimumab and placebo groups at baseline with a similar change from baseline to Week 52 (0.0 and 0.1 for belimumab and placebo groups, respectively: p=0.1174)[45]. At Week 52, 203 (72.5%) patients in the placebo group and 446 (80.2%) patients in the belimumab group experienced no worsening (change ≤0) in the SDI compared with baseline and the odds of not experiencing an SDI worsening significantly favoured belimumab (OR 1.54; 95% CI: 1.10–2.16; p=0.0123)[45].

B.2.6.2.7 BLISS-SC: FACIT Fatigue Scale

FACIT Fatigue is a patient-reported outcome measure that assesses the individual's level of fatigue during their daily activities over the previous week. Add-on treatment with belimumab was associated with a significant reduction in patient-reported fatigue. While patients in the placebo and belimumab groups had an increased (improved) mean FACIT Fatigue score at Week 52 (Figure 5), the adjusted mean change from baseline was significantly greater with belimumab compared with placebo (4.4 vs 2.7, respectively; treatment difference 1.6; 95% CI 0.3–2.9; p=0.0130)[45]. In addition, at Week 52, a higher proportion of patients in the belimumab group (246 [44.4%]) than the placebo group (101 [36.1%]) experienced improvement in FACIT Fatigue score exceeding the minimal clinically important difference (MCID), i.e. improvement ≥4 (OR: 1.42, 95% CI: 1.05–1.94, p=0.0245)[45].

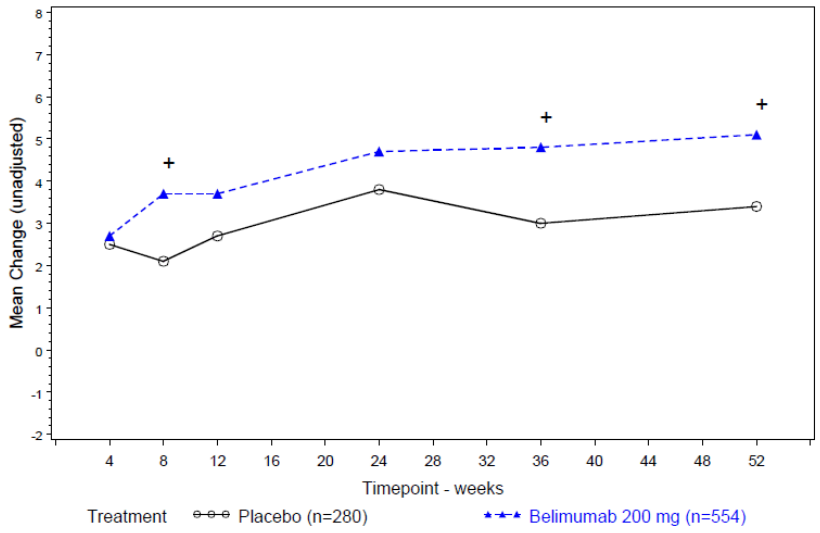


Figure 5. FACIT-Fatigue score change from baseline by visit (ITT population)[45]

FACIT: Functional Assessment of Chronic Illness Therapy; ITT: intention-to-treat.

B.2.6.3 Long-term evidence from clinical studies

B.2.6.3.1 BLISS-SC LTE

BLISS-SC LTE was a 6-month open-label extension phase to evaluate the long-term efficacy, safety and tolerability profile of belimumab SC in adult patients who completed the double-blind phase of BLISS-SC[49]. Details of the trial methodology are provided in Appendix M.

A total of 662 out of the 677 patients that completed the double-blind phase were enrolled in the open-label extension study and received at least one dose of belimumab 200 mg SC during the open-label phase[49]. Of these, 206 patients who had received placebo during the double-blind phase were switched to belimumab 200 mg SC (placebo-to-belimumab group) and 456 patients who had received belimumab 200 mg SC during the double-blind phase continued to do so in the open-label extension phase (belimumab group)[49].

A higher percentage of patients who were randomised to belimumab in the double-blind phase achieved a SRI-4 response at open-label phase Week 24/Exit, and continued to demonstrate improvement in other key areas compared with patients randomised to placebo in the double-blind phase (Table 29). Efficacy of belimumab was maintained over the course of the 6-month open-label phase.

Table 29. BLISS-SC LTE-Key results at Week 24/Exit[49]

	Placebo-to-belimumab 200 mg SC N=206	**Belimumab 200 mg SC N=456**	**Total N=662**
SRI-4 and composite responses[a]			
SRI-4 responder, N/N (%)	23/143 (16.1)	332/435 (76.3)	355/578 (61.4)
SELENA-SLEDAI ≥4 point reduction, N/N (%)	25/143 (17.5)	345/435 (79.3)	370/578 (64.0)
No worsening in PGA, N/N (%)	125/143 (87.4)	426/435 (97.9)	551/578 (95.3)
No new 1A/2B BILAG domain scores, N/N (%)	134/143 (93.7)	417/435 (95.9)	551/578 (95.3)
SFI flares			
Any flare, mild/moderate or severe, N/N (%)	38.206 (18.4)	58/456 (12.7)	96/662 (14.5)
Severe flares, N/N (%)	2/206 (1.0)	12/456 (2.6)	14/662 (2.1)
Time to first severe flare, days (median)	169.0	169.0	169.0
SDI over time			
SDI change from baseline (median)	0.0	0.0	0.0
SDI worsening (change>0), N/N (%)	6/206 (2.9)	19/456 (4.2)	25/662 (3.8)
SELENA-SLEDAI			
≥4 point reduction from baseline, N/N (%)	25/147 (17.0)	356/448 (79.5)	381/595 (64.0)
Percent change from baseline[b], mean (SD)	–9.2 (44.94)	–58.0 (37.98)	–43.8 (45.83)
Change from baseline, mean (SD)	–0.7 (2.75)	–6.3 (4.04)	–4.6 (4.52)
Prednisone use over time			
Change from baseline, mean (SD)	–0.1 (2.11)	–2.3 (6.78)	–1.6 (5.85)
Reduction from >7.5 to ≤7.5 mg/day, N/N (%)	10/102 (9.8)	67/275 (24.4)	77/377 (20.4)

Increase from ≤7.5 to >7.5 mg/day, N/N (%)	2/104 (1.9)	9/181 (5.0)	11/285 (3.9)
FACIT-Fatigue Score			
Change from baseline, mean (SD)	0.7 (7.07)	5.6 (10.63)	4.0 (9.88)

[a]1 belimumab 200 mg to belimumab 200 mg patient did not have a baseline PGA assessment, 2 belimumab 200 mg to belimumab 200 mg patients and 55 placebo to belimumab 200 mg patients had a baseline SELENA-SLEDAI score <4, and 18 belimumab 200 mg to belimumab 200 mg patients and 8 placebo to belimumab 200 mg patients had a missing visit or missing SRI-4 components, and therefore do not contribute to SRI-4/component analyses. [b]Patients with a baseline score of zero are excluded from the analyses due to division by zero.
BILAG: British Isles Lupus Assessment Group; FACIT: Functional Assessment of Chronic Illness Therapy; PGA: Physician's Global Assessment; SD: standard deviation; SDI, Systemic Lupus International Collaborating Clinics (SLICC)/ACR Damage Index; SELENA: Safety of Estrogen in Lupus Erythematosus National Assessment; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index; SFI: SLE Flare Index; SRI-4: SLE responder index-4.

B.2.6.3.2 BLISS-76 US LTE

BLISS-76 US LTE was a multicentre, continuation trial of belimumab in SLE patients who completed the Phase 3 BLISS-76 study in the US[51]. The primary outcome measure was long-term safety of belimumab and organ damage assessed using the SDI. Details of the trial methodology are provided in Appendix M. *A total of 268 patients received at least 1 dose of belimumab in the study.*

In the early years of this LTE, there was a *marked difference in the number of* SRI-4 responders *depending on the parent study (BLISS-76) treatment assignment*[51]*.* Over the duration of the LTE, this difference diminished and a gradual increase in the proportion of SRI-4 responders was observed (Table 30). Organ damage accrual was low during the course of the study, and efficacy was maintained[51].

Table 30. BLISS-76 US LTE: Key results at Year 1 and Year 7[51]

	Placebo-to-belimumab 10 mg/kg IV N=91		Belimumab IV[a] N=177		Total N=268	
	Year 1[b]	**Year 7[b]**	**Year 1[b]**	**Year 7[b]**	**Year 1[b]**	**Year 7[b]**
SRI-4 and composite responses						
SRI-4 responder, N/N (%)	10/60 (16.7)	6/7 (85.7)	86/169 (50.9)	84/112 (75.0)	96/229 (41.9)	90/119 (75.6)
SELENA-SLEDAI ≥4 point reduction, N/N (%)	13/65 (20.0)	7/7 (100.0)	91/169 (53.8)	86/112 (76.8)	104/234 (44.4)	93/119 (78.2)
No worsening in PGA, N/N (%)	67/84 (79.8)	11/12 (91.7)	166/177 (93.8)	108/115 (93.9)	233/261 (89.3)	119/127 (93.7)
No new 1A/2B BILAG domain scores, N/N (%)	84/88 (95.5)	12/12 (100)	174/177 (98.3)	113/115 (98.3)	258/265 (97.4)	125/127 (98.4)
SFI flares						
Any flare, mild/moderate or severe, N/N (%)	22/88 (25.0)	3/12 (25.0)	37/174 (21.3)	33/115 (28.7)	59/262 (22.5)	36/127 (28.3)
Severe flares, N/N (%)	7/90 (7.8)	21/90 (23.3)	8/177 (4.5)	34/177 (19.2)	15/267 (5.6)	55/267 (20.6)
SDI over time[c]						
SDI change from baseline, N/N, mean (SD)	88/91 0.1 (0.31)	11/91 0.5 (0.69)	175/177 0.1 (0.26)	115/177 0.4 (0.68)	263/268 0.1 (0.28)	126/268 0.4 (0.68)
SDI worsening (change>0), N/N (%)	6/88 (6.8)	5/11 (45.5)	9/175 (5.1)	33/115 (28.7)	15/263 (5.7)	38/126 (30.2)
Prednisone use over time						
% change from baseline[d], median (min, max)	0.000 (–100.00, 380.84)	–51.282 (–100.00; 0.00)	0.000 (–48.96, 280.38)	–47.106 (–100.00, 300.00)	0.000 (–100.00, 380.84)	–47.106 (–100.00, 300.00)
Reduction from >7.5 to ≤7.5 mg/day, N/N (%)	0/19 (0.0)	0/1 (0.0)	2/61 (3.3)	16/41 (39.0)	2/80 (2.5)	16/42 (38.1)
Quality of life, change from baseline, N/N, mean (SD)[c,e]						
SF36v2 PCS	86/91 0.41 (7.88)	58/91 0.88 (7.79)[d]	173/177 4.90 (8.58)	127/177 6.57 (9.57)[d]	259/268 3.41 (8.60)	185/268 4.79 (9.41)[d]
SF36v2 MCS	86/91 –0.30 (9.18)	58/91 –0.58 (9.29)[d]	173/177 3.85 (10.29)	127/177 4.21 (11.81)[d]	259/268 2.47 (10.11)	185/268 2.71 (11.27)[d]
FACIT-Fatigue Score	87/91 1.07 (9.71)	58/91 –0.37 (9.54)[d]	173/177 6.85 (10.81)	126/177 5.58 (12.27)[d]	260/268 4.91 (10.79)	184/268 3.70 (11.79)[d]

[a]patients who had received belimumab 1 mg/kg or 10 mg/kg in BLISS-76 continued on the same dose in this trial (the protocol was later amended to increase the 1 mg/kg dose to the licensed 10 mg/kg. [b]values reflect the year midpoint (Week 24), apart from SDI and quality of life measurements which were taken at Week 48; [c]SDI and quality of life measurements were taken at Week 48; [d]in patients taking prednisone at baseline, [e]Note quality of life measures were recorded at Week 48 of Year 6, not Year 7.
BILAG: British Isles Lupus Assessment Group; FACIT: Functional Assessment of Chronic Illness Therapy; LTE: long-term extension; MCS: Mental component score; PCS: Physical Component Score; PGA: Physician's Global Assessment; SELENA: Safety of Estrogen in Lupus Erythematosus National Assessment; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index; SD: standard deviation; SDI, Systemic Lupus International Collaborating Clinics (SLICC)/ACR Damage Index; SFI: SLE Flare Index; SRI-4: SLE responder index-4; US: United States.

B.2.6.3.3 SLICC(ACR)/SDI Indirect Cohort Study

Organ damage progression in SLE patients who received belimumab in the BLIS-LTE was compared with propensity score matched (PSM) patients treated with ST from the Toronto Lupus Cohort (TLC). The primary endpoint was the difference in change in Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index (SDI) score from baseline to 5 years. A total of 259 patients from the BLISS-76 US LTE study[51] and 706 patients from the Toronto Lupus Cohort (TLC)[26, 35] were included in the PSM analysis, and 99 patients from each of the studies were 1:1 PS-matched[48].

Over a 5-year period, patients treated with belimumab experienced significantly less organ damage than patients treated with ST alone (Table 31). Further, patients receiving belimumab were 61% less likely to progress to a higher SDI score over any given year of follow-up compared with patients treated with ST (HR 0.391; 95% CI 0.253 to 0.605; p<0.001). A patient receiving belimumab had a 3.5% annual probability of organ damage progression compared with an 8.7% annual probability of progression with ST alone[48].

When the magnitude of year-to-year organ damage progression was explored, it was found that of those patients treated with belimumab there were 33 instances of an SDI score increase of ≥1 compared with 72 instances in patients treated with ST.

A higher proportion of patients treated with ST experienced an SDI score increase ≥2 compared with patients treated with belimumab (p=0.006)[48].

Table 31. PSM analysis: Change in SDI from baseline to 5-years

	ST N=99	**Belimumab N=99**	**Difference**
5-year SDI change, mean (SE)	0.717	0.283	–0.434 (0.119)
95% CI	0.500 to 0.934	0.166 to 0.400	–0.667 to –0.201
p-value			<0.001

CI: confidence interval; PSM: propensity score matching; SD: standard deviation; SDI: Systemic Lupus International Collaborating Clinics (SLICC)/ACR Damage Index; ST: standard therapy.

B.2.6.3.4 BLISS-52/76 non-US LTE

BLISS-52/76 non-US LTE[50] was a multicentre continuation trial to evaluate the long-term safety and tolerability of belimumab in patients with SLE who completed the Phase 3 BLISS-52[46] or BLISS-76 trials[47]. Patients were monitored for safety and disease activity, including organ damage accrual (assessed using the SDI)[50]. The study was continued until belimumab was commercially available. Details of the trial methodology are provided in Appendix M. A total of 735 patients received at least 1 dose of belimumab during the study[50]. Observed organ damage, assessed using the changes in SDI over time, was low in patients treated with belimumab (Table 32).

Table 32. BLISS-52/76 non-US LTE: Change in SDI from baseline to Year 8[50]

	Placebo-to-belimumab 10 mg/kg IV N=232		**Belimumab[a] N=503**		**Total N=735**	
	Year 1[b]	**Year 8[b]**	**Year 1[b]**	**Year 8[b]**	**Year 1[b]**	**Year 8[b]**
SDI over time						
SDI change from baseline, mean (SD)	0.1 (0.22)	0.0 (0.00)	0.1 (0.31)	0.2 (0.58)	0.1 (0.29)	0.2 (0.56)
SDI worsening (change>0), N/N (%)	11/220 (5.00)	0/5 (0.00)	28/496 (5.6)	8/60 (13.3)	39/716 (5.4)	8/65 (12.3)

[a]patients who had received belimumab 1 mg/kg or 10 mg/kg in BLISS-76 continued on the same dose in this trial (the protocol was later amended to increase the 1 mg/kg dose to the licensed 10 mg/kg. [b]values reflect the year endpoint (Week 48).
LTE: long-term extension; SD: standard deviation; SDI, Systemic Lupus International Collaborating Clinics (SLICC)/ACR Damage Index; US: United States.

B.2.6.4 Real-world evidence

B.2.6.4.1 OBSErve registry series

The impact of belimumab on patient outcomes in the real-world setting has been captured in the OBSErve series of non-randomised, single-arm, retrospective, observational studies conducted in Argentina, Canada, Germany, Spain, Switzerland, and the US (see Appendix N). A total of 830 patients were included in the pooled analysis[68]

With regards to the primary endpoint of physician-assessed overall clinical response to belimumab at 6 months, 82.8% of patients had ≥20% improvement and 48.1% had ≥50% improvement in their overall condition (Table 33). Belimumab was

steroid-sparing; most patients receiving steroids at belimumab initiation decreased their steroid dose after 6 months of belimumab treatment[68], with improvements continuing through 24 months in the US and Argentina OBSErve studies[55, 74].

Table 33. OBSErve registry series: Key results at 6 months (pooled analysis)[54, 68, 75]

	Pooled dataset[a] N=830
Physician-assessed overall clinical response, N (%)	
Worse	10 (1.20)
No improvement	29 (3.49)
<20%	104 (12.53)
20–49%	288 (34.70)
50–79%	292 (35.18)
≥80%	107 (12.89)
Prednisone use over time (mg)	
Dose change from baseline, mean (SD)	–8.5 (10.7)
Reduction from >7.5 to ≤7.5 mg/day, N/N (%)	258/491 (52.6)
SLEDAI score **(N=344)**	
Change from baseline, mean (SD)	–5.7 (4.5)

[a]pooled data from OBSErve registries conducted in Argentina, Canada, Germany, Spain, Switzerland, and the US.
SLEDAI: Systemic Lupus Erythematosus Disease Activity Index; SD: standard deviation.

B.2.6.4.2 BILAG-BR

The BILAG-BR presented data from the HDA-1 subgroup, therefore further information will be presented in the subgroup analyses section (see Section B.2.7.4).

B.2.6.5 Additional relevant evidence

This section contains an overview of key additional evidence supporting the SC belimumab formulation.

B.2.6.5.1 Phase 2B open-label, single-arm, repeat-dose study to evaluate the reliability of the SC autoinjector

An open-label, single-arm, multi-dose Phase 2B study was conducted to assess the suitability of the autoinjector for self-administration of belimumab by patients with SLE (primary objective)[69]. A total of 95 patients were enrolled in the study[69].

The majority of injections were completed with the first attempt: only 5 patients required a second attempt, each on a single occasion, in order to achieve a successful injection. Only the data related to the second successful injection for these patients were used for the summary (Table 34)[69]. In terms of device reliability, there were 2 reported device malfunctions, both reviewed by the device development group[69]. One of these was substantiated as an actual device error, representing a functional performance rate of 99.9% for the autoinjector considering the total number of attempted injections (N=736)[69]. Further results are provided in Appendix O.

One of the objectives of the autoinjector reliability study was to characterise the change in belimumab trough concentrations when switching from IV to SC administration[69]. With the protocol-specified time window of 1 to 4 weeks between the last IV and the first SC dose (violated by several patients), levels close to steady-state were achieved by Week 3 or earlier for most patients[69]. The average steady-state exposure for weekly 200 mg belimumab SC was similar to the average concentration over the dosing interval for 10 mg/kg belimumab IV administered every 4 weeks[69].

Table 34. SC autoinjector study: Successful injections by week

	Belimumab 200 mg Autoinjector N=95
Primary efficacy endpoint, N/N (%)	
Weeks 1 and 2 (inside clinic)	89/90 (99)
Secondary efficacy endpoints, N/N (%)	
Weeks 4 and 8 (inside clinic)	85/87 (98)
Weeks 3, 5, 6 and 7 (outside clinic)	81/87 (93)
SC: subcutaneous.	

In a follow-up study that explored patient experiences with the autoinjector, and those of switching from IV to SC belimumab (N=21), the majority of participants indicated they preferred the autoinjector to the IV, and were confident in the use of the autoinjector, rating it as convenient and easy to use[76, 77].

B.2.6.5.2 Indirect treatment comparison between SC and IV belimumab formulations

An ITC was performed to compare the efficacy of SC and IV belimumab formulations plus standard therapy in patients with autoantibody-positive SLE with HDA (Ramachandran et al, 2018)[70]. Overall, belimumab IV and SC were found to have similar efficacy for the percentage of patients with an SRI-4 response, ≥4-point reduction in SELENA-SLEDAI and rate of severe SFI flares at Week 52 in patients with HDA (Table 35)[70]. Further results are provided in Appendix O.

Table 35. ITC of SC and IV belimumab: SRI-4 response rates at Week 52[70]

Endpoints, N/N (%)	Criteria I[a]			Criteria II[b]		
	Belimumab IV	Belimumab SC	Placebo	Belimumab IV	Belimumab SC	Placebo
SRI-4 response	313/596 (52.5)	159/246 (64.6)	188/530 (35.5)	398/738 (53.9)	269/421 (63.9)	282/731 (38.6)
≥4-point reduction in SELENA-SLEDAI	324/596 (54.4)	162/246 (65.9)	198/530 (37.4)	411/738 (55.7)	273/421 (64.8)	294/731 (40.2)
Severe SFI flare	100/596 (16.8)	35/246 (14.2)	155/530 (29.2)	116/738 (15.7)	52/421 (12.4)	190/731 (26.0)

[a]Low complement (C3 or C4) AND anti-dsDNA positive; [b]Low complement (C3 or C4) OR a SELENA-SLEDAI score ≥10.
ITC: indirect treatment comparison; IV: intravenous; SC: subcutaneous; SELENA: Safety of Estrogen in Lupus Erythematosus National Assessment-SLE Disease Activity Index; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index; SFI: SLE flare index; SLE: systemic lupus erythematosus; SRI-4: SLE Responder Index-4.

B.2.6.5.3 Phase 2 LBSL02 study and its LTE

A Phase 2, randomised, double-blind, placebo-controlled, dose-ranging study of belimumab was conducted to evaluate the safety, efficacy, and biologic activity of belimumab in patients with active SLE[62]. An integrated analysis of this Phase 2 study and Phase 3 IV LTE studies (BLISS-52 and BLISS-76) were used to estimate the year 2 onwards discontinuation rate in the economic models (see section B.3.3.4.2). Further details are provided in Appendix L. A ten-year continuation study was conducted in patients who had completed the Phase 2 study[63]. Further details are provided in Appendix M.

B.2.6.5.4 BASE study

Adverse events of special interest were specifically assessed in the post-marketing Phase 4 BASE safety trial. Further details are provided in Appendix F, with steroid use data presented in Appendix O.

B.2.6.5.5 EMBRACE study

Patients of black race have more severe SLE and more frequent lupus nephritis than other racial groups. EMBRACE was a multicentre, double-blind, placebo-controlled trial in patients of self-identified black race, aged ≥18 years, with active SLE. Further details are provided in Appendix O.

B.2.6.5.6 North East Asia

Patients with SLE of Eastern Asian origin may have a higher incidence of haematological disorders and kidney disease compared with European cohorts. This Phase 3, multicentre, randomised, double-blind, placebo-controlled study was conducted to assess the safety and efficacy of belimumab as an add-on to ST in patients with autoantibody-positive SLE in North East Asia. Further details are provided in Appendix O.

B.2.6.5.7 Treatment holiday study

At the time of TA397, there was no efficacy or safety data on the effects of temporary discontinuation of belimumab therapy in patients with stable low disease activity and subsequent reintroduction of belimumab therapy (so-called 'treatment holidays') or data on rebounds of SLE activity following belimumab cessation. This open-label, non-randomised, 52-week study investigated the potential for rebound upon temporary discontinuation of belimumab IV[78]. Further details are provided in Appendix O.

B.2.7 *Subgroup analyses*

This section provides results in two HDA populations:

- HDA-1 (SELENA-SLEDAI) score ≥10 AND low complement AND positive anti-dsDNA) – current NICE guidance population

- HDA-2 (SELENA-SLEDAI score ≥10 AND at least one of the following serological features: low complement OR positive anti-dsDNA). HDA-2 forms the base case of the economic analysis in Section B.3.2.1.

Previous subgroup analyses have demonstrated that patients with serological markers of high disease activity (low complement and anti-dsDNA antibodies) are more likely to respond to treatment with belimumab[79, 80]. However, the rheumatology community have highlighted that not all patients experiencing high disease activity will have both of these serological markers. Furthermore, some patients with high disease activity will have neither biomarkers. Therefore, clinical advisors to the company agreed that patients with a SELENA-SLEDAI score ≥10 and one of the serological biomarkers (anti-dsDNA antibodies or low complement) should be considered as having HDA. Consequently, patients who continue to have high disease activity who would potentially derive benefit from belimumab do not have access to this important SLE specific treatment option.

Real-world experience within the NHS has proven that the current NICE-approved HDA-1 population is overly restrictive as demonstrated by the slow recruitment into the BENLYSTA sub study of the BILAG-BR. The HDA-2 population in the BLISS trials (47.3 and 52.3% of the ITT population in the IV and SC populations, respectively) compared with the current NICE-approved HDA-1 population (35.2 and 31.6 % of the ITT population in the IV and SC populations, respectively)[45-47], still clearly defines a subgroup of patients that are more likely to benefit from belimumab compared with the overall ITT population enrolled in the BLISS trials. Amending the current NICE guidance to allow for a more clinically relevant subgroup (HDA-2 population) would allow more patients with high disease activity to benefit from treatment with belimumab without placing excessive strain on NHS resources.

Data for the NICE-approved HDA-1 subgroup based on pooled BLISS-52 and BLISS-76 trials were presented in TA397 ([redacted]; results from BLISS-SC were not available at the time and are therefore presented in this submission. Data for the HDA-2 subgroup have not been presented to NICE or published before [redacted].

B.2.7.1 Identification of HDA-1 and HDA-2 populations

Identification of the HDA populations through a series of pre-planned and post-hoc analyses of BLISS-52 and BLISS-76 was described in TA397. Briefly, baseline factors that are predictive of response at Week 52 irrespective of treatment received were identified using a logistic regression main effects model developed based on the pooled data from the Phase 3 BLISS-52 and BLISS-76 studies.

Among the model-derived predictors, SELENA-SLEDAI was chosen for the HDA definition as a direct measure of disease activity and the most significant predictor of Week 52 response. In addition, anti-dsDNA and complement were chosen, as they are objective, widely considered as important measures of disease activity, used routinely in SLE, and easily accessible to physicians. Furthermore, patients with positive anti-dsDNA or low complement are at an increased risk of flares[81].

B.2.7.2 Baseline characteristics of HDA-1 and HDA-2 populations

Baseline characteristics of HDA-1 and HDA-2 populations are summarised in Table 36 and Table 37, respectively.

Table 36. Baseline characteristics of HDA-1 population included in pivotal belimumab trials[45]

	BLISS-SC[45]		**Pooled BLISS-52 and BLISS-76 data[46, 47]**	
	Belimumab 200 mg SC N=186	**Placebo N=78**	**Belimumab 10 mg/kg IV N=193**	**Placebo N=203**
Demographics				
Female, N (%)	[redacted]	[redacted]	186 (96.4)	187 (92.1)
Age (years), mean (SD)	[redacted]	[redacted]	34.2 (9.6)	34.3 (10.6)
Ethnicity, N (%)				
White	[redacted]	[redacted]	77 (39.9)	90 (44.3)
Asian	[redacted]	[redacted]	57 (29.5)	45 (22.2)
Black	[redacted]	[redacted]	13 (6.7)	14 (6.9)
Alaska Native or American Indian from North/Central/ South America	[redacted]	[redacted]	46 (23.8)	54 (26.6)
Native Hawaiian or Other Pacific Islander	[redacted]	[redacted]	0	0
Multiracial	[redacted]	[redacted]	0	1 (0.5)
Disease characteristics				

	BLISS-SC[45]		Pooled BLISS-52 and BLISS-76 data[46, 47]	
	Belimumab 200 mg SC N=186	Placebo N=78	Belimumab 10 mg/kg IV N=193	Placebo N=203
SLE Disease duration (years), mean (SD)	[redacted]	[redacted]	6.38 (6.28)	7.04 (6.69)
BILAG organ domain involvement, N (%)				
At least 1A or 2B	[redacted]	[redacted]	136 (70.5)	143 (70.4)
At least 1A	[redacted]	[redacted]	32 (16.6)	39 (19.2)
No A or B	[redacted]	[redacted]	12 (6.2)	10 (4.9)
SELENA-SLEDAI category, N (%)				
10–11	[redacted]	[redacted]	81 (42.0%)	76 (37.4%)
≥12	[redacted]	[redacted]	112 (58.0)	127 (62.6%)
SELENA-SLEDAI score, mean (SD)	[redacted]	[redacted]	12.6 (3.3)	12.8 (3.3)
SFI, N (%)				
At least 1 flare	[redacted]	[redacted]	40 (20.7)	62 (30.5)
At least 1 severe flare	[redacted]	[redacted]	3 (1.6)	4 (2.0)
SDI score, mean (SD)	[redacted]	[redacted]	0.6 (1.0)	0.7 (1.2)
Proteinuria level (g/24 h), mean (SD)	[redacted]	[redacted]	0.7 (1.0)	0.7 (1.2)
Biomarker levels				
Anti-dsDNA (IU/mL), mean (SD)	[redacted]	[redacted]	152.9 (58.9)	149.3 (64.5)
C3 (mg/dL)	[redacted]	[redacted]	74.29 (22.80)	74.61 (23.88)
C4 (mg/dL)	[redacted]	[redacted]	9.8 (5.2)	10.0 (5.0)
Medication use, N (%)				
Steroid only	[redacted]	[redacted]	30 (16)	18 (9)
Immunosuppressant only	[redacted]	[redacted]	7 (4)	3 (1)
Anti-malarial only	[redacted]	[redacted]	8 (4)	6 (3)
Steroid and immunosuppressant	[redacted]	[redacted]	42 (22)	33 (16)
Steroid and anti-malarial	[redacted]	[redacted]	54 (28)	68 (33)
Immunosuppressant and anti-malarial	[redacted]	[redacted]	0	8 (4)
Steroid and immunosuppressant and anti-malarial	[redacted]	[redacted]	51 (26)	64 (32)
Average daily prednisone dose (mg/day) category, N (%)				
0	[redacted]	[redacted]	16 (8.3)	20 (9.9)
>0 to ≤7.5	[redacted]	[redacted]	51 (26.4)	57 (28.1)
>7.5	[redacted]	[redacted]	126 (65.3)	126 (62.1)
Average daily prednisone dose (mg/day)	[redacted]	[redacted]	12.3 (9.6)	11.6 (8.6)

[a]The large difference from IV studies is due to skewed data: median anti-dsDNA level for belimumab 200mg SC: [redacted] (range [redacted]). [b]The large difference from IV studies is due to skewed data: median level for placebo: [redacted] (range [redacted]).

dsDNA: double-stranded DNA; IV: intravenous; SC: subcutaneous; SELENA-SLEDAI: Safety of Estrogen in Lupus Erythematosus National Assessment-SLE Disease Activity Index; SFI: SLE flare index; SLE: systemic lupus erythematosus.

Table 37. Baseline characteristics of HDA-2 population included in pivotal belimumab trials[45]

	BLISS-SC[45]		Pooled BLISS-52 and BLISS-76 data[46, 47]	
	Belimumab 200 mg SC N=296	**Placebo N=141**	**Belimumab 10 mg/kg IV N=262**	**Placebo N=270**
Demographics				
Female, N (%)	[redacted]	[redacted]	[redacted]	[redacted]
Age (years), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Ethnicity, N (%)				
White	[redacted]	[redacted]	[redacted]	[redacted]
Asian	[redacted]	[redacted]	[redacted]	[redacted]
Black	[redacted]	[redacted]	[redacted]	[redacted]
Alaska Native or American Indian from North/Central/ South America	[redacted]	[redacted]	[redacted]	[redacted]
Native Hawaiian or Other Pacific Islander	[redacted]	[redacted]	[redacted]	[redacted]
Multiracial	[redacted]	[redacted]	[redacted]	[redacted]
Disease characteristics				
SLE Disease duration (years), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
BILAG organ domain involvement, N (%)				
At least 1A or 2B	[redacted]	[redacted]	[redacted]	[redacted]
At least 1A	[redacted]	[redacted]	[redacted]	[redacted]
No A or B	[redacted]	[redacted]	[redacted]	[redacted]
SELENA-SLEDAI category, N (%)				
10–11	[redacted]	[redacted]	[redacted]	[redacted]
≥12	[redacted]	[redacted]	[redacted]	[redacted]
SELENA-SLEDAI score, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
SFI, N (%)				
At least 1 flare	[redacted]	[redacted]	[redacted]	[redacted]
At least 1 severe flare	[redacted]	[redacted]	[redacted]	[redacted]
SDI score, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Proteinuria Category (g/24hr)				
≥2	[redacted]	[redacted]	[redacted]	[redacted]
Proteinuria level (g/24 h), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Clinical characteristics				
Low C3 and/or C4 n (%)				
No	[redacted]	[redacted]	[redacted]	[redacted]
Yes	[redacted]	[redacted]	[redacted]	[redacted]
Positive Anti-dsDNA n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Biomarker levels				

	BLISS-SC[45]		Pooled BLISS-52 and BLISS-76 data[46, 47]	
	Belimumab 200 mg SC N=296	Placebo N=141	Belimumab 10 mg/kg IV N=262	Placebo N=270
Anti-dsDNA (IU/mL), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
C3 (mg/dL)	[redacted]	[redacted]	[redacted]	[redacted]
C4 (mg/dL)	[redacted]	[redacted]	[redacted]	[redacted]
Medication use, N (%)				
Steroid only	[redacted]	[redacted]	[redacted]	[redacted]
Immunosuppressant only	[redacted]	[redacted]	[redacted]	[redacted]
Anti-malarial only	[redacted]	[redacted]	[redacted]	[redacted]
Steroid and immunosuppressant	[redacted]	[redacted]	[redacted]	[redacted]
Steroid and anti-malarial	[redacted]	[redacted]	[redacted]	[redacted]
Immunosuppressant and anti-malarial	[redacted]	[redacted]	[redacted]	[redacted]
Steroid and immunosuppressant and anti-malarial	[redacted]	[redacted]	[redacted]	[redacted]
Average daily prednisone dose (mg/day) category, N (%)				
0	[redacted]	[redacted]	[redacted]	[redacted]
>0 to ≤7.5	[redacted]	[redacted]	[redacted]	[redacted]
>7.5	[redacted]	[redacted]	[redacted]	[redacted]
Average daily prednisone dose, mg/day, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]

[a]The large difference from IV studies is due to skewed data: median anti-dsDNA level for belimumab 200mg SC: [redacted] (range [redacted]. [b]The large difference from IV studies is due to skewed data: median level for placebo: [redacted] (range [redacted].

dsDNA: double-stranded DNA; IV: intravenous; SC: subcutaneous; SELENA-SLEDAI: Safety of Estrogen in Lupus Erythematosus National Assessment-SLE Disease Activity Index; SFI: SLE flare index; SLE: systemic lupus erythematosus.

B.2.7.3 Results in HDA-1 and HDA-2 populations

B.2.7.3.1 Evidence from RCTs

The results of the primary endpoint analysis and its components, major secondary endpoints, and further key endpoints of interest for both BLISS-SC and pooled BLISS-52 and BLISS-76 trials are presented in Table 38 for the HDA-1 population and Table 39 for the HDA-2 population. In both HDA populations, the efficacy of belimumab was greater than compared with placebo, and this difference was more pronounced than in the overall ITT populations of the BLISS trials (See Section B.2.6 for BLISS-SC and Appendix L for pooled BLISS-52 and BLISS-76).

Table 38. SRI-4 responder rate and individual components at Week 52 in the HDA-1 population

	BLISS-SC		Pooled BLISS-52 and BLISS-76*	
	Placebo N=78	Belimumab 200 mg SC N=186	Placebo N=203	Belimumab 10 mg/kg IV N=193
SRI-4[a] (Primary endpoint)				
Response, N (%)	[redacted]	[redacted]	77 (37.9)	121 (62.7)
Observed difference vs placebo (%)	[redacted]	[redacted]	–	24.8
OR (95% CI) vs. placebo	[redacted]	[redacted]	–	2.7 (1.8, 4.1)
p-value	[redacted]	[redacted]	–	<0.0001
4-point reduction in SELENA-SLEDAI[a] (Primary endpoint component)				
Response, N (%)	[redacted]	[redacted]	84 (41.4)	125 (64.8)
Observed difference vs placebo (%)	[redacted]	[redacted]	-	23.4
OR (95% CI) vs. placebo	[redacted]	[redacted]	-	2.6 (1.7, 3.9)
p-value	[redacted]	[redacted]	-	<0.0001
No worsening in PGA[b] (Primary endpoint component)				
Response, N (%)	[redacted]	[redacted]	119 (58.6)	142 (73.6)
Observed difference vs placebo (%)	[redacted]	[redacted]	-	15.0
OR (95% CI) vs. placebo	[redacted]	[redacted]	-	2.0 (1.3, 3.1)
p-value	[redacted]	[redacted]	-	0.0015
No new 1A/2B BILAG domain scores[c] (Primary endpoint component)				
Response, N (%)	[redacted]	[redacted]	125 (61.6)	145 (75.1)
Observed difference vs placebo (%)	[redacted]	[redacted]	-	13.6
OR (95% CI) vs. placebo	[redacted]	[redacted]	-	1.9 (1.2, 3.0)
p-value	[redacted]	[redacted]	-	0.0034
SELENA-SLEDAI change from baseline				
Mean (SD or SE)	[redacted]	[redacted]	–4.1 (SE 0.3)	–5.8 (SE 0.3)
LS mean (SE)[d]	[redacted]	[redacted]	–4.9 (0.4)	–6.5 (0.4)
Treatment difference (95% CI) vs placebo[d]	[redacted]	[redacted]	–	–1.7 (–2.6, –0.7)
p-value[d]	[redacted]	[redacted]	–	0.0005
Time to first SFI flare[e]				
Patients with flare over 52 weeks, N (%)[f]	[redacted]	[redacted]	176 (86.7)	149 (77.2)
Median days (IQR)[g,j]	[redacted]	[redacted]	68.0 (range 1,368)	109.0 (range 1,329)
HR (95% CI) vs. placebo[h]	[redacted]	[redacted]	–	0.70 (0.56,0.88)
p-value[h]	[redacted]	[redacted]	–	0.0017
Time to first severe SFI flare[e] (major secondary endpoint)				
Patients with severe flare over 52 weeks, N (%)[f]	[redacted]	[redacted]	67 (33.0)	39 (20.2)

Median days (IQR or range)[g,j]	[redacted]	[redacted]	NA (range 5, 363)	NA (range 10, 366)
HR (95% CI) vs. placebo[h]	[redacted]	[redacted]	–	0.55 (0.37,0.81)
p-value[h]	[redacted]	[redacted]	–	0.0028
Prednisone reduction by ≥25% from baseline to ≤7.5 mg/day during weeks 40–52 in patients with baseline prednisone dose >7.5 mg/day (Major secondary endpoint)				
Patients with prednisone reduction to ≤7.5 mg/day n/N (%)[i]	[redacted]	[redacted]	9/126 (7.1)[h]	20/126 (15.9)[h]
OR (95% CI) vs placebo[j]	[redacted]	[redacted]	–	2.43 (1.05, 5.65)
p-value[j]	[redacted]	[redacted]	–	0.0389
FACIT-Fatigue Scale Score Change from Baseline				
Mean (SD or SE)	[redacted]	[redacted]	3.33 (SE 0.74)	4.90 (SE 0.82)
LS mean (SE) [d]	[redacted]	[redacted]	3.28 (0.88)	5.03 (0.88)
Treatment difference (95% CI) vs. placebo [d]	[redacted]	[redacted]	–	1.75 (–0.18, 3.67)
p-value [d]	[redacted]	[redacted]	–	0.0748
EQ-5D UK Score change from baseline				
Mean (SE)	[redacted]		0.10 (0.02)	0.11 (0.03)
LS mean (SE)[k]	[redacted]		0.09 (0.02)	0.12 (0.02)
Treatment difference (95% CI) vs. placebo[k]	[redacted]		–	0.03 (-0.02,0.08)
p-value[k]	[redacted]		–	0.2526

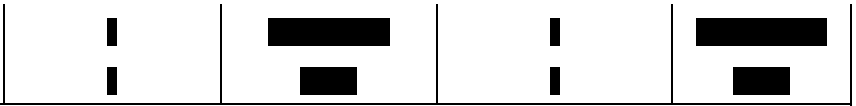
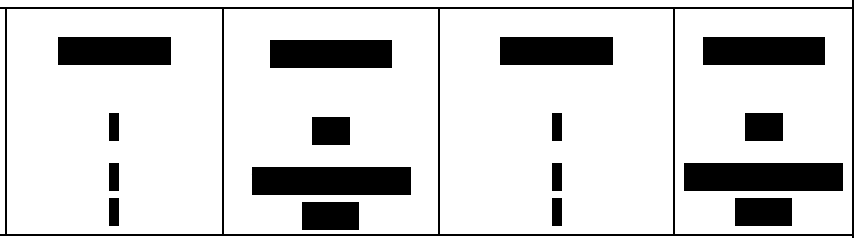
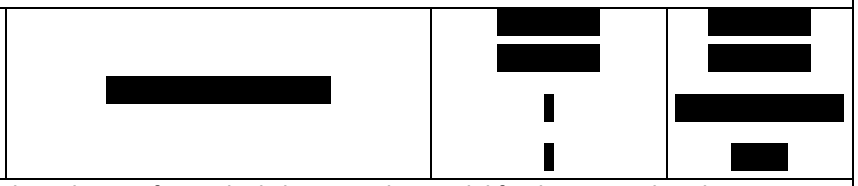
[*] Results re-presented from the previous NICE submission TA397, note this is based on the interim data for BLISS-76. [a]OR (95% confidence interval) and p-value are from a logistic regression model for the comparison between Belimumab and Placebo with covariates treatment group, baseline SELENA-SLEDAI score, race (Black vs other) and baseline proteinuria level (<2 g/24-hour vs ≥2 g/24-hour) For pooled data analysis, study was also included as an additional covariate. [b]Baseline PGA score is also included in the model. [c]Baseline BILAG domain involvement (at least 1A/2B versus at most 1B) is also included in the model. [d]All statistics are from an analysis of covariance (ANCOVA) model comparing Belimumab and Placebo with covariates for treatment group, baseline SELENA SLEDAI score, race (black vs. other), and baseline proteinuria level (<2 g/24 hour vs≥2 g/24 hour). [e]Severe flares that were triggered only by an increase in SELENA SLEDAI score to >12 are reported as mild/moderate flares if the change from the previous visit was at least three points and are excluded otherwise. Data censored at last available visit by week 52 visit. For patients who died, data are censored at death if no flares occurred before death. Time to first flare is defined as (event date – treatment start date + 1). [f]Only includes post-baseline flares. [g]Statistics will be missing when the number of events is too low to estimate the value. [h]From Cox proportional hazards model for the comparison between Belimumab and Placebo adjusting for baseline SELENA SLEDAI score, race (black vs. other), and baseline proteinuria level (<2 g/24 hour vs. ≥2 g/24 hour). [i]Includes only patients with baseline prednisone >7.5 mg/day. [j]Median days will be missing if the estimated probability of a flare is <50%.[k]From ANCOVA for the comparison between each belimumab dose and placebo, adjusted for the corresponding baseline EQ-5D score.

BILAG: British Isles Lupus Assessment Group; CI: confidence interval; HDA: high disease activity; HR: hazard ratio; IQR: interquartile range; NA: not available; OR: odds ratio; PGA: Physician's Global Assessment; SC: subcutaneous; SD: standard deviation; SE: standard error; SELENA: Safety of Estrogen in Lupus Erythematosus National Assessment; SFI: SLE Flare Index; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index; SRI-4: SLE responder index-4.

Table 39. SRI-4 responder rate and individual components at Week 52 in the HDA-2 population

	BLISS-SC	Pooled BLISS-52 and BLISS-76

	Placebo N=141	Belimumab SC 200 mg N=296	Placebo N=270	Belimumab IV 10 mg/kg N=262
SRI-4[a] (Primary endpoint)				
Response, N (%)	[redacted]	[redacted]	[redacted]	[redacted]
Observed difference vs placebo (%)		[redacted]		[redacted]
OR (95% CI) vs. placebo		[redacted]		[redacted]
p-value		[redacted]		[redacted]
4-point reduction in SELENA-SLEDAI[a] (Primary endpoint component)				
Response, N (%)	[redacted]	[redacted]	[redacted]	[redacted]
Observed difference vs placebo (%)		[redacted]		[redacted]
OR (95% CI) vs. placebo		[redacted]		[redacted]
p-value		[redacted]		[redacted]
No worsening in PGA[b] (Primary endpoint component)				
Response, N (%)	[redacted]	[redacted]	[redacted]	[redacted]
Observed difference vs placebo (%)		[redacted]		[redacted]
OR (95% CI) vs. placebo		[redacted]		[redacted]
p-value		[redacted]		[redacted]
No new 1A/2B BILAG domain scores[c] (Primary endpoint component)				
Response, N (%)	[redacted]	[redacted]	[redacted]	[redacted]
Observed difference vs placebo (%)		[redacted]		[redacted]
OR (95% CI) vs. placebo		[redacted]		[redacted]
p-value		[redacted]		[redacted]
SELENA SLEDAI change from baseline at Week 52				
Mean (SD or SE)	[redacted]	[redacted]	[redacted]	[redacted]
LS mean (SE) [d]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment difference (95% CI) vs. placebo[d]		[redacted]		[redacted]
p-value[d]		[redacted]		[redacted]
Time to first SFI flare[e]				
Subjects with flare over 52 weeks, N (%)[f]	[redacted]	[redacted]	[redacted]	[redacted]
Median days (IQR or range)[g]	[redacted]	[redacted]	[redacted]	[redacted]
HR (95% CI) vs. Placebo[h]		[redacted]		[redacted]
p-value[h]		[redacted]		[redacted]
Time to first severe SFI flare[e] (Major secondary endpoint)				
Subjects with severe flare over 52 weeks, N (%)[f]	[redacted]	[redacted]	[redacted]	[redacted]
Median days (IQR or range)[g]	[redacted]	[redacted]	[redacted]	[redacted]
HR (95% CI) vs. Placebo[h]		[redacted]		[redacted]
p-value[h]		[redacted]		[redacted]
FACIT-Fatigue Scale Score Change from Baseline				
Mean (SD or SE)	[redacted]	[redacted]	[redacted]	[redacted]
LS mean (SE)[i]	[redacted]	[redacted]	[redacted]	[redacted]

Treatment difference (95% CI) vs. placebo[i]	
p-value[i]	
Prednisone reduction by ≥25% from baseline to ≤7.5 mg/day during weeks 40–52 in patients with baseline prednisone dose >7.5 mg/day (Major secondary endpoint)	
Patients with prednisone reduction to ≤7.5 mg/day n/N (%), n[j]	
Observed difference vs. placebo (%)	
OR (95% CI) vs. Placebo[k]	
P-value[k]	
EQ-5D UK Score change from baseline	
Mean (SE)	
LS mean (SE)[l]	
Treatment difference (95% CI) vs. placebo[l]	
p-value[l]	

[a]OR (95% confidence interval) and p-value are from a logistic regression model for the comparison between Belimumab and Placebo with covariates treatment group, baseline SELENA-SLEDAI score, race (Black vs other) and baseline proteinuria level (<2 g/24-hour vs ≥2 g/24-hour) For the pooled IV data analysis, study was also included as an additional covariate. [b]Baseline PGA score is also included in the model. [c]Baseline BILAG domain involvement (at least 1A/2B versus at most 1B) is also included in the model. [d]All statistics are from an analysis of covariance (ANCOVA) model comparing Belimumab and Placebo with covariates for 1) BLISS-SC: treatment group, baseline SELENA SLEDAI score, race (black vs. other), and baseline proteinuria level (<2 g/24 hour vs≥2 g/24 hour) 2) pooled IV trials: baseline proteinuria level (<2 g/24 hour vs. ≥2 g/24 hour equivalent), race (African descent or indigenous-American descent vs. other) and study (BLISS-52 vs BLISS-76). [e]Severe flares that were triggered only by an increase in SELENA SLEDAI score are reported as mild/moderate flares if the change from the previous visit was at ≥3 points and are excluded otherwise. Data censored at last available visit by week 52 visit. For subjects who died, data are censored at death if no flares occurred before death. Time to first flare is defined as (event date – treatment start date + 1). [f]Only includes post-baseline flares. [g]Statistics will be missing when the number of events is too low to estimate the value. [h]From Cox proportional hazards model for the comparison between Belimumab and Placebo adjusting for 1) BLISS-SC: baseline SELENA SLEDAI score, race (black vs. other), and baseline proteinuria level (<2 g/24 hour vs. ≥2 g/24 hour) 2) pooled IV trials: baseline proteinuria level (<2 g/24 hour vs. ≥2 g/24 hour equivalent), race (African descent or indigenous-American descent vs. other) and study (BLISS-52 vs BLISS-76). [i]All statistics are from an analysis of covariance (ANCOVA) model comparing Belimumab and Placebo with covariates 1) BLISS-SC: treatment group, baseline FACIT-Fatigue Scale score, baseline SELENA SLEDAI score, race (black vs. other), and baseline proteinuria level (<2 g/24 hour vs. ≥2 g/24 hour) 2) pooled IV trials: baseline FACIT-Fatigue score, baseline proteinuria level (<2 g/24 hour vs. ≥2 g/24 hour equivalent), race (African descent or indigenous-American descent vs. other) and study (BLISS-52 vs BLISS-76). [j]Includes only subjects with baseline prednisone > 7.5 mg/day. All corticosteroids are converted to a prednisone equivalent average daily dose (mg/day). [k]From a logistic regression model for the comparison between belimumab and placebo with covariates including 1) for BLISS-SC: treatment group, baseline prednisone dose, baseline SELENA SLEDAI score, and race (black vs. other) and 2) for pooled IV trials: baseline prednisone dose, baseline proteinuria level (<2 g/24 hour vs. ≥2 g/24 hour equivalent), race (African descent or indigenous-American descent vs. other) and study (BLISS-52 vs BLISS-76). [l]From ANCOVA for the comparison between Belimumab and Placebo, adjusted for the corresponding baseline EQ-5D score

BILAG: British Isles Lupus Assessment Group; CI: confidence interval; HDA: high disease activity; HR: hazard ratio; IQR: interquartile range; NA: not available; OR: odds ratio; PGA: Physician's Global Assessment; SC: subcutaneous; SD: standard deviation; SE: standard error; SELENA: Safety of Estrogen in Lupus Erythematosus National Assessment; SFI: SLE Flare Index; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index; SRI-4: SLE responder index-4.

B.2.7.4 Real-world evidence: BILAG-BR

B.2.7.4.1 BILAG-BR: Overview

The BILAG Biologics Register (BILAG-BR) is a multicentre prospective cohort study that has been ongoing since March 2010 at hospitals throughout the UK. Its main objective is to investigate the safety of biologics in the treatment of SLE. The BILAG-BR collects information on patient demographics, disease severity, quality of life, and safety measures of biologic (and biosimilar) and other non-biologic treatments for patients presenting with SLE. Patients join the study if they have a diagnosis of SLE and are aged ≥5 years. Clinical assessment is performed prior to start of treatment (baseline), at 3, 6, and 12 months, and annually thereafter. The Benlysta sub-study collected the same data as the main BILAG-BR study, providing information on real-world effectiveness, safety, and quality of life for all patients prescribed belimumab in England. Eligibility for treatment was defined by the HDA-1 subgroup criteria. This study fulfilled the requirements of the managed access agreement that resulted from TA397.

Three cohorts were considered as part of this study, comprising patients who received belimumab, rituximab, and non-biologics. Patients could stop treatment at any point during follow up and switch to other treatment or restart the same one. This is identified as a second or subsequent round of treatment, resulting in the potential for multiple “rounds” of treatment per participant. However, baseline characteristics were only captured at enrolment into the registry and were not updated before a second or subsequent treatment round, so that comparisons are made against baseline of “round 1” (a patient’s initial registry treatment). Results reported in the tables in Appendix P are provided for round 1 (with patients as denominator) and/or any treatment round (with patient-rounds as denominator) for baseline data, efficacy, safety and QoL data.

As outlined in the report (provided as Appendix P), there is a high likelihood of confounding, including selection bias in the treatment groups, so that the data captured is not suitable to test the causal efficacy of the treatment or compare treatment efficacy. As described below, there were also substantial differences in

follow-up duration and cohort size, making a meaningful comparison of the two treatments difficult. We have therefore focused on reporting the BILAG-BR data for the belimumab cohort herein and provided the full report (including results for the non-biologic and rituximab cohorts) in Appendix P.

B.2.7.4.2 BILAG-BR: Cohort size

Between March 2010 and July 2020, [redacted] patients were included in the BILAG-BR of which [redacted] received rituximab, [redacted] received a non-biologic treatment and [redacted] received belimumab. There were a total of [redacted] rounds of treatment, of which [redacted] rounds of treatment were with rituximab, [redacted] rounds with belimumab, and [redacted] with a non-biologic. For belimumab, this meant that [redacted] patients had another treatment (either rituximab, a non-biologic, or both) before switching to belimumab. When the October 2013 date is taken as a start date for the analysis (as detailed in the data analysis plan), which is based on when the NHS England interim rituximab SLE policy was published, and after applying eligibility criteria for the HDA-1 subgroup to all cohorts, the resultant patients numbers are [redacted], [redacted] and [redacted] for the non-biologic, rituximab and belimumab cohorts respectively considering all rounds of treatment. However, in total 85 distinct patients received belimumab of which, 1 patient received a second round of belimumab, hence 86 patient rounds. For this one patient, only baseline data was recorded and therefore there is no contribution to the presented efficacy, safety and patient reported outcomes.

Table 40. BILAG-BR: Derivation of the BILAG-BR treatment cohorts

Cohort derivation step	Patients at baseline of any round				Patients at baseline of round 1			
	Belimumab, N (%)	Rituximab, N (%)	Non-biologic, N (%)	Total patient-rounds	Belimumab, N (%)	Rituximab, N (%)	Non-biologic, N (%)	Total patients
All patient-rounds	[redacted]							
Classified into treatment groups*	[redacted]							
Registry entry from 1st Oct 2013**	[redacted]							
No BILAG renal A score†	[redacted]							
No BILAG CNS A score†	[redacted]							
Positive anti-dsDNA test	[redacted]							
Low C3 or C4	[redacted]							
SLEDAI ≥10	[redacted]							

BILAG: British Isles Lupus Assessment Group; CNS: central nervous system; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index

**There were [redacted] patients at round 1 and [redacted] patient-rounds at any round who received other biologic treatment or with missing baseline data entry; these patients/ patient-rounds were not analysed further*

*** A small number of patients had entered the registry prior to June 2016 (MAA as part of TA 397) †Used as a proxy to identify patients with CNS lupus and lupus nephritis, respectively, who currently fall outside of the license for belimumab*

Table 41. BILAG-BR: Breakdown of belimumab treatment rounds

Patient group	Treatment round							Total
	1st	2nd	3rd	4th	5th	6th	7th	
Number of patients receiving belimumab at any round	[redacted]							

B.2.7.4.3 BILAG-BR: Follow-up duration

The median date of study entry was May 2018 for belimumab-treated patients, considerably later than for both non-biologic (March 2017) and rituximab (October 2016) cohorts, so that differences in available follow-up should be taken into account when interpreting the results provided here and in Appendix P. For example, longer follow-up may translate into additional treatment rounds being recorded, making it difficult to compare second or later rounds of treatment.

B.2.7.4.4 BILAG-BR: Disease activity in patients receiving belimumab

Three disease activity measures were assessed: BILAG-2004, SLEDAI-2K, and the SDI. This data is summarised in Table 42, restricted to 12 months of follow-up, in line with the available data for patients receiving belimumab. [redacted] in SLEDAI-2K and BILAG scores were observed at [redacted] and were [redacted] Average within-person change in the belimumab group was [redacted] points for BILAG, and [redacted] points for SLEDAI-2K at 12 months. A 4-point reduction in SLEDAI-2K, considered a clinically important effect, was observed in [redacted] of patients at 3 months and was [redacted] (Table 43).

Table 42. BILAG-BR: BILAG, SLEDAI-2K, and SDI in patients receiving belimumab (any treatment round)

SLE Activity Measures	Baseline	Within Follow Up						Within-Person Change					
		3mths		6mths		12mths		3mths		6mths		12mths	
	N Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)
BILAG score	[redacted]												
Average SLEDAI-2K score	[redacted]												
SDI score	[redacted]												

Table 43. BILAG-BR: Number and percentage of patient-rounds on belimumab with ≥4 point reduction in SLEDAI-2K from baseline (any treatment round)

3mths		6mths		12mths	
Patient-rounds, N	**≥4-point reduction, N (%)**	**Patient-rounds, N**	**≥4-point reduction, N (%)**	**Patient-rounds, N**	**≥4-point reduction, N (%)**
[redacted]					

The mean SDI score changed from [redacted] at baseline to [redacted] at 12 months follow up and the equivalent within-patient change was [redacted] (Table 42). At baseline [redacted] of patients had a [redacted] SDI score; this was [redacted] to [redacted] at 12 months. The within-patient change in SDI score indicated that [redacted] of belimumab patients reported a [redacted] in SDI score at 12 months, and [redacted] occurred in approximately [redacted] (Table 42). However, SDI score should be interpreted with caution. The relatively short follow-up in the belimumab group limits the conclusions that can be drawn from the data, as organ damage accumulates over several years and is best observed in longer-term studies.

B.2.7.4.5 BILAG-BR: Flares in belimumab-treated patients

In the BILAG-BR, flares were identified using the BILAG score and a flare was defined as a score of A in one or more additional (i.e. not present at baseline) domains, or a score of B in 2 or more additional domains. This definition is the same as the definition of the BILAG component of SRI-4, the primary endpoint of pivotal belimumab RCTs. It is, however, different from the definition of flares used in the time to first severe flare secondary endpoint of pivotal belimumab RCTs, where flares were identified using the SFI.

The rate of new flares was [redacted] patient-rounds/years. In other words, for [redacted] every 100 patients receiving belimumab, there would be on average [redacted] new flares per year. The number of new flares identified at the 3, 6, and 12-month follow-up time points is shown in Table 44.[redacted]

Table 44. BILAG-BR: BILAG flares in belimumab-treated patients compared to baseline (any treatment round)

Follow up time point	Patients, N	BILAG flares (% of patient at time point)
Baseline	[redacted]	[redacted]
3mths	[redacted]	[redacted]
6mths	[redacted]	[redacted]
12mths	[redacted]	[redacted]
Total over 12 months	[redacted]	[redacted]

B.2.7.4.6 BILAG-BR: Quality of life of belimumab-treated patients

Quality of life of BILAG-BR patients was measured using the LupusQoL, the SF36, and the EQ-5D. Responses were available for [redacted] of patients at baseline and for [redacted] at 12 months.

B.2.7.4.6.1. LupusQoL

LupusQoL reports patient quality of life for eight domains relating to quality of life in terms of Physical, Pain, Planning, Intimate, Burden to others, Emotional Health, Body Image, and Fatigue. In each domain, a greater score indicates improved quality of life. The number of responses, the mean score, and standard deviation for

the individual LupusQoL domain scores at baseline and follow up (3, 6, and 12 months) are reported in Table 45.

The mean within-person change over 12 months of follow-up shows that, on average, patients experienced an [redacted] across [redacted] LupusQoL domains [redacted] and [redacted] for [redacted] from [redacted], which is considered a clinically meaningful improvement[82].

B.2.7.4.6.2. SF36

The SF-36 reports quality of life scores for eight domains: Social, Role Health, Role Emotion, Physical, Pain, General Health, and Energy. In each domain, an increase in score indicates improved quality of life. The number of patients with records at [redacted] was [redacted] so the results [redacted] domains [redacted] showed [redacted] particularly the [redacted] domain, where a mean [redacted] was observed. The results are reported in Table 46.

B.2.7.4.6.3. EQ-5D

The EQ-5D health status is a score between 0 and 100 with a greater score indicating better health status. Treatment with belimumab was associated with an [redacted] (Table 47). The average within patient change was [redacted]

Table 45. BILAG-BR: LupusQoL scores in belimumab-treated patients

LUPUS-QoL domain	Within Follow Up N; Mean (sd)				Within Person Change N; Mean (sd)		
	Baseline	3mths	6mths	12mths	3mths	6mths	12mths
Physical	[redacted]						
Pain							
Planning							
Intimate relationships							
Burden to Others							
Emotional Health							
Body Image							
Fatigue							

Table 46. BILAG-BR: SF36 health survey scores in belimumab-treated patients

SF36 domain	Within Follow Up N; Mean (sd)				Within Person Change N; Mean (sd)		
	Baseline	3mths	6mths	12mths	3mths	6mths	12mths
Social score	[redacted]						
Role Health Score							
Role Emotion							
Physical score							
Pain score							
General health score							
Energy score							
Emotional Score							

Table 47. BILAG-BR: EQ-5D health status

Within Follow Up N; Mean (sd)				Within Person Change N; Mean (sd)		
Baseline	3mths	6mths	12mths	3mths	6mths	12mths
[redacted]						

B.2.7.4.7 Steroid use

Both regular and irregular steroid use was assessed; regular use representing treatment according to a regular dosage schedule (e.g. once a day for three weeks, or three times a week for 6 weeks) and irregular use representing one-off dosing. Regular use is presented below and for details of irregular use, please refer to full report in Appendix P. Regular use was expressed as equivalent daily dose, and the median averaged over the number of different regimes that patients received during the follow up period.

There was a [redacted] in regular steroid dose at 3 months and 12 months on treatment with belimumab (Table 48). However, the average dose reported at 6 months was [redacted] Overall [redacted] In each case, [redacted] the mean (SD) dose at baseline [redacted] at 6 months to [redacted] and the within person change [redacted]

Table 48. BILAG-BR: Steroid treatment in belimumab-treated patients

	Within time-point N; mean (sd)				Within Person Change N; mean (sd)		
	Baseline	**3mths**	**6mths**	**12mths**	**3mths**	**6mths**	**12mths**
Regular Treatment Dose (mg)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

B.2.8 *Meta-analysis*

No meta-analysis was performed as part of this submission.

B.2.9 *Indirect and mixed treatment comparisons*

There are no studies directly comparing belimumab with rituximab. Differences in the endpoints and the patient populations preclude the conduct of any meaningful indirect and mixed treatment comparisons between belimumab and rituximab. For example, the inclusion criteria of the published Phase 2/3 randomised, double-blind study of rituximab required SLE patients to have significantly active disease at screening [83] likely to correspond to a more severe patient population than the Phase 3 belimumab trials. Also, changes in SELENA-SLEDAI, an important short-term outcome which can be linked to longer term impact on organ damage, were not collected in the rituximab trial, making an indirect comparison difficult.

Furthermore, in July 2020, NHS England published a clinical commissioning policy 'Rituximab for refractory Systemic Lupus Erythematosus (SLE) in adults and post-pubescent children [200402P]'[44] in which the rituximab implementation criteria states:

"Rituximab should be considered for adults and post-pubescent children with moderate or severe refractory SLE with active disease, who have failed to respond or have had adverse events to 2 or more immunosuppressive therapies (one of which must be either mycophenolate or cyclophosphamide, unless contraindicated) and have:

EITHER

• Disease activity with at least one BILAG A and/or two B scores or a SLEDAI-2K score > 6

Or

• Requiring unacceptably high levels of oral glucocorticoids e.g. more than 7.5mg prednisolone in an adult per day, to maintain a lower disease activity state

AND

• been assessed as not eligible for clinical trials or belimumab."

Therefore, rituximab should be considered after eligibility for belimumab has been assessed (based on the criteria of TA397) and ruled out, i.e. a sequential approach to treatment is recommended with belimumab followed by rituximab, therefore making rituximab not a direct comparator. At the same time, GSK remains mindful of current clinical practice, where patients with more severe, highly active SLE and usually managed in tertiary centres, do have access to rituximab.

However, assessing the benefit of rituximab is problematic; it failed the primary endpoint in its Phase 2/3 EXPLORER study; and published observational data (including many studies with small numbers of patients, which may incorporate clinician selection bias) report variable levels of benefit with this medicine, which is used off-label for SLE. An indirect comparison with rituximab based on data available from the literature, as previously outlined in TA397, was therefore not considered appropriate and has not been incorporated into the cost-effectiveness model. Despite the updated systematic review of clinical evidence, high-quality data that permits a reliable indirect comparison is lacking. In lieu of an indirect comparison of RCTs, real-world comparative evidence was sought from the BILAG-BR. As part of the analysis of the BILAG-BR, the University of Manchester did undertake a multilevel regression modelling exercise to explore the patient outcomes across the three cohorts: belimumab, rituximab and non-biologic. In these regression models, treatment effect estimates are compared with rituximab due to the availability of the largest sample size and results are reported as effect coefficients. The results suggest that for most outcome measures, a similar level of change was observed between belimumab and rituximab. However, the regression modelling could only be conducted out to 12 months because of the limited follow-up data available for belimumab patients in this study. It remains that reducing the risk of long-term organ damage is a key treatment goal for SLE patients and of most interest to clinicians. Whilst there is published data to support this for belimumab, there is limited equivalent evidence for rituximab (see Section 2.7.4.1).

A formal indirect treatment comparison of rituximab and belimumab based on BILAG-BR data was not conducted due to the observational, exploratory nature of the data and the differences in cohort sizes, patient characteristics and duration of follow-up.

B.2.10 *Adverse reactions*

Note: The safety profile of belimumab was consistent across all studies conducted and is comprehensively described herein, so that no additional adverse reactions are described in Appendix F.

B.2.10.1 Overview of AEs and SAEs

The safety of belimumab in patients with SLE has been evaluated in 3 pre-registration placebo-controlled IV studies (Phase 2 LBSL01 study [Appendix L], BLISS-52 [Table 5] and BLISS-76 [Table 6]), 1 placebo-controlled SC study (BLISS-SC [Table 4]), and one post-marketing, placebo-controlled IV study (BASE [Appendix F])[6]. Overall, adverse reactions were reported in 87% of belimumab-treated patients and 90% of placebo-treated patients[6]. The most frequently reported adverse reactions (≥5% of patients with SLE treated with belimumab plus ST and at a rate ≥1% greater than placebo) were viral upper respiratory tract infections, bronchitis, and diarrhoea. The proportion of patients who discontinued treatment due to adverse reactions was 7% for belimumab-treated patients and 8% for placebo-treated patients[6].

In the three Phase 3 placebo-controlled studies, belimumab in combination with ST therapies had an overall safety profile that was similar to placebo plus ST with regard to frequency, severity, and types of AEs (Table 49). Further safety data from the pivotal Phase 3 trials are provided in Appendix F.

B.2.10.1.1 Discontinuation of belimumab due to AEs

Phase 3 RCTs: in BLISS SC, the overall incidence of AEs leading to discontinuation of study agent was 8.9% for the placebo group and 7.2% for the belimumab 200 mg SC group (Table 49). The most common system organ class for AEs leading to

discontinuation in BLISS SC was infections and infestations (2.5%) for the placebo group and renal and urinary disorders (1.4%) for the belimumab group[45]. In BLISS 76, study agent was discontinued for 7.8% of all patients due to 1 or more AEs, most frequently nervous system disorders (1.7%), general disorders and administration site conditions (1.1%), and renal and urinary disorders (1.1%). Neither a treatment effect nor a belimumab dose relationship was apparent[47]. In BLISS 52, study agent was discontinued for 5.8% of all patients due to 1 or more AEs, most frequently renal and urinary disorders (1.3%), and infections and infestations (0.7%). Neither a treatment effect nor a belimumab dose relationship was apparent[46]. Appendix F provides further details of AEs leading to discontinuation of study drug in BLISS-SC, BLISS-52, and BLISS-76.

Long-term extension studies: in BLISS-SC LTE, 17 (2.6%) patients discontinued study drug in the open-label phase due to AEs[49]. In BLISS-76 US LTE, 26 (10%) patients discontinued study drug or withdrew from the study due to an AE, the majority during the first four interval years[51]. In BLISS-52/76 non-US LTE, 69 (9%) patients discontinued study drug or withdrew from the study due to an adverse event, most during the first three interval years[50].

Real-world evidence studies: In the OBSErve pooled analysis of Argentina, Germany, Spain, and Switzerland a total of 12 of 227 patients discontinued treatment with belimumab before month 6, 3 of these (25%) due to AEs. Other reasons for discontinuation included death, disease progression, lack of compliance, inefficacy, and patient request[68].

B.2.10.1.2 Deaths

In BLISS SC, the overall incidence of death for the belimumab 200 mg SC group (0.5%) was similar to that for the placebo group (0.7%) (Table 49). Two deaths in the belimumab group (one tuberculosis of the central nervous system, the other pneumonia following urosepsis) were judged to be possibly related to study agent[45]. In BLISS 76, 3 patients died during the study (2 in the belimumab 1 mg/kg, group, and 1 in the belimumab 10 mg/kg group), all judged to be not related to study

agent[47]. In BLISS 52, 9 patients died during the study, 3 in the placebo group, 2 in the belimumab 1 mg/kg group, and 4 in the belimumab 10 mg/kg group. Two deaths in the 10 mg/kg belimumab group (bacterial sepsis and infectious diarrhoea) and 1 death in the placebo group (myocardial infarction) were considered possibly or probably related to study agent[46].

Table 49. Number of patients with treatment-emergent AEs in Phase 3 trials of belimumab[45-47]

	BLISS-SC		BLISS 76			BLISS 52		
	Belimumab 200 mg SC N=556	**Placebo N=280**	**Belimumab 1 mg/kg IV N=271**	**Belimumab 10 mg/kg IV N=273**	**Placebo N=275**	**Belimumab 1 mg/kg IV N=288**	**Belimumab 10 mg/kg IV N=290**	**Placebo N=287**
AE, N (%)	449 (80.8)	236 (84.3)	253 (93.4)	253 (92.7)	253 (92.0)	264 (91.7)	266 (91.7)	263 (91.6)
Related AE, N (%)	173 (31.3)	73 (26.1)	120 (44.3)	104 (38.1)	123 (44.7)	91 (31.6)	105 (36.2)	113 (39.4)
SAE, N (%)	60 (10.8)	44 (15.7)	63 (23.2)	61 (22.3)	54 (19.6)	47 (16.3)	41 (14.1)	36 (12.5)
Severe AE, N (%)	55 (9.9)	40 (14.3)	51 (18.8)	54 (19.8)	52 (18.9)	36 (12.5)	33 (11.4)	34 (11.8)
Serious and/or severe AE, N (%)	82 (14.7)	59 (21.1)	76 (28.0)	82 (30.0)	72 (26.2)	57 (19.8)	50(17.2)	48 (16.7)
AE resulting in study agent discontinuation, N (%)	40 (7.2)	25 (8.9)	18 (6.6)	23 (8.4)	23 (8.4)	16 (5.6)	15 (5.2)	19 (6.6)
Death, N (%)	3 (0.5)	2 (0.7)	2 (0.7)	1 (0.4)	0 (0.0)	2 (0.7)	4 (1.4)	3 (1.0)
Severe refers to Grade 3 and Grade 4. *AE: adverse event; IV: intravenous; SAE: serious adverse event; SC: subcutaneous.*								

B.2.10.2 Safety of belimumab in the Phase 4 BASE study – Adverse events of Special Interest

The BASE study was a double-blind, placebo-controlled, randomised (1:1), Phase 4 safety study to evaluate all-cause mortality and adverse events of special interest (AESI) in adults with SLE receiving belimumab IV 10 mg/kg versus placebo over 52 weeks[64]. Differences in rates of mortality and other pre-specified AESI (malignancies, serious infections, opportunistic infections and other infections of interest, serious depression, suicidality, and serious infusion/hypersensitivity reactions) on-treatment (first to last dose +28 days) were assessed[64].

A total of 4,003 patients received at least 1 dose of trial medication. Overall rates of on-treatment AESIs were similar between groups, except for serious depression and serious infusion/hypersensitivity reactions, which were more frequently reported in the belimumab IV group[64] (Table 50).

Table 50. BASE study: Pre-specified AESI endpoints

	Placebo N=2001	**Belimumab 10 mg/kg IV N=2002**	**Difference (%) versus placebo (95% CI)**
Deaths, N (%)	8 (0.40)	10 (0.50)	0.10 (–0.31, 0.51)
Serious infections, N (%)	82 (4.10)	75 (3.75)	–0.35 (–1.55, 0.85)
Opportunistic infections and other infections of interest, N (%)	50 (2.50)	36 (1.80)	–0.70 (–1.60, 0.20)
Malignancies (excluding NMSC), N (%)	5 (0.25)	5 (0.25)	0 (–0.31, 0.31)
NMSC, N (%)	3 (0.15)	4 (0.20)	0.05 (–0.21, 0.31)
Serious depression, N (%)	1 (0.05)	7 (0.35)	0.30 (0.02, 0.58)
Suicidality[a] (C-SSRS), N (%)	23 (1.16)	28 (1.42)	0.26 (–0.44, 0.96)
Serious infusion, hypersensitivity reactions, N (%)	2 (0.10)	8 (0.40)	0.30 (–0.01, 0.61)

[a]Treatment-emergent suicidal ideation/behaviour.
AESI: adverse events of special interest; CI: confidence interval; C-SSRS: Colombia Suicide Severity Rating Scale; NMSC: non-melanoma skin cancer.

On-treatment deaths were most frequently caused by infection (3 [0.15%] placebo versus 9 [0.45%] belimumab); on-study deaths occurred in 22 (1.10%) placebo and 13 (0.65%) belimumab patients (difference [95% CI]: –0.45 [–1.03, 0.13]). However, fatal infections (e.g. pneumonia and sepsis) occurred in 0.45% of belimumab-treated

patients vs 0.15% of placebo-treated patients. Most fatal infections were observed during the first 20 weeks of treatment with belimumab.

On-treatment serious suicidal ideation/behaviour and self-injury events were reported for 5 (0.25%) placebo and 15 (0.75%) belimumab patients (difference [95% CI]: 0.50 [0.06, 0.94]); on-study suicidal ideation/behaviour occurred in 39 (1.96%) placebo and 48 (2.43%) belimumab patients (difference [95% CI]: 0.47 [–0.44, 1.38]). No suicide-related deaths were reported[64]. Further details of AESI recorded in the BASE study are provided in Appendix F.

B.2.10.3 Real-world safety experience

Safety assessment was not among the objectives of the pooled analysis of OBSErve studies and the individual publications report limited safety data. Available safety information from OBSErve studies can be provided upon request.

B.2.10.3.1 Safety data from the BILAG-BR registry

Safety data beyond 12 months on treatment were available for very few patients treated with belimumab, so that 12-month data is presented. The number of AEs recorded over 12 months of follow-up was very low. SAEs of interest are listed below:

- [redacted].

The frequency of specific SAE types of interest is provided in Table 51. Over 12 months, ≥1 SAE was observed for [redacted] of belimumab patient-rounds. However, [redacted] experienced SAEs due to their SLE ([redacted] or due to belimumab ([redacted]; infections were also rare [redacted]

Hospitalisations due to SLE and any cause are provided in Table 52. Average hospital stay due to SLE was [redacted], except for the 3-month follow-up,

where a [redacted]. When all hospitalisations were considered, mean hospital stay [redacted]. Overall, however, few hospitalisations occurred among belimumab-treated patients in the BILAG-BR.

Table 51. BILAG-BR: SAEs of interest in BILAG-BR patients receiving belimumab (any treatment round)

	Follow-up point	Patient-rounds with ≥1 SAE, N (%)	Total patient-rounds
Any SAE	Baseline	[redacted]	[redacted]
	3mths	[redacted]	[redacted]
	6mths	[redacted]	[redacted]
	12mths	[redacted]	[redacted]
SAE due to SLE	Baseline	[redacted]	[redacted]
	3mths	[redacted]	[redacted]
	6mths	[redacted]	[redacted]
	12mths	[redacted]	[redacted]
Hospitalisation due to SLE	Baseline	[redacted]	[redacted]
	3mths	[redacted]	[redacted]
	6mths	[redacted]	[redacted]
	12mths	[redacted]	[redacted]
Infection	Baseline	[redacted]	[redacted]
	3mths	[redacted]	[redacted]
	6mths	[redacted]	[redacted]
	12mths	[redacted]	[redacted]
SAE due to biologic	Baseline	[redacted]	[redacted]
	3mths	[redacted]	[redacted]
	6mths	[redacted]	[redacted]
	12mths	[redacted]	[redacted]

SAE: serious adverse event

Table 52. BILAG-BR: Duration of hospital stay due to SLE and due to any cause (any treatment round)

Follow up	Any Treatment Round		
	N	Mean (SD)	Max
Hospitalisation due to SLE			
Baseline	[redacted]	[redacted]	[redacted]
3mths	[redacted]	[redacted]	[redacted]
6mths	[redacted]	[redacted]	[redacted]
12mths	[redacted]	[redacted]	[redacted]
Any hospitalisation			
Baseline	[redacted]	[redacted]	[redacted]
3mths	[redacted]	[redacted]	[redacted]
6mths	[redacted]	[redacted]	[redacted]
12mths	[redacted]	[redacted]	[redacted]

B.2.11 *Ongoing studies*

Ongoing studies of belimumab in SLE, for which the results are not yet available, are listed in Table 53. Please note that studies for non-SLE indications are not included, as they are not directly relevant to the current appraisal. Further, BLISS-LN study evaluating the efficacy and safety of IV belimumab in patients with active lupus nephritis completed in March 2020 and was recently published in the New England Journal of Medicine (September 2020). However, this is also out of scope of this appraisal.

Table 53. Ongoing studies of belimumab in patients with SLE

Study Name	Phase	Study Type	Study Description	Belimumab formulation	Estimated study completion	[redacted]
BLISS-BELIEVE	IIIA	Interventional Clinical Trial	Phase 3, 104-week, safety and efficacy study of belimumab-rituximab combination in patients with SLE	SC	July 2021	[redacted]
SABLE	NA	Observational registry established as a post-marketing commitment	Multicentre, prospective, observational cohort study to evaluate the incidence of AESI and effectiveness in SLE patients	Either	Jan 2025	[redacted]
114256	NA	Pregnancy registry – post-marketing commitment	A registry to investigate the safety of belimumab in pregnancy. Due to the very slow recruitment, GSK is currently in discussions with the EMA around alternative relevant studies that could be conducted.	Either	Nov 2021	[redacted]
116559	NA	Meta-analysis of the elderly SLE patients – post-marketing commitment	Meta-analysis conducted under study ID BEL116559 to assess belimumab efficacy and safety in elderly patients treated in selected belimumab studies. This is a post-marketing commitment with the EMA.	Either	Dec 2025	[redacted]
BASE	IV	Interventional Clinical Trial	Global, multicentre, placebo-controlled RCT to evaluate AESI in SLE patients treated with belimumab. Primary analysis of this study is now complete.	IV	Aug 2022	[redacted]

PLUTO-SC	IIA	Interventional Clinical Trial	PK-PD study in paediatric patients with SLE	SC	[redacted]
PLUTO	IIB	Interventional Clinical Trial	Safety, PK and efficacy study of belimumab in paediatric patients with SLE. Please note that the primary analysis of this study is complete, and the results are provided in Appendix O. The open-label continuation phase and safety follow-up are still ongoing.	IV	[redacted]

AESI: adverse events of special interest; EMA: European Medicines Agency; IV: intravenous; NA: not applicable; PD: pharmacodynamics; PK: pharmacokinetics; RCT: randomised controlled trial; SC: subcutaneous; SLE: systemic lupus erythematosus

B.2.12 *Innovation*

Belimumab is a biologic therapy, targeting the BLyS pathway associated with an immune response in SLE, that addresses a substantial unmet need in a chronic and potentially debilitating disease. When added to ST, belimumab reduces not only short-term disease activity but, as demonstrated by the recently published PSM analysis[48], also long-term organ damage. Belimumab also reduces steroid usage, which is a crucial benefit considering the long-term adverse consequences of steroid use are well known, and that chronic steroid use contributes substantially to the organ damage that patients with SLE accumulate over the years[26].

Belimumab also impacts disease signs and symptoms that are important for patients, reducing the incidence of disease flares (including severe flares) and fatigue. Fatigue in SLE patients can be significantly debilitating and have a severe adverse impact on QoL. Clinical advice suggests that among all of the SLE symptoms, patients consider fatigue to be the most important. The reduction in fatigue observed with belimumab, as well as prevention of SLE flares, may significantly improve the QoL of SLE patients.

While patients with other autoimmune conditions, such as rheumatoid arthritis, have seen a substantial growth in the number of treatment options and could experience the associated clinical and QoL benefits, treatment of SLE still relies on older, non-specific therapies. The heterogeneity of SLE has contributed to the repeated failures of clinical trials in this disease, restricting patients to older, non-specific therapies as the mainstay of treatment. Multiple targeted treatments e.g. (rituximab, ustekinumab, anifrolumab, abatacept, atacicept, lupuzor) working through various mechanisms (anti-CD20, anti-IFN, T-cell modulators) have been studied for the treatment of lupus in the recent years (reviewed by Touma and Gladman[84] and Vukelic et al.[85]), yet failed to show benefit or meet their protocol specific primary endpoints. The positive results observed with belimumab are an exception among the ever-growing list of failed treatments for SLE.

Finally, a SC formulation has been developed in addition to IV belimumab, to offer patients a choice of treatment modalities. While some patients may prefer to receive

their treatment once every 4 weeks in clinic rather than having to self-administer it weekly, others may appreciate not having to travel to clinic appointments, especially if regular appointments are a burdensome interruption to their everyday lives, or if they have to rely on external help to attend the appointment. Therefore, the additional SC formulation broadens access to treatment promoting equality. In addition to offering more flexibility for patients, SC belimumab also reduces the burden on NHS resources compared with the IV formulation, as clinic time is not required for administration.

The availability of the SC formulation to NHS patients is particularly topical during the COVID-19 pandemic. Patients with SLE receiving belimumab IV were deemed to be at a high-risk for infection and required to self-isolate and potentially shield[86]. To enable patients to continue treatment with belimumab through self-administration at home, GSK has made the SC formulation (belimumab 200 mg solution for injection in pre-filled pen) temporarily available to allow treatment continuation through self-administration at home. This clearly demonstrates the importance of having an additional formulation available that does not require administration in the hospital setting.

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

B.2.13.1 Key aspects of the evidence

- **Several areas of uncertainty raised in TA397 have been addressed:** The additional efficacy and safety data collected since TA397 addresses most of the key areas of uncertainty identified by the committee and the ERG during that submission. This includes, rate of development of organ damage (see SLICC/SDI Indirect Cohort Comparison Study[48] described above), beneficial impact of belimumab on QoL (particularly with respect to improvement in fatigue) and UK standard therapy, with the latter data derived from the UK BILAG-BR. LTE studies and the BILAG-BR provide information on discontinuations, including patient numbers and reasons for discontinuation. There remain evidence gaps regarding the length of treatment required with belimumab. With regards to treatment duration in responders, we propose that patients are treated with

belimumab for as long as they continue to derive a benefit and do not experience adverse events, so that time on treatment is likely to vary between individuals. Where a patient's disease is well controlled on belimumab, some physicians may consider a period of treatment cessation. The BILAG-BR data has only been collected since 2016 and recruitment was slower than anticipated. Consequently, due to the relatively short study duration and low patient numbers, this real-world UK source provides only limited information on discontinuations and length of treatment.

- **Belimumab is efficacious in HDA-1 and HDA-2 populations:** Clinical trial data demonstrated efficacy of belimumab both in the currently NICE-approved HDA-1 population, and in the more clinically relevant HDA-2 population, which included patients with a SELENA-SLEDAI score ≥10 AND at least one of the following serological features: low complement OR positive anti-dsDNA. Analyses of data from the Phase 3 trials presented in Section B.2.7 unequivocally support the use of belimumab in this population compared with the full ITT population of BLISS trials. Compared to the previously approved and restricted target population (HDA-1), the HDA-2 population is more clinically applicable and accurately reflects those patients who experience high disease activity and who are likely to still derive a benefit from treatment with belimumab. This notion is supported by the much slower than anticipated recruitment into the BILAG-BR, which suggested that the HDA-1 population posed excessive eligibility restrictions, excluding patients who have HDA, but do not necessarily present with both low complement and anti-dsDNA antibodies concomitantly.

  Amending the criteria will support SLE patients with high disease activity in England and Wales, providing them with access to an important licensed treatment option to better manage their disease and potentially minimise detrimental organ damage in the longer term. Moreover, this will continue to provide a more cost-effective use of NHS resources, as outlined in the economic evaluation.

B.2.13.2 Key conclusions from the evidence

- **Clinical trial data is relevant to the UK patient population:** This submission presents a wealth of comprehensive evidence on the efficacy and safety of belimumab in both clinical trial and real-world settings, including long-term data pertaining to disease control and minimisation of organ damage. The BILAG-BR data collected for UK SLE patients who received belimumab provides reassurance that the benefits seen in the clinical trials can also be realised in the real-world setting.

- **Long-term disease control has been demonstrated:** A substantial body of additional evidence has been collected since TA397. Belimumab has been shown to provide effective long-term disease control, based on up to 13 years of data from the Phase 2 LBSL02 study LTE and up to 8 years of data from LTEs of Phase 3 trials. Furthermore, the SLICC/SDI Indirect Cohort Comparison Study[48] comparing patients enrolled in the BLISS-76 US LTE against those from the real-world TLC has demonstrated that add-on treatment with belimumab significantly reduces the risk of irreversible, long-term organ damage that accumulates over the years from poor disease control and cumulative intake of corticosteroids, leading to the development of serious and potentially life-threatening comorbidities[26].

  LTE studies consistently showed that add-on treatment with belimumab has the potential to reduce the dose of steroids which is critical in the management of this complex, chronic disease given the known detrimental and irreversible long-term adverse consequences of steroid use[26]. Although reductions in steroid dose did not consistently meet statistical significance in the belimumab trials, the RCTs were a blinded design, and no steroid tapering regime was implemented in any of the trial protocols. Given these design features, a hesitancy to reduce steroid dose in the trials may have led to an underestimate of the potential steroid-sparing benefit of belimumab.

- **SC belimumab is comparable to IV belimumab:** This submission introduces SC belimumab, delivered through an autoinjector device. SC belimumab provides

comparable efficacy results to the IV formulation as demonstrated in the ITC[70], is easy to self-administer, and offers a choice for patients for whom travelling to the clinic to receive monthly IV infusions is difficult or poses a burdensome interruption to their everyday lives. The availability of a formulation that patients can self-administer at home has also proven instrumental during the COVID-19 pandemic, where, as discussed previously, GSK temporarily made SC belimumab available to enable patients, who were required to self-isolate or shield, to continue their treatment. Furthermore, compared with IV dosing, the use of SC belimumab will reduce the burden on NHS resources as no clinic time is involved in drug administration. Both the IV and SC belimumab will be available to ensure physicians and patients can choose the formulation that is best suited to their circumstances, which should translate to better adherence and improved outcomes.

- **The HDA-2 subgroup is the population under consideration in this submission**: Recruitment into the BILAG-BR showed that the numbers of patients who met the HDA-1 definition to receive belimumab in England was substantially smaller than anticipated and suggested that the HDA-1 population was too restrictive. To better address the unmet need in SLE and more accurately reflect patients with high disease activity, we focus on the HDA-2 population.

B.3 Cost effectiveness

A cost-effectiveness analysis was undertaken to evaluate two formulations of belimumab (Benlysta) in adults, taking an NHS and PSS perspective in England and Wales in the treatment of active, autoantibody-positive systemic lupus erythematosus (SLE) with a high degree of disease activity despite standard therapy.

Populations

The cost-effectiveness analysis explores two **High Disease Activity (HDA)** populations:

- HDA-1: SELENA-SLEDAI score ≥10 AND low complement AND positive anti-dsDNA – **current NICE guidance population.** HDA-1 results are presented in Appendix Q.
- HDA-2: SELENA-SLEDAI score ≥10 AND at least one of the following serological features: low complement AND/OR positive anti-dsDNA – **the base-case for this appraisal**

The Intervention: Belimumab

Belimumab IV

For the IV formulation, belimumab 10 mg/kg IV plus standard therapy (ST) is compared to ST alone. Both the HDA-1 and HDA-2 populations in this analysis are a subgroup of the total pooled Intent-To-Treat (ITT) patient population recruited into the two Phase 3 IV clinical trials: BLISS-52 and BLISS-76, excluding the unlicensed belimumab 1mg/kg treatment arm.

Belimumab SC

For the SC formulation, belimumab 200mg SC plus ST is compared to ST alone.
Both the HDA-1 and HDA-2 patients in this analysis are a subgroup of the ITT population recruited into the Phase 3 SC clinical trial; BLISS-SC.

Economic Analysis

A micro-simulation cost utility model simulating individual patients over a lifelong period is presented; a model for the IV formulation and a model for the SC formulation.

Trial-based model inputs

Clinical efficacy

Data for the treatment effect inclusive of the year 1 discontinuation rate for the IV formulation HDA-2 base-case are calculated from BLISS-52, and BLISS-76 trials, whilst for the SC formulation HDA-2 base-case is calculated from the BLISS-SC trial. Year 2 discontinuation rates for both formulations are from an integrated analysis of Phase 2 and Phase 3 IV LTE studies.

Health related Quality of life (HRQoL)

A linear regression model was used to calculate HRQoL using EQ-5D measurements from BLISS-52 and BLISS-76 data in both the IV and SC models.

Resource utilisation

Utilisation data are derived from the Phase II LBSL02 belimumab study and unit costs are derived from published UK sources.

Results for the base-case HDA-2 subgroup

Belimumab IV

- Incremental costs [redacted]
- Additional life years [redacted]
- The resultant incremental cost effectiveness ratio (ICER) is £30,001 per QALY.

Belimumab SC

- Incremental costs [redacted]
- Additional life years [redacted]
- The resultant incremental cost effectiveness ratio (ICER) is £30,566 per QALY.

Scenario Analysis
A number of scenario analyses were conducted to examine the effects of alternative plausible scenarios. The results of scenario analyses ranged from £19,818 per QALY gained to £28,095 per QALY gained across both the IV and SC models for the HDA-2 patient subgroup.

Sensitivity Analysis
Both probabilistic sensitivity analyses (PSA) and deterministic sensitivity analyses were conducted. The results of the PSA showed that results were robust with mean ICERs for the SC and IV formulations compared to ST at £29,264 and £31,629 respectively. Results from the one-way deterministic sensitive analyses showed that the parameter that was most sensitive to the ICER was the year two natural discontinuation rate.

Conclusion
Both belimumab IV and belimumab SC have ICERs of approximately £30,000 per QALY gained in the base-case HDA-2 population, and under £30,000 per QALY gained across all scenario analyses. Some conservative assumptions have been applied to the modelling, for example, limiting the duration of benefit of belimumab on slowing the progression of organ damage, and assuming long durations of treatment, so the true ICERs are likely to be lower than the base case ICERs presented. Therefore, belimumab continues to represent an efficient use of NHS resources.

B.3.1 *Published cost-effectiveness studies*

A systematic literature (SLR) review was conducted to identify economic evaluations of belimumab against any other comparator. A full description of the SLR is provided in Appendix G (including search strategy, included, and excluded records with reasons and data extraction tables) and the PICOS criteria are summarised in Table 54.

A previous search for relevant cost-effectiveness studies to support the NICE submission for belimumab (included as part of TA397, 2016) did not identify any relevant economic studies. Eight bibliographic databases were searched between 28th January 2020 and 19th February 2020. We restricted our search to English-language studies and placed no restrictions on the time period in which studies may have been published. Conference abstracts were restricted to 2017 onwards.

Additional records were identified via grey literature searches and references identified in SLR studies. Two reviewers independently assessed the eligibility of records based on title and abstract and full text. One reviewer extracted data from each eligible study, with a second reviewer checking the extracted data.

Table 54. Published cost-effectiveness studies

	Inclusion Criteria	**Exclusion Criteria**
Population	Patients with SLE	Studies which included more than 25% of patients with significant renal involvement (lupus nephritis) or CNS involvement (central nervous system lupus). Belimumab is not currently licensed for the management of lupus nephritis or CNS lupus.
Intervention	Belimumab	No reference to belimumab
Comparators	Standard Therapy alone; belimumab; cyclophosphamide; rituximab	
Outcomes	Total costs; Summary health outcomes (Quality-adjusted Life Years (QALYs)); Incremental cost-effectiveness ratio (ICER).	
Study Design	• Cost-utility analysis • Cost-effectiveness analysis	• Case reports • Case studies • News • Comments • Editorials • Letters • Budget impact, cost comparisons
Limits	Reported in English language Conference abstracts published from 2017 onwards.	• Non-English language studies. • Full text unavailable • Duplicate studies • ERG report on the original NICE submission • Published in error & withdrawn. • Societal perspective analysis • Conference papers published before 2017.

HTA, health technology assessment; ICER, incremental cost effectiveness ratio; QALY, quality-adjusted life year; SLR, systematic literature review.

B.3.1.1 Results

Figure 6 shows the flow diagram of records retrieved. The search identified 224 records, with an additional record identified through other sources. Following deduplication, 175 records were assessed for relevance. Screening by abstract removed a further 125 records. Full text screening excluded a further 47 records. Excluded articles and rationale can be found in Appendix G. Three records were included in the final review, and all derived from a GSK sponsored SLE cost-utility analysis.

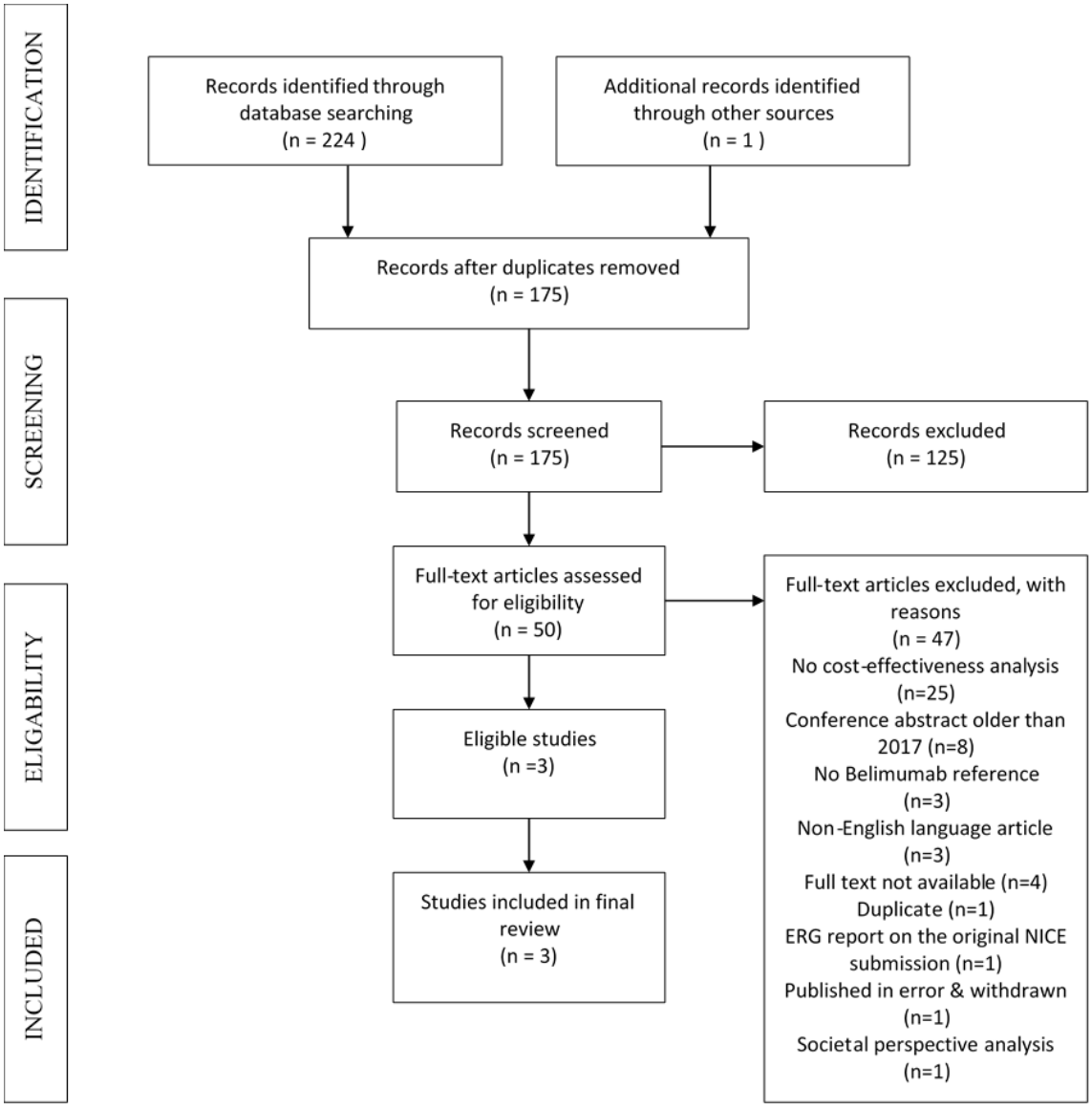


Figure 6. Flow diagram of identification of records retrieved

B.3.1.2 Summary of published belimumab cost-effectiveness studies

A summary of the three included cost-effectiveness publications is provided in the following Table 55.

Table 55. Summary of the cost-effectiveness studies

Study	Year	Summary of model	Patient population, (average age in years)	QALYs (intervention, comparator)	Costs (currency) (Intervention. Comparator)	ICER (per QALY gained)
Specchia[87]	2014	Cost-utility analysis conducted to estimate the cost-effectiveness of belimumab in patients with a high level of SLE disease activity compared to usual ST in Italy over a lifetime horizon. A microsimulation model was developed in Microsoft Excel. This paper summarises a health technology assessment performed in Italy.	34.7 (mean)	Belimumab:11.31 ST: 10.78	Belimumab: €142,921 ST: €125,234	The ICER for belimumab compared to ST was €32,859/QALY gained in the base-case. Sensitivity analysis and scenario analysis results were not provided.

Pierotti[88]	2015	Cost-utility analysis conducted to estimate the cost-effectiveness of belimumab in patients with a high level of SLE disease activity compared to usual ST in Italy over a lifetime horizon. A microsimulation model was developed in Microsoft Excel. The model used in this analysis is adapted from the model submitted to NICE as part of TA397. This paper details the cost-utility analysis reported in Specchia et al (2014).	34.7 (mean)	Belimumab:11.31 ST: 10.78	Belimumab: €142,921 ST: €125,234	The ICER for belimumab compared to ST was €32,859/QALY gained in the base-case. The ICERs for the one-way sensitivity analysis ranged from €25,408/QALY gained to €49,825/QALY gained. The ICERs for the scenario ranged from €28,754/QALY gained to €39,515/QALY gained.
CADTH[89]	2012	Cost-utility analysis comparing belimumab plus standard therapy against ST alone, in a subgroup of patients over a lifetime time horizon.	Not stated	Not stated	Not stated	Belimumab plus ST compared with ST alone was associated with an ICER of Can$112,883 per QALY gained
Abbreviations: QALYs, quality-adjusted life years; ICER, incremental cost-effectiveness ratio; ST, Standard Therapy						

Specchia et al. (2014)[87] reported on a cost utility analysis performed as part of an HTA on belimumab in patients with SLE in an Italian setting (in line with population recruited to the BLISS studies, and in a separate subgroup with both low complement and anti-dsDNA.) This paper briefly reported the results of the economic analysis, providing base-case and summary sensitivity analysis only. Specchia et al concluded that in the Italian setting and according to the guidelines of the Italian Association of Health Economics (IAHE), belimumab is shown to be cost-effective in terms of both ICER and ICUR, (at the threshold of €25,000 – €40,000 per QALY gained).

Pierotti et al. (2015)[88], a GSK sponsored study, provided detailed reporting of the economic analysis undertaken and presented in Specchia et al. (2014)[87]. The study assessed the cost-effectiveness of belimumab with standard therapy (typically glucocorticoids and various immunosuppressants, mostly unlicensed for SLE), in a lifetime micro-simulation, based on the UK model that was submitted to NICE as part of TA397, and subsequently adapted to the Italian setting. In addition to the information reported by Specchia et al. (2014)[87], Pierotti shared several scenario analysis which included a maximum 10-year treatment duration, where the ICER varied from €28,754/QALY gained to €39,515/QALY gained (see Table 28 in Appendix G.)

The third included study, CADTH (2012)[89], reported on the HTA decision of Canadian Agency for Drugs and Technologies in Health (CADTH), for the use of belimumab in patients with SLE. This decision considered the comparison of belimumab with standard therapy (which was defined as any of the following: prednisone or equivalent, antimalarials, NSAIDs, or any immunosuppressive therapy). Very limited details of the economic analysis performed were reported as part of this decision.

No conference abstracts (based on those conferences searched, including ISPOR) published from 2017 onwards were identified.

B.3.2 *Economic analysis*

The economic evaluations of relevance that were identified in the literature reflect the de novo model developed by GSK for TA397. No new evaluations were identified.

To best reflect the costs and benefits of belimumab added to standard therapy compared to standard therapy alone in adults with active, autoantibody-positive systemic lupus erythematosus with a high degree of disease activity despite standard therapy, an updated version of the micro-simulation economic model provided as part of TA397 is presented. An NHS and personal social service perspective for the analysis is adopted and discounting is applied at 3.5% for both costs and benefits.

As this submission is a re-appraisal of TA397, following advice from the NICE team and the ERG at the Decision Problem Meeting (August 2020), the general approach taken in Sections B.3.2- B3.11 is to provide information on updates to the evaluation since TA397 (structural, input parameters, subgroups etc) and therefore the company submission provided to NICE in 2011 as part of TA397 will be referred to as 'the previous submission' herewith. Please refer to the appropriate sections of the previous submission for rationale of the model development, structure, functioning and modelled outcomes of SLE for disease activity and organ damage.

Note, throughout the remainder of B.3, "belimumab treatment" refers to treatment with belimumab plus Standard Therapy (ST) regardless of formulation, "placebo" refers to treatment with placebo plus ST, whereas "ST" refers to ST alone.

B.3.2.1 Patient population

In TA397, GSK presented analyses for the total pooled SLE patient population recruited into the two Phase 3 clinical trials evaluating IV belimumab: BLISS-52 and BLISS-76, (excluding the unlicensed belimumab 1mg/kg treatment arm). GSK also presented a High Disease Activity (HDA) subpopulation of the total pooled SLE patient population. TA397 recommends the use of add-on belimumab in a HDA subpopulation of the total pooled SLE patient population from BLISS-52 and BLISS 76, defined as:

- (SELENA-SLEDAI) score ≥10 AND low complement AND positive anti-dsDNA
    - Referred to as HDA-1 in this submission.

HDA-1 represents a total of 23.5% of patients of the two pooled Phase 3 IV clinical trials (for the licensed dose 10mg/kg) and 31.6% of patients in the Phase 3 SC clinical trial. Note: the higher proportion seen in the SC trial is likely due to a higher baseline disease severity (SS score) inclusion criterion of 8 (vs 6 for the IV studies).

The cost-effectiveness of add-on belimumab compared with ST alone in HDA-1 is presented for completeness (see Appendix Q).

As discussed in Section B.2.7, since TA397 and throughout the Managed Access Agreement (MAA) data collection period, clinical experts highlighted the challenges in identifying patients satisfying all three selection criteria. As a result, a proportion of SLE patients with HDA despite ST were unable to access treatment with belimumab. Following advice from lupologists, patients with an SS score of ≥10 and either low complement or anti-dsDNA would still be considered as having high disease activity and therefore the requirement for both biomarkers is too stringent.

Therefore, whilst mindful of NHS resources, to ensure appropriate SLE patients with HDA have access to this licensed treatment we propose an alternative more clinically relevant subgroup; this is the base-case for this submission, defined as:

- (SELENA-SLEDAI) score ≥10 AND at least one of the following serological features: low complement AND/OR positive anti-dsDNA
    - Referred to as HDA-2 in this submission.

HDA-2 represents a total of 31.6% of patients of the two Phase 3 IV clinical trials and 52.3% of patients in the Phase 3 SC clinical trial. Note: the higher proportion seen in the SC trial is likely due to a higher baseline disease severity (SS score) inclusion criterion of 8 (vs 6 for the IV studies).

Relative to the previously recommended HDA-1 population, the HDA-2 population better defines, clinically, a relevant SLE population with significant disease activity

that could benefit from belimumab to manage their SLE and avoid detrimental organ damage in the longer term. It also still constitutes a cost-effective use of NHS resources. Therefore, the base-case for the current economic evaluation is the HDA-2 population.

B.3.2.2 Model

The microsimulation model structure fundamentally remains unchanged since TA397 and is presented in Figure 7.

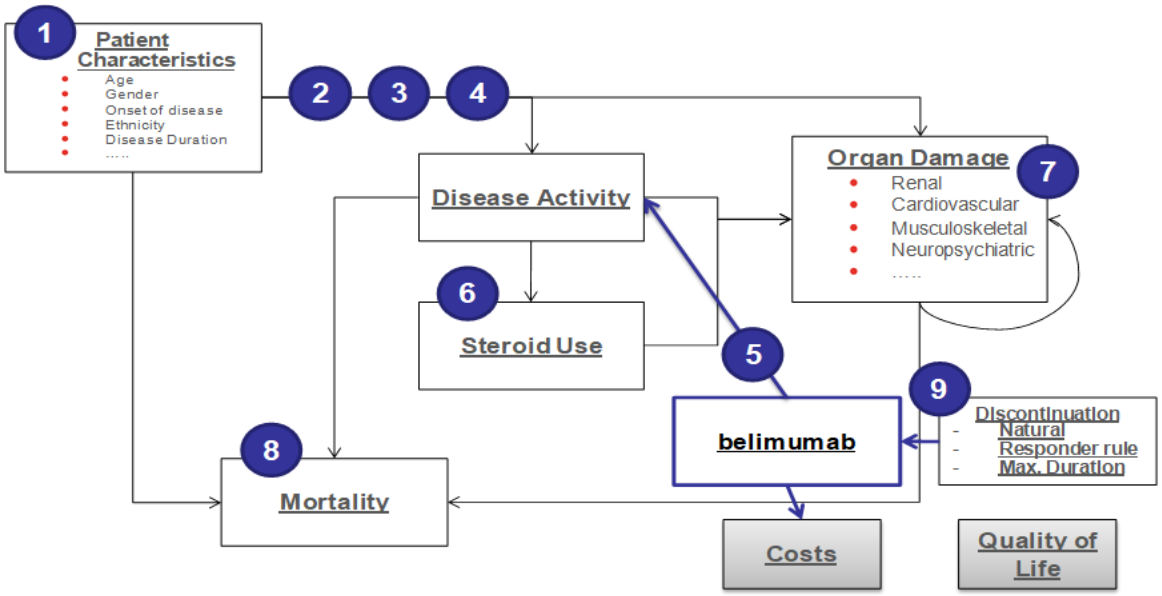


Figure 7. Schematic overview of interdependencies between baseline characteristics, treatment and outcomes in the micro-simulation model (presented in TA397)

1. *Simulation of a patient: Baseline characteristics are sampled from the baseline characteristics of the relevant population in the BLISS trials (by formulation).*
2. *Response at 24 weeks (defined as a decrease in SS score of ≥4 points after 24 weeks): Determined from the probabilities of response in the BLISS trials, stratified by baseline SS score.*
3. *Disease activity in the first year: A regression model produced from BLISS trial data to explain the change in SS score after 52 weeks, based on treatment, baseline SS score and SS score response at 24 weeks (yes or no).*
4. *Disease activity over time: SS score over time (after the first year) for a standard therapy (ST) patient is determined with a statistical model developed using the Johns Hopkins cohort longitudinal database.*
5. *Effect of belimumab on SS score: The regression model for SS score at 52 weeks (3) is used to determine the difference between a ST and a belimumab patient. This is subtracted from the disease activity over time (4). A patient discontinuing belimumab treatment returns to ST disease activity levels.*
6. *Steroid use: determined by a model developed on the Johns Hopkins cohort. The model explains steroid use at a time point based on the average disease activity in the last year.*
7. *Organ damage: The Johns Hopkins database was also used to estimate the time to organ damage outcomes. Yearly organ damage probabilities are calculated based on patient characteristics, disease activity (adjusted [average] mean SLEDAI [AMS]) and steroid use. A propensity score matched comparative analysis has since provided an estimate of the long-term reduction in SDI for patients on add-on belimumab compared with a matched cohort (Toronto Lupus Cohort) on ST. In the current appraisal, we incorporate the findings from the propensity score matched analysis to model the long term organ damage reduction*

treatment effect shown by belimumab (see Section B.3.3.6), by means of a calibration factor.

8. *Mortality: yearly mortality risk is calculated by combining average population life tables with an increased mortality in SLE patients and a statistical model explaining the influence of patient characteristics, disease activity and organ damage on mortality.*

Refer to Section 6.2.2 of the previous submission for details on:

- Justification of the chosen structure in line with the clinical pathway of care.
- How the model structure and its health states capture the disease or condition for patients.

The model continues to use a cycle length of one year with a lifetime horizon, as this best captures the changes in overall disease activity and the accumulation of organ damage. A half cycle correction was not included.

We present a separate model (replicate models of the structure shown in Figure 7) for each formulation, belimumab intravenous (IV) and subcutaneous (SC).

B.3.2.2.1 Key measures of SLE for disease activity and organ damage

A key measure of disease activity in SLE is SELENA-SLEDAI (SS), and of organ damage, is the Systemic Lupus International Collaborating Clinics/American College of Rheumatology (SLICC/ACR) Damage index (SDI). Please see Section 6.3.1 of the previous submission for a discussion of these key measures.

B.3.2.2.2 Response definition

The primary endpoint across the IV and SC pivotal Phase III studies (BLISS-52, BLISS-76 and BLISS-SC) was response in SLE Responder Index (SRI-4) at week 52, defined as a composite of:

i) a ≥ 4-point reduction from baseline in SS score and
ii) no new BILAG A organ domain score or 2 new BILAG B organ domain scores compared with baseline, and
iii) no worsening (increase of < 0.30 points from baseline) in Physician's Global Assessment (PGA)

The SS score component of the composite SRI-4 endpoint at week 24 alone remains the most appropriate and the only feasible methodological approach to link to long-term outcomes in the belimumab IV and SC models. Please see Section 6.3.1 of the previous submission for a detailed discussion of the rationale of this assumption.

In the previous submission, a series of linear regressions on the pooled BLISS trial data explained the difference between the SS score at baseline and week 52. This was dependent on baseline SS score combined with a treatment indicator variable, and a "response" indicator variable identifying whether or not patients were classified as satisfying the treatment continuation rule at week 24 with belimumab. Table 6.5 in Section 6.3.1 of the previous submission shows the linear regression which explains the change in SS score at week 52 for the pooled total population, whilst Figure 6.5 from the same Section shows the plots of correlation between baseline SS and difference after 52 weeks for ST patients, and belimumab responders and non-responders in the pooled total population. This approach continues to be used in both the current IV and SC models, with updated linear regressions explaining the change in SS score after 52 weeks compared to ST for the HDA-1 and HDA-2 populations presented in Appendix Q and Table 61 respectively.

B.3.2.3 Intervention technology and comparators

B.3.2.3.1 The intervention: Belimumab IV and SC

In TA397, GSK presented belimumab IV formulation. The current submission introduces a SC formulation of belimumab, administered via a pre-filled pen (autoinjector device) (please see Section B.2.6.5.1 for details). As discussed in Section B.2.13, belimumab SC provides comparable efficacy results with the IV formulation, can be self-administered outside of hospital setting, and offers a choice for patients for whom travelling to the hospital to receive a monthly IV infusion is difficult or poses a burdensome interruption to their everyday lives.

The availability of the SC formulation that patients can self-administer at home has also proven instrumental during the COVID-19 pandemic, where GSK temporarily made SC belimumab available to enable patients, who were required to self-isolate or shield, to continue their treatment to reduce the risk of flares. Furthermore,

compared to the IV formulation, the use of SC belimumab will reduce the burden on NHS resources as no clinic time is involved in drug administration following initial education. GSK would like to provide both the IV and SC formulations to ensure patients and their clinicians can choose the formulation that is best suited to their circumstances, which should translate into improved equality of access to treatment.

Table 56 provides an overview of the intervention and comparators considered in this economic evaluation.

Table 56. Overview of intervention and comparators in the current submission

	IV model		SC model	
Subgroup	**HDA-1**	**HDA-2 [Base-case]**	**HDA-1**	**HDA-2 [Base-case]**
Formulation	120 mg or 400 mg powder for concentrate for solution for infusion • GSK Summary of Product Characteristics: Benlysta 120 mg powder for concentrate for solution for infusion.*[6] • GSK Summary of Product Characteristics: Benlysta 400 mg powder for concentrate for solution for infusion.*[6]		1-ml pre-filled pen contains 200 mg of belimumab: available in packs of four pens • GSK Summary of Product Characteristics: Benlysta 200 mg solution for injection in pre-filled pen.*[6]	
Dosing	Belimumab 10 mg/kg is administered as an IV infusion over a one-hour period on days 0, 14 and 28, and at 4-week intervals thereafter in addition to standard therapy in a clinic centre by trained nurses. In Year 1 there are 14 administrations and in Year 2 onwards there are 13 administrations per year		Belimumab 200 mg solution for injection in pre-filled pen administered via SC route each week. 53 doses in the first year, and 52 doses each year thereafter.	
Comparator	ST alone			
*Last updated 2020				

B.3.2.3.1.1. Belimumab IV

Belimumab IV is dosed by patient weight, with patients receiving 10mg/kg of body weight by infusion. In the previous submission, the source for average patient weight (65.4 kg) was derived from the pooled BLISS studies “high disease activity” subgroup (HDA-1). In the current submission, the average patient weight of 70.4 kg is derived from 151 patients prescribed belimumab who had their weight captured in the BILAG-BR registry. (Weights of these patients ranged from 39.0kg to 97.9kg.)

As belimumab IV is administered within a hospital setting it seems reasonable that compliance will be reasonably high. In the base-case, the average exposure to the trial product is assumed to be 100%. Level of compliance (i.e. exposure) can be changed in the model settings. However, this will affect only drug costs in the model; no adjustment of efficacy is made. This is because there is a lack of data to model the effect a reduced exposure would have on disease activity and longer-term outcomes.

It is assumed that vial sharing between patients will not occur. As the number of patients with moderate to severe SLE is relatively small, vial sharing may not be easy to manage in tertiary care units due to storage requirements.

B.3.2.3.1.2. Belimumab SC

Patients who receive belimumab SC are trained to self-administer a single belimumab 200 mg solution for injection in pre-filled pen subcutaneously, once each week. In contrast to belimumab IV, dosage for belimumab SC is not based on patient weight.

Fifty-three administrations per year are required. It is assumed that patients require up to an hour with a specialist nurse within the first year of receiving belimumab SC in order to receive training and education on how to self-administer effectively, and to assess any associated adverse reactions. Following initial education by a specialist nurse on how to self-administer the SC formulation, it is assumed that patients can competently self-administer throughout the duration of their treatment. Although belimumab SC is a self-administered formulation, the model assumes that average exposure to the product is 100%, as patients in the BLISS-SC trial had an exposure of 97% to the trial product.

B.3.2.3.2 Comparators

The Final Scope for the current appraisal considers the following comparators:

- Standard therapy alone

For people in whom it is considered appropriate:

- Rituximab plus standard therapy
- Cyclophosphamide plus standard therapy

B.3.2.3.2.1. Standard therapy

As per TA397, standard therapy continues to include the use of antimalarials (i.e. hydroxychloroquine), NSAIDs, corticosteroids and immunosuppressants such as azathioprine, methotrexate and mycophenolate mofetil. Many of the treatments used for SLE are unlicensed, with only hydroxychloroquine, corticosteroids and azathioprine licensed for use in SLE.

B.3.2.3.2.2. Cyclophosphamide

As per TA397, our current economic analysis does not consider cyclophosphamide plus standard therapy. See Section B.2.2 for the justification of the exclusion of this medicine.

B.3.2.3.2.3. Rituximab

The current economic analysis does not consider rituximab plus standard therapy as a comparator in the economic evaluation. Please see Section B.2.9 for justification.

We also acknowledge our commitment to the Managed Access Agreement of TA397, to collect data via the BILAG Biologics Registry (BR) on the use of both belimumab and rituximab. In the FAD for TA397, the Committee supported the data collection with the potential to provide additional information in the future technology appraisal of belimumab.

The FAD from TA397 concluded 'The Committee heard from the ERG that there were 3 outcomes for which an indirect comparison could be completed (that is, BILAG, SLEDAI and SF-36 scores), but data were only available in the public domain for the SF-36. The ERG also highlighted the differences in the trial populations, which it considered meant that the results of an indirect comparison were not meaningful. The Committee concluded that there were no data that would allow a robust calculation of the relative clinical efficacy of belimumab compared with

rituximab. Following on from TA397, the update to the systematic literature review did not identify any studies that directly compared belimumab with rituximab. Differences in the patient populations and measurement in end points from previous studies still precludes the conduct of any meaningful indirect and mixed treatment comparisons between belimumab and rituximab as outlined in Section 5.7.1 of the previous submission and Section B.2.9 of the current appraisal.

The data available from the BILAG-Registry shows that since 2016 (and before the NHS E policy of July 2020[44]), there is an overlap of patients i.e. some patients who receive rituximab would be eligible to receive belimumab. The patient characteristics are shared in Section B.2.3.4. As part of the analysis of the BILAG-BR, the University of Manchester did undertake a multilevel regression modelling exercise to explore the patient outcomes across the three cohorts: belimumab, rituximab and non-biologic. In these regression models, treatment effect estimates are compared with rituximab (reference) due to the availability of the largest sample size and results are reported as effect co-efficients. The results suggest that for most health outcome measures patients in the belimumab cohort demonstrate a similar level of improvement to rituximab. However, the regression modelling could only be conducted out to 12 months because of the limited follow-up data available for belimumab patients in this study. It remains that reducing the risk of long-term organ damage is a key treatment goal for SLE patients and of most interest to clinicians. Whilst there is published data to support this for belimumab, there is limited equivalent evidence on impact for rituximab.

As GSK does not feel that a robust comparison can be made with rituximab, we have concentrated on the comparator of ST only in this economic analysis.

B.3.2.4 Features of the economic analysis

A summary of the features of the economic analysis is presented in Table 57.

Table 57. Features of the economic analysis

	Previous Appraisal	Current Appraisal	
Factor	**TA397**	**Chosen values**	**Justification**
Time horizon	Lifetime	Lifetime	The economic evaluation estimates costs and health benefits over the full lifetime of each individual. This time horizon is necessary for the key health outcomes and resource use to be fully explored in this chronic disease and is consistent with the NICE reference case

	Previous Appraisal	Current Appraisal	
Factor	**TA397**	**Chosen values**	**Justification**
Cycle length	Yearly	Yearly	SLE is a chronic, incurable disease. The changes in overall disease activity and the accumulation of organ damage are believed to be adequately captured with a yearly cycle over a lifetime horizon. However, if long-term data on the incidence and severity of flares had been available, a shorter cycle length may have been more appropriate to capture the pattern of flares over time.
Half-cycle correction	Not included	Not included	Not applicable

	Previous Appraisal	Current Appraisal	
Factor	**TA397**	**Chosen values**	**Justification**
Measurement of health effects	QALYs	QALYs	This is consistent with the reference case.
Discount rate	3.5% for both benefits and costs	Base-case: 3.5% for both benefits and costs Scenario analysis: (1) 1.5% for both benefits and costs (2) 1.5% benefits /3.5% costs	The values of 3.5% for both benefits and costs are consistent with the reference case.
Perspective	The analysis took an NHS and PSS perspective	The analyses take an NHS and PSS perspective	This is consistent with the reference case.
Patient characteristics	Subset of pooled Phase 3 BLISS-52 and BLISS-76 population	<u>IV model</u> HDA-1 and HDA-2 populations based on pooled Phase 3 BLISS-52 and BLISS-76 population – Note that average patient weight is taken from the BILAG biologics registry <u>SC model</u> HDA-1 and HDA-2 populations based on Phase 3 BLISS-SC trial	<u>IV model</u>

			In the previous appraisal, only the IV formulation was available. The previous model used the HDA-1 population subset for baseline patient characteristics from the Phase 3 IV trials for the analysis. In the current submission for the IV formulation, we continue to use the same data for baseline patient characteristics. However, as UK relevant patients' weights are now available from the BILAG biologics registry, we update the model with this data accordingly for HDA-1 and HDA-2

	Previous Appraisal	Current Appraisal	
Factor	**TA397**	**Chosen values**	**Justification**
			populations. This is relevant to calculate the dosage for each patient (and associated number of vials). <u>SC model</u> The SC model presented in the current submission uses the relevant data to the HDA population subset under consideration for baseline patient characteristics from the Phase 3 BLISS-SC trial.

Treatment continuation rule	Patients on belimumab who do not satisfy the treatment continuation criterion (demonstrating a SS score decrease of 4 or greater) at week 24 remain in the belimumab arm of the model but continue to receive ST treatments after this time-point and assume the average ST level of disease activity for the	Patients on belimumab who do not satisfy the treatment continuation criterion (demonstrating a SS score decrease of 4 or greater) at week 24 remain in the belimumab arm of the model but continue to receive ST treatments after this time-point and assume the average ST level of disease activity for the remainder of the model horizon.	Withdrawing patients from belimumab due to inadequate response to the drug is consistent with the SmPC for belimumab. If patients do not demonstrate a sufficient level of response after six months of treatment with belimumab they would not continue on this drug.

	Previous Appraisal	Current Appraisal	
Factor	**TA397**	**Chosen values**	**Justification**
	remainder of the model horizon.		

Discontinuation rate for patients on belimumab	**Year 1 values** IV model HDA-1: 8% **Annual Year 2+ values** IV model HDA-1:11.7%	**Year 1 value** [redacted]	In the previous analysis, year 1 discontinuation rates for the IV model were calculated from the pooled Phase 3 BLISS-52 and BLISS-76 studies, whereas year 2 onwards discontinuation rates were calculated from a 7-year analysis of the LTE for the Phase 2 (LBSL02) study. In our current economic analysis, discontinuation rate data for year 1 of the IV model were calculated for the HDA-1 and HDA-2 populations from the Pooled BLISS-52 and BLISS-76 data. A more recent analysis has

	Previous Appraisal	Current Appraisal	
Factor	**TA397**	**Chosen values**	**Justification**
			provided an updated value for the year 1 IV model for the HDA-1 subgroup. The discontinuation rate for the year 1 of the SC model for the HDA-1 and HDA-2 populations was derived from the Phase 3 BLISS SC study. Year 2+ annual discontinuation rate data across all models were calculated from an integrated analysis of Phase 2 (LBSL02) and Phase 3 IV (BLISS-52 and BLISS-76) LTE studies.

Treatment effect of belimumab on management of SLE	Lifetime	Lifetime The impact of belimumab on organ damage is included for the first 6-year period, incorporating the results of the propensity score matching comparative analysis by means of a calibration factor Scenario analysis: • The calibration factor is applied to both belimumab and ST arm of the model for a 6-year period. • The calibration for the belimumab arm of the model is applied for patient lifetime.	In addition to the base-case provided as part of TA397, the model also incorporates the benefit of belimumab on the reduction in long term organ damage accrual. Following a full validation exercise, the benefit is applied as a calibration factor over a 6-year period (observation period) derived from a propensity score matched analysis to examine the effects of long-term organ damage reduction for patients receiving belimumab. A scenario analysis will be

	Previous Appraisal	Current Appraisal	
Factor	**TA397**	**Chosen values**	**Justification**
			explored where a calibration factor from the same analysis is also applied to the ST treatment group in the model for a period of 6 years. A further scenario analysis will see calibration factors for belimumab extrapolated and applied for patient lifetime.

	Previous Appraisal	Current Appraisal	
Factor	**TA397**	**Chosen values**	**Justification**
Treatment duration of belimumab	Lifetime	Base-case: Lifetime Scenario analysis: • 10 years for both treatment duration and effect of belimumab on disease.	The base-case remains consistent with the analysis provided in TA397. The scenario analysis allows the further exploration of the assumption that belimumab only exerts an effect on disease for 10 years, and patients do not receive belimumab beyond this duration.
Treatment waning effect	Not applied	Not applied	A treatment waning effect is not applied to the base-case as there is no evidence to date to support this assumption. This approach is consistent with TA397.

	Previous Appraisal	Current Appraisal	
Factor	**TA397**	**Chosen values**	**Justification**
Disease flares	Not included	Not included	The Johns Hopkins cohort database did not record data on disease activity flares so these data could not be modelled directly. However, the protocol for the JH cohort requires patients to visit the clinic every three months or more during flares and so flares to some extent will be captured in the SLEDAI-2K instrument and therefore in AMS.

	Previous Appraisal	Current Appraisal	
Factor	**TA397**	**Chosen values**	**Justification**
			Disease activity at time of organ damage is reflected in the individual system involvement covariates in the natural history of disease (NHD) models; these data would complement the AMS score by describing current disease activity and type of activity.

	Previous Appraisal	Current Appraisal	
Factor	**TA397**	**Chosen values**	**Justification**
Adverse events	Not included	Not included	There was little difference between treatment groups in the BLISS trials in the incidence of all reported adverse events or all serious events and hence there would not be an important cost and utility differentiation between the arms in the health economic model with regards to adverse events.

	Previous Appraisal	Current Appraisal	
Factor	**TA397**	**Chosen values**	**Justification**
Source of utilities	Utility data was searched in Health Technology Assessments (HTAs) available on the NICE website. If the required information was unavailable from NICE, additional searches were carried out on Pubmed.	A targeted literature search was performed to update values from the TA397 submission.	A comprehensive systematic literature search was not deemed feasible because of the breadth of organ systems that would need to be searched for (TA397).

	Previous Appraisal	Current Appraisal	
Factor	**TA397**	**Chosen values**	**Justification**
Source of costs	• PSSRU 2007 (inflated to 2010 costs using the CPI), • NHS reference costs 2005-06 (Department of Health 2006) and inflated to 2010 costs using the CPI (OECD 2010a).	A targeted literature search using a two-staged approach was performed to update values from the TA397 submission. Values were inflated to 2018/2019 values, using the consumer price index for health as published by PSSRU in 2019.	Further in this submission, searches were updated and restricted to seven key organ systems which were shown to contribute most to organ damage related loss in quality-adjusted life years (QALYs) and additional costs, based on initial modelling work.

			A two-staged approach was used to source updated costs (and utilities) for the current submission. First, costs (and utilities) were searched on the NICE website (http://www.nice.org.uk/). Where available, National Institute for Health Research (NIHR) HTA were identified. The technology assessments were searched for health economic information or useful references which would provide relevant cost (or utility data). Where no relevant data was identified through the NICE search, further data

	Previous Appraisal	Current Appraisal	
Factor	**TA397**	**Chosen values**	**Justification**
			were collected using PubMed by searching relevant keywords and MeSH-terms.
NHS, National Health Service; PSS, Personal Social Services; QALYs, quality-adjusted life years.			

B.3.3 *Clinical parameters and variables*

As noted in Section B.3.2.2, the current economic evaluation considers two different formulations of belimumab - intravenous (IV) and subcutaneous (SC). For each of these formulations, separate models consider two High Disease Activity (HDA) subgroups, HDA-1 and HDA-2 (our base case) and variables for these populations are drawn from relevant pivotal Phase 3 clinical trials.

Values for the belimumab IV subgroups are drawn from the total pooled SLE patient population recruited into the two Phase 3 clinical trials: BLISS-52 and BLISS-76, (excluding the belimumab 1mg/kg treatment arm). The pooling of the trial data for belimumab IV is considered appropriate given that the trials were essentially identical in design and in the analysis of the primary endpoint and its three separate components there were no evidence of a treatment-by-study interaction. Pooling the studies increased the sample size and provided more power for the statistical analyses.

Values given for patients in the belimumab SC subgroups are drawn from the full study population who participated in the Phase 3 BLISS-SC clinical trial.

As the HDA-2 patient subgroup is the focus of our economic evaluation, we present here details for this subgroup. Equivalent details for the HDA-1 patient subgroup can be found in Table 1 in Appendix Q.

B.3.3.1 Baseline Characteristics of the study population

The baseline characteristics of the HDA-2 patient subgroups for both the IV and SC models are shown in Table 58, Table 59 and Table 60.

Demographics for the IV and SC studies were similar. The distribution and corresponding parameters used to simulate each characteristic are included. Figure 8. shows the baseline weight distribution for belimumab IV patients obtained from the BILAG biologics registry, which is taken into consideration for patients who receive belimumab IV. Please see Section 6.3.1 of the previous submission for rationale of the development of the baseline characteristics and selection of distributions; this remains unchanged.

Table 58. Baseline patient demographics for HDA-2 patients

Patient demographics	IV model HDA-2 subgroup			SC model HDA-2 subgroup		
	Mean	Distribution	Value	Mean	Distribution	Value
Age (years)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gender (% females)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Black Ethnicity (% black)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SLE disease duration (years)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SLICC damage index score (SDI)*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

**Note that Instead of simulating a patient's total SDI score, the scores simulated for each individual item presented in Table 60 †Probability for each age*

Table 59. Baseline disease activity parameters and steroid use simulated at baseline for HDA-2 population

	IV model HDA-2 subgroup			SC model HDA-2 subgroup		
	Mean (SD)	Distribution	Parameter	Mean (SD)	Distribution	Parameter
Baseline SLEDAI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Increased DNA binding	91.4%	Bernoulli	0.914	92.4%	Bernoulli	0.924
Low Complement	83.1%	Bernoulli	0.831	66.6%	Bernoulli	0.666
Vasculitis	11.8%	Bernoulli	0.118	10.5%	Bernoulli	0.105
Neuropsychiatric involvement	0.6%	Bernoulli	0.006	0.0%	Bernoulli	0.000
Renal involvement	6.4%	Bernoulli	0.064	4.8%	Bernoulli	0.048

Serositis involvement	1.1%	Bernoulli	0.011	5.7%	Bernoulli	0.057
Haematological Involvement	6.4%	Bernoulli	0.064	1.8%	Bernoulli	0.018
Skin Involvement	57.0%	Bernoulli	0.570	77.6%	Bernoulli	0.776
Daily steroid use (mg/day)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

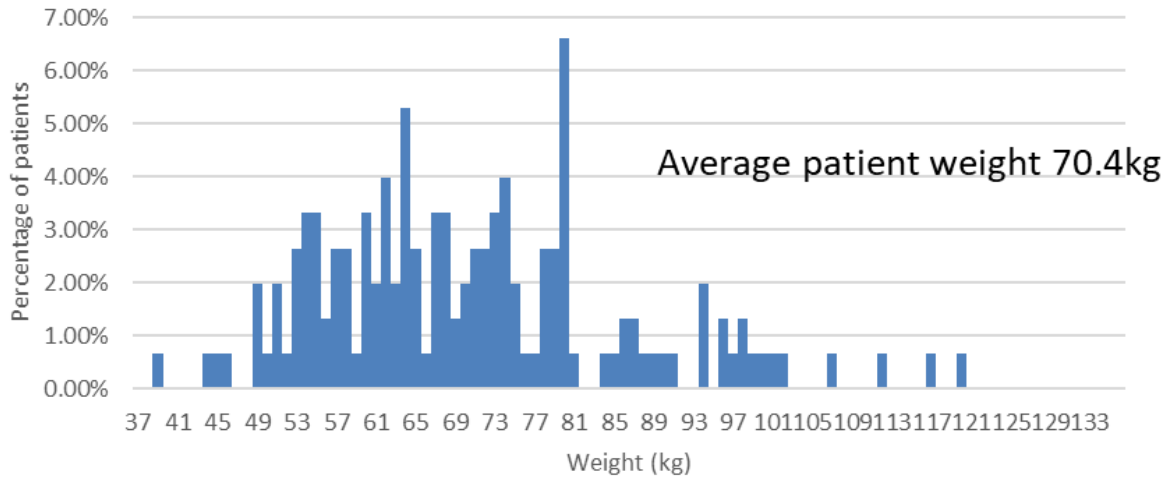


Figure 8. Baseline weight distributions for pooled belimumab IV patients from the BILAG-BR (n=151).

Aligned to the previous submission, an individual organ damage item score was drawn from a multinomial distribution with each category having the probability as outlined in Table 60. This reflects the baseline SLICC/ACR Damage Index (SDI) item score occurrences observed in the HDA-2 subgroups. Therefore, after simulating a patient's baseline characteristics they enter the model in which their remaining lifetime SLE history is simulated.

Table 60. Individual SLICC item scores simulated at baseline for the HDA-2 population

SLICC damage item	IV model HDA-2 subgroup						SC model HDA-2 subgroup					
	Score 0	Score 1	Score 2	Score 3	Score 4	Distribution	Score 0	Score 1	Score 2	Score 3	Score 4	Distribution
Cardiovascular	94.0%	5.1%	0.9%	0.0%	0.0%	Multinomial	96.8%	2.7%	0.5%	0.0%	0.0%	Multinomial
Diabetes	97.6%	2.4%	0.0%	0.0%	0.0%	Multinomial	98.4%	1.6%	0.0%	0.0%	0.0%	Multinomial
Gastrointestinal	96.4%	3.4%	0.2%	0.0%	0.0%	Multinomial	96.8%	3.2%	0.0%	0.0%	0.0%	Multinomial
Malignancy	99.6%	0.4%	0.0%	0.0%	0.0%	Multinomial	99.1%	0.9%	0.0%	0.0%	0.0%	Multinomial
Musculoskeletal	87.4%	8.8%	3.2%	0.4%	0.2%	Multinomial	90.6%	7.6%	1.6%	0.2%	0.0%	Multinomial
Neuropsychiatric	88.9%	9.2%	1.5%	0.4%	0.0%	Multinomial	93.6%	6.2%	0.2%	0.0%	0.0%	Multinomial
Ocular	93.6%	6.2%	0.2%	0.0%	0.0%	Multinomial	90.2%	9.4%	0.5%	0.0%	0.0%	Multinomial
Peripheral vascular	94.4%	5.1%	0.4%	0.2%	0.0%	Multinomial	95.9%	3.4%	0.7%	0.0%	0.0%	Multinomial
Premature gonadal failure	98.9%	1.1%	0.0%	0.0%	0.0%	Multinomial	98.2%	1.8%	0.0%	0.0%	0.0%	Multinomial
Pulmonary	97.0%	2.8%	0.2%	0.0%	0.0%	Multinomial	97.0%	2.7%	0.2%	0.0%	0.0%	Multinomial
Renal	97.4%	2.6%	0.0%	0.0%	0.0%	Multinomial	98.4%	1.6%	0.0%	0.0%	0.0%	Multinomial
Skin	92.1%	7.1%	0.6%	0.2%	0.0%	Multinomial	94.1%	5.3%	0.7%	0.0%	0.0%	Multinomial

B.3.3.2 Year one treatment effects

Unchanged from the previous submission (TA397), in the first year of the simulation, the effects on disease activity as observed in the relevant BLISS trials are applied (BLISS-52 and BLISS-76 in the IV model, and BLISS-SC in the SC model). These can be divided into an effect on total SS score.

B.3.3.3 Change in SELENA-SLEDAI (SS) score at week 52

The methodology used to determine a patient's change in SS score at week 52 is consistent with the previous submission. To robustly determine a patients change in SS score at week 52, it is important to acknowledge the dependence with baseline score, the effect of treatment (whether a patient gets belimumab plus ST or ST alone) and the difference between patients on belimumab with and without a response (defined as a reduction of ≥ 4 points SS at 24 weeks). This is achieved by fitting a linear regression on the pooled BLISS IV trial data (or utilising the BLISS-SC study for SC) that explains the difference between the SS score at baseline and week 52, depending on baseline SS score combined with a treatment indicator variable, and a "response" indicator variable identifying whether or not patients are classified as satisfying the treatment continuation rule at week 24 with belimumab. The results of the regression for estimating change in SS score at Week 52 for the HDA-2 subgroups from the Phase 3 trial data for IV and SC respectively are presented in Table 61.

Table 61. Linear regression explaining change in SELENA-SLEDAI score after 52 weeks compared to ST for the HDA-2 population

Parameter	IV model – HDA-2				SC model – HDA-2			
	Estimate	Std Error	t-value	p-value	Estimate	Std Error	t-value	p-value
SS_0 ST	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SS_0 all belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SS_0 belimumab responders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Note "responders" are patients on belimumab who satisfy the treatment continuation rule.								

B.3.3.4 Treatment continuation probabilities with belimumab and natural discontinuation probabilities

B.3.3.4.1 Treatment continuation probabilities

Reasons for treatment discontinuation in the current submission remain consistent with the reasons for treatment discontinuation provided in the previous submission: natural discontinuation, and no longer deriving clinical benefit from treatment. (Please see Section 6.3.1 of the previous submission provided as part of TA397 for further details).

In brief, in both the IV model and the SC model for both the HDA-1 and HDA-2 subpopulations, patients on belimumab had to satisfy the treatment continuation criterion, defined as demonstrating a SS score decrease of 4 points or greater at week 24. Patients on belimumab who did not satisfy the treatment continuation criterion at week 24 remain in the belimumab arm of the model but continue to receive ST treatments after this time-point and assume the average ST level of disease activity for the remainder of the model horizon.

B.3.3.4.2 Natural discontinuation probabilities for patients receiving belimumab

To derive year 1 natural discontinuation rates for patients receiving belimumab, an analysis for HDA-1 and HDA-2 populations was conducted on the relevant pivotal Phase 3 BLISS trials for each formulation. Kaplan-Meier survival estimates at week 76 were derived from the trial for responders at week 24. A constant daily hazard rate was assumed for belimumab discontinuation during this period (week 24-week 76). Daily hazard rates were converted into a 28-week probability for discontinuing between week 24 and week 52 in the first year. Responders can only discontinue in that time period, as response is defined at week 24. Table 62 presents the percentage of patients continuing treatment with belimumab and the discontinuation rates for patients in the HDA-2 subgroup. Natural discontinuation is not relevant to patients who do not meet the treatment continuation rule at week 24. (The same data is presented for the HDA-1 subgroup in Table 3 of Appendix Q).

As no long term randomised controlled trial exists beyond 76 weeks for the IV trials and 52 weeks for the SC trial, data to calculate the natural discontinuation probability in years subsequent to year 1 were derived from an integrated P2 and P3 LTE studies analyses.

For the IV model, [redacted] of patients who satisfy the treatment continuation rule at week 24 discontinue belimumab in the first year, whilst in subsequent years [redacted] of patients discontinued belimumab. For the SC model, [redacted] of patients who satisfy the treatment continuation rule at week 24, discontinue belimumab in the first year, whilst in subsequent years [redacted] of patients discontinued belimumab. As the model only considers patients who satisfy the treatment continuation rule in the first year, subsequent years treatment continuation rates for non-responders are not relevant to these analyses.

For the HDA-2 population, on IV belimumab and SC belimumab, there were [redacted] and [redacted] of patients who satisfied the treatment continuation rule respectively. [redacted] Table 62. Summary of percentage belimumab continuations and natural discontinuation for HDA-2

	IV model HDA-2 subgroup	SC model HDA-2 subgroup
% belimumab patients satisfying treatment continuation rule at 24 weeks	[redacted]	[redacted]
Natural discontinuation	**Patients satisfying treatment continuation at 24 weeks**	
KM estimate week 76 IV, week 52 SC	[redacted]	[redacted]
Daily hazard rate (wk24-wk76 IV, wk 24-52 SC)	[redacted]	[redacted]
Year 1	[redacted]	[redacted]
Subsequent years	[redacted]	[redacted]

B.3.3.5 Extrapolation to long-term SLE outcomes

As discussed in Section 6.3.1 of the previous submission (TA397), the Phase 3 BLISS-52 and BLISS-76 trials were not designed to capture long-term effects of belimumab due to their relatively short duration. This also applies to the BLISS-SC trial newly presented in this submission. Therefore, GSK examined multiple real-world registries through a literature review, and determined that the Johns Hopkins (JH) cohort was the most appropriate database to develop a natural history model

(NHM) for patients with SLE based on detailed information captured and availability of the dataset. Time to event (TTE) models, were used to identify the relationship between disease activity (SLEDAI) and organ damage or mortality.

In both the IV and SC models, rather than using SS scores to reflect disease severity over time, the scores are used to calculate the Adjusted Mean SLEDAI (AMS) score. The AMS score was developed to measure disease severity over time[90] whereas the SS score only reflects disease activity over the preceding 10 days.

Using the JH cohort data, a Weibull survival model was developed explaining the risk of death with AMS included and SELENA-SELDAI item involvement effects removed. The model does not include the incidence and severity of flares in the disease activity and organ damage models.

Both the IV and SC models use the Johns Hopkins natural history model of SLE to extrapolate to long-term outcomes.

B.3.3.6 Organ damage reduction on belimumab

The original IV cost-effectiveness model presented in TA397 was populated using up to 1.5 years of observed effectiveness data derived from the Phase 3 BLISS-52 and BLISS-76 clinical studies. In the absence of long-term clinical effectiveness data, the corresponding long-term effects on disease progression (e.g. organ damage and mortality), were simulated by using the natural disease history model based on the Johns Hopkins Lupus cohort.

Since then, long-term clinical-effectiveness of belimumab has reported, namely from the long-term extension studies to BLISS-52 and BLISS-76. Further, a propensity score matched (PSM) analysis has been undertaken to estimate the long-term comparative effectiveness of belimumab plus ST compared with ST from a matched population. This has provided the opportunity to validate, and subsequently calibrate, organ damage model results using observed long-term evidence. The steps for its application in the IV and SC model are outlined in the following sections.

B.3.3.6.1 Long-term clinical effectiveness of belimumab

The long-term safety and efficacy of belimumab was observed in two long-term extension (LTE) open-label studies (BLISS-52 and BLISS-76 Non-US patients and BLISS-76 US patients) in patients who previously completed one of the BLISS IV studies (further information about BLISS-52 and BLISS-76 Non-US can be found in Section B.2.6.3.4 and on the BLISS-76 US can be found in Section B.2.6.3.2. The methodology of both studies is described in Section B.2.3.2). Patients who received placebo in the parent study received 10mg/kg belimumab in the continuation study. Patients randomised to receive belimumab continued to receive the same dose as in the parent study (1 or 10 mg/kg IV every 28 days) plus ST. Following a protocol amendment (March 9, 2011), patients receiving 1 mg/kg belimumab had their dose increased to 10 mg/kg. Data on all patients receiving belimumab during this study were pooled for analysis.

The primary analysis of the PSM was conducted using the BLISS-76 US open label extension study population to compare organ damage progression (SDI score) from baseline (defined as first exposure to belimumab) to Year 5 in patients treated with belimumab or ST [48]. The SLICC/ACR Damage Index (SDI) is a measure of organ damage and contains 41 damage items in 12 systems that are specific comorbidities associated with SLE or damage due to toxicity of SLE treatment. Damage items have to persist for a minimum of 6 months or be associated with an immediate pathological scar indicative of damage. The total score is the sum of the marked scores and ranges from 0 to 47. Since damage is irreversible, items that are marked will stay marked for the lifetime of the patient.

In the absence of a control arm, BLISS LTE patients were propensity score matched post-hoc 1:1 to an SLE patient cohort to obtain comparative evidence on organ damage progression compared with ST alone. Following a systematic literature review, the Toronto Lupus Cohort (TLC), was identified as the preferred SLE cohort primarily due to the size of the cohort, the extent of organ damage seen in the patients and the severity of SLE disease activity which was comparable to the BLISS LTE inclusion criteria. Similar to the BLISS trials, ST for patients in the TLC included the use of antimalarials (i.e. hydroxychloroquine), NSAIDs, corticosteroids and

immunosuppressants such as azathioprine and methotrexate[26]. The TLC collected patient data at each visit and at 3–4-month intervals, and the scales used within the TLC for recording disease severity and organ damage progression were similar to those used within the BLISS studies.

The primary end point of the PSM comparative analysis was the difference in change of total SDI score from baseline to 5 years between patients on belimumab compared with those on ST from the TLC in patients with ≥5 years of follow-up. To ensure standardisation in the PSM analysis, baseline in the belimumab treatment arm was defined as first exposure to belimumab. Therefore, the matched belimumab cohort, based on a total of 99 patients, consisted of those who commenced belimumab in the pivotal P3 trials or were switched to belimumab on completion of the ST arm of the P3 trial and switching to belimumab on LTE study.

The results of the PSM analysis (Table 63) demonstrated that over a 5-year period, patients treated with belimumab experienced a five-year SDI change of 0.283 (95% CI 0.166 to 0.400), which represented less organ damage compared with patients treated with ST alone (who had a five-year SDI change of 0.717 [95% CI 0.500 to 0.934).

Table 63. PSM analysis 5-year SDI increase

	Belimumab N=99	**ST** N=99	**Difference**	**p-value**
5-year SDI change [95% CI]*	0.283 [0.166; 0.400]	0.717 [0.550; 0.934]	-0.434 [-0.667; -0.201]	P<0.001
**SDI increase between t=1.5 and y=6.5, as LTE patients already had 52 to 76 weeks of prior treatment. CI, Confidence Interval*				

The methodology of the PSM analysis is described in Section B.2.3.3. For further detailed information on the TLC and the PSM analysis, please refer to the clinical study report (CSR) of the PSM analysis[91].

B.3.3.6.2 Robustness of the PSM analysis

- Patients in the TLC who would have been eligible for belimumab treatment did not receive belimumab solely because it was not available at the time they were enrolled in the patient registry.

- Patients from the TLC were excluded for matching if their baseline data preceded 1990 to enhance comparability of the period of treatment across the groups. Further those with ≥ 15 years of follow-up were also excluded.

- As the BLISS LTEs included patients with different exposure durations to belimumab, for the purpose of the comparative analysis baseline was classified as first exposure time to belimumab.

- Credibility of the PSM analysis in a limited number of matched patients was confirmed by Inverse PS weighting (IPSW) and regression-augmented IPSW sensitivity analyses, which used whole available population samples and produced similar results to the PSM analysis.

- To take into consideration changes in SLE management over the study period, the PS-matched model was re-estimated to adjust for baseline corticosteroid dose, immunosuppressive use and decade of study entry. The change in SDI score from baseline to Year 5 for PS-matched patients was similar to the primary PS-matched analysis which did not adjust for these factors.

- The IPSW method aimed to confirm the robustness of the PSM method. Regression-augmented IPSW was also conducted as an additional sensitivity analysis to overcome any inadequate balance with the IPSW analysis, adding variables with bias >10% as covariates in the regression model. The main PSM methodology demonstrated that, over a 5-year period, patients treated with belimumab experienced less organ damage compared with patients treated with ST alone. The IPSW and regression-augmented IPSW results were similar to the PS-matched results, which demonstrates the robustness of the findings across alternative PS adjustment methodologies.

- Although the primary analysis was conducted on the US cohort only (allowed matching on most predictors), a secondary exploratory analysis was performed on the more geographically dispersed pooled BLISS-52 and non-US BLISS-76 LTE population. The results of the pooled analyses of US and non-US patients were similar.

B.3.3.6.3 Model validation

The cost-effectiveness model was validated by comparing the modelled long-term organ damage progression results to the observed 5-year SDI progression data for belimumab and ST (Table 63).

To ensure comparability of the simulated model results with the long-term evidence the baseline characteristics of the model population were re-adjusted to reflect the BLISS LTE population. An overview of the model settings used for the model validation and supporting rationale is provided in Table 64.

Table 64. Model settings used in the model validation to the PSM

Model setting	Value	Rationale
Subgroup	Total BLISS population	The data from the long-term extension study were on the total patient population, without restrictions in terms of SS score or complement levels.
Responder rule	No	In the open-label extension study, patients were not moved from belimumab to ST if they did not have a treatment response. Therefore, the data from the long-term extension study were on all patients that started in the study, rather than the 24-week responders only. However, patients who had not demonstrated a sufficient response with belimumab during the Phase 3 studies would unlikely have continued into the extension study.
Discontinuation	0%	Reported 5-year SDI change (Table 63) was based on observed cases, who were still on belimumab after 5 years.
Maximum treatment duration Benlysta	Lifetime	In the BLISS open-label extension study, no maximum belimumab treatment duration was in place.
Maximum duration treatment effect	Lifetime	In the BLISS open-label extension study, real world treatment effect was measured, with no maximum treatment duration being in place.

Baseline patient characteristics from the PSM analysis and the cost-effectiveness model were compared. Using only a subset of the BLISS population that completed the BLISS-52 or BLISS-76 (N=99) results in small differences in patient

characteristics. The minor differences in patient characteristics were discussed and were assumed not to impact on the relevance of this validation exercise.

As the IV model captures the observed pooled analysis results from the pooled P3 studies, it was decided that the validation exercise of the deterministic model should be simulated as a 5-year increase in SDI score (further from the baseline duration of 1.5 years). The model starts at the beginning of the BLISS trial, hence the period from 1.5 to 6.5 years from the model was chosen to compare with the PSM analysis results. This simulated an SDI score increase of 0.568 in the belimumab arm and 0.611 in the ST arm, respectively (Table 65, Figure 9).

Table 65. 5-year SDI increases, modelled versus real world data

5-year SDI increase	Belimumab + ST	ST
Cost-effectiveness model; matched LTE ITT population	0.568	0.611
Propensity score-matched analysis	0.283	0.717

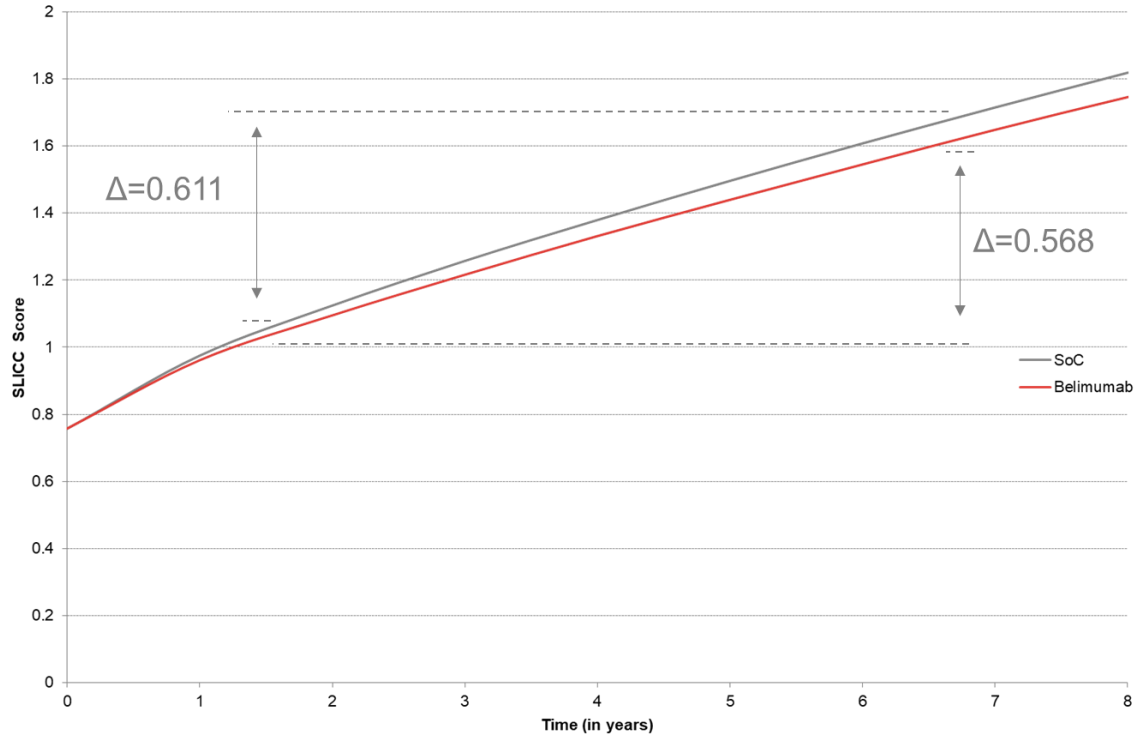


Figure 9 Modelled SDI increases over time

Compared with the results from the PSM analysis, it was apparent that the existing cost-effectiveness model overestimated SDI progression in the belimumab arm and underestimated SDI progression in the ST arm.

This finding is to some extent expected as a recognised area of the underestimated benefit of belimumab in TA397. Organ damage progression is currently incorporated using regression models based on the natural history of SLE patients captured in the Johns Hopkins Cohort and describes the relationship between disease activity and other covariates on the risk of developing organ damage. Therefore, the effect of belimumab on the prevention of organ damage could not be measured directly and so was previously only indirectly captured through the reduction of disease activity. In the absence of evidence from a randomised study, a model validation exercise based on the PSM analysis represents a more robust way to estimate this relationship.

B.3.3.6.4 Model calibration

To account for the difference in the model's predicted SDI progression and the results from the long-term evidence, a calibration factor was derived and applied to allow for adjustment of the existing natural history model in the cost-effectiveness model, and therefore better reflect the findings from the PSM analysis. The adjustments were made by multiplying the original organ damage probabilities from the time-to-event risk equations with the derived calibration factor.

To derive the calibration factor, the model was simulated several times with varying calibration factors, until the model's results matched the observed results from the PSM up to 3 decimals. These analyses were undertaken manually, as the simulation characteristics of the model do not support What-If Analyses. Nonetheless, a starting point for the calibration factor could be derived based on the ratio of the observed and current model outputted values of SDI score after five years (e.g. 0.717/0.611=1.17 for ST). As can be observed in Table 66, the derived ratio and identified calibration factors reflect similar values.

Table 66. Calibrated 5-year increase in SDI score

5-year SDI increase*	Belimumab + ST	ST
Model results with no calibration	0.568	0.611
Observed 5-year SDI increase from PSM	0.283	0.717
Ratio of observed vs. current SDI value	0.498	1.173
Calibration factors	0.491	1.186
Model results with calibration factors	0.283	0.717
* SDI increase between t=1.5 and t=6.5.		

The model calibrations resulted in the amendment of the original organ damage probabilities in the time-to-event risk equations in the model. For each treatment, these were multiplied with the derived calibration factor. For ST, this implies that the annual risk of organ damage for ST was adjusted upwards with 18.6%, in order to reflect the observed long-term organ damage progression after 5 years with ST. On the other hand, for belimumab, this implies that the annual risk of organ damage for belimumab was adjusted downwards with 50.9% in order to resemble the observed long-term organ damage progression after 5 years with belimumab.

Additional model settings have been introduced to enable the user to choose how long the calibration factors will need to be applied for in the simulation. Although it is reasonable to apply the calibration factor to both the standard therapy and belimumab arms of the economic models, in the base-case, a conservative approach is taken where the calibration factor is only applied to belimumab, and only for a period of 6 years to reflect the data collection period of the original Phase 3 BLISS trial and the 5 year open-label study extension. This approach would likely underestimate the incremental benefit of belimumab in terms of long-term organ damage prevention as compared to standard therapy.

B.3.3.6.5 Limitations of the application of the PSM results to the economic analysis

The model validation and calibration exercise whilst clinically plausible is associated with the following limitations:

- The LTE included patients treated with placebo, belimumab 1 mg/kg (until protocol amend), and belimumab 10 mg/kg. For this reason, baseline was defined as time to first exposure to belimumab. Further, the PSM analysis

showed that there was no significant change in time to first SDI change between the treatment arms in the parent study (PSM Analysis CSR)[91]. No information on the differences in total SDI is available.

- The time-to-event risk equations are available by each organ system, but only total SDI values were available for adjustment from the PSM analysis. The assumption was therefore made that the same relative decrease would be applied to each organ system. The relatively few number of patients who were able to be matched for analysis resulted in a reduced power to observe organ-domain specific reductions.

- The model is calibrated to predict the 5-year SDI increase in line (from 1.5 years) with what is observed in the PSM analysis using real world data for belimumab and ST from the LTE belimumab studies and the Toronto Lupus Cohort, respectively. It is unclear how the SDI increase after 5 years (t=6.5 in the model) should be extrapolated beyond this point. Hence, the most conservative assumption was made to apply the calibration factor for a maximum of 6 years.

- The validation of the IV model was undertaken from 1.5 years to 6.5 years and the resultant SLICC score and delta to the 5-year SDI reduction reported from the PSM study were compared. The baseline SDI score in the PSM analysis was at first exposure to belimumab. This means that for patients who were already on belimumab in the BLISS-76 study, the reference period in the model would be model entry till 5 years. Hence, the model validation exercise was also conducted by comparing the SDI score increase at 0 to 5 years. This resulted in a SLICC score delta of 0.682 for belimumab and 0.739 for ST. If this duration from T=0 had been applied, the calibration factors derived would have been 0.41 for belimumab and 0.97 for ST. A conservative approach was taken since these calibration factors would have further increased the organ damage prevention benefit of belimumab relative to ST.

- The validation was conducted on a ‘like-BLISS LTE ITT’ population. The subsequent calibration factors are applied to the more severe populations i.e.

HDA-2 (and HDA-1). We believe that given the clinical benefit (reduction in SS at 24 weeks) is greater in HDA-2 (and HDA-1) populations, compared with the ITT, the application of the calibration based on an ITT population represents a likely conservative approach.

B.3.4 *Measurement and valuation of health effects*

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

Both SF-36 and EQ-5D-3L generic quality of life instruments were collected during the BLISS-52 and BLISS-76 Phase 3 studies. However, no generic quality of life instrument was included in the BLISS-SC study.

As discussed in the previous submission for TA397, the impact on HRQoL is very likely to have been underestimated in BLISS-52 and BLISS-76. Collected EQ-5D-3L values were translated to utility values using the Dolan algorithm[92] to obtain UK general public related scores, and values used to populate the model. Please see Section 6.4.3 of the previous submission for a discussion of the use of instruments, and information on how the results of the EQ-5D-3L for the pooled BLISS-52 and BLISS-76 dataset were used to inform the baseline utilities used in both the IV and SC economic models.

B.3.4.2 Health-related quality-of-life data used in the cost-effectiveness analysis

In brief, in TA397, a statistical model was estimated including all baseline variables (i.e. baseline characteristics, organ damage and organ involvement) that were included in the health economic model. This linear regression was made with the linear mixed effects package in R, correcting for the multiple observations per patient. This analysis included 1,125 patients with 9,051 EQ-5D measurements from BLISS-52 and BLISS-76.

To reduce complexity in calculating utilities due to all types of organ damage, the regression analysis was used to determine a patient's 'clean' utility (U), i.e. free of damage items, using the following equation:

Baseline quality of life was determined by the following regression equation:

$$U = 1.275 - 0.140 * \log e(AGE) - 0.036 * BLACK - 0.009 * SS$$

Where age = current age of patient, black is 1 if a patient is of black African ethnicity, or 0 otherwise, and SS = SELENA-SLEDAI score during the particular model yearly cycle.

Please see Section 6.4.16 of the previous submission for further details of the variables in the equation, how this equation was developed and supporting rationale.

As the BLISS-SC trial did not collect utility data, the current submission uses the same regression equation to estimate utility for patients simulated in both the IV and SC models. This approach is appropriate as recruitment criteria, and the profile of patients recruited between the IV and SC trials and the benefit seen with belimumab for both formulations were very similar.

B.3.4.3 Mapping

As per GSK's submission that formed part of TA397, no mapping techniques were used to transform any of the utilities or quality-of-life data collected in the clinical trials.

B.3.4.4 Health-related quality-of-life studies

Please see Section 6.4.5 of the previous submission for reasons why a formal systematic review for HRQoL data was not conducted. A description of the search process conducted can be found in Section 9.12, Appendix 12 of the previous submission. In brief, a comprehensive systematic literature review was not deemed feasible because of the breadth of organ systems that would need to be searched for.

In the current submission, organ damage utility multipliers were updated through a targeted literature search. Searches were restricted to key organ systems of the SLICC/ACR Damage Index (SDI) which were shown to contribute most to organ damage related loss in quality-adjusted life years (QALYs) and additional costs, based on initial modelling work. Due to restriction of the updated searches to key

organ systems, no searches for the utility multipliers of diabetes, gastrointestinal, ocular, premature gonadal failure and skin organ systems were conducted, and utility multipliers thus remained unchanged. Details are provided in Appendix H.

Utility data was searched in Health Technology Assessments (HTAs) available on the NICE website. If the required information was unavailable from NICE, additional searches were carried out in PubMed.

Of the seven systems that may be subject to organ damage and searched for updated utilities, six yielded new data in items that contributed to the system and resulted in an overall change in utility value for: Cardiovascular, Musculoskeletal, Neuropsychiatric, Pulmonary, Malignancy and Peripheral vascular. Only the Renal system did not yield any new utilities.

Table 67 was originally presented as Table 6.14 in Section 6.4.10 of the previous submission. In the current submission, we present an adapted version of this table showing updated values, (light grey highlight in the relevant cells). Table 67 also shows the items that contribute to each organ damage system. A summary of the updated utility multipliers is provided in Table 67 with a comparison to the original disutility data. Weightings for each of the domains of the organ damage system were calculated and normalised. This approach is consistent with the previous submission.

Both the IV and SC models use the same utility values, regardless of the population or HDA subpopulation under consideration.

Table 67. Summary of quality-of-life values for the cost-effectiveness analysis from the previous submission, updated with values from the current literature search update

<table>
<tr><th rowspan="2">Organ Damage System</th><th colspan="3">Disutilities Year</th><th rowspan="2">SD</th><th rowspan="2">Assumption/justification</th></tr>
<tr><th>1</th><th>2</th><th>Subsequent</th></tr>
<tr><td>Cardio-vascular</td><td>0.779</td><td>0.806</td><td>Same as Y2</td><td>assumed 10%</td><td>Weighted average of:<table><tr><th>Item</th><th>Utility Y1 / Y2</th><th>Weight</th></tr><tr><td>Angina or coronary artery bypass</td><td>0.77 / 0.85</td><td>22%</td></tr><tr><td>Myocardial infarction</td><td>0.949 / 0.963</td><td>25%</td></tr><tr><td>Cardiomyopathy (ventricular dysfunction)</td><td>0.77 / 0.77</td><td>25%</td></tr><tr><td>Valvular disease (diastolic or a systolic murmur > 3/6)</td><td>0.77 / 0.77</td><td>18%</td></tr><tr><td>Pericarditis x 6 months or pericardiectomy</td><td>1 / 1</td><td>10%</td></tr></table></td></tr>
<tr><td>Diabetes</td><td>0.91</td><td>0.91</td><td>Same as Y2</td><td>assumed 10%</td><td>Phase 3 BLISS trials</td></tr>
<tr><td>Gastro-intestinal</td><td>0.79</td><td>0.91</td><td>Same as Y2</td><td>assumed 10%</td><td>Weighted average of:<table><tr><th>Item</th><th>Utility Y1 / Y2</th><th>Weight</th></tr><tr><td>Infarction or resection of bowel below duodenum, spleen, liver or gall bladder ever, for whatever cause (score 2 if > one site)</td><td>0.77 / 0.9</td><td>85%</td></tr><tr><td>resection > 1 site</td><td>0.77 / 0.9</td><td>1%</td></tr><tr><td>Mesenteric insufficiency</td><td>1 / 1</td><td>3%</td></tr><tr><td>Chronic peritonitis</td><td>1 / 1</td><td>3%</td></tr><tr><td>Stricture or upper gastrointestinal tract surgery ever</td><td>1 / 1</td><td>5%</td></tr><tr><td>Pancreatic insufficiency requiring enzyme replacement or with pseudocyst</td><td>1 / 1</td><td>3%</td></tr></table></td></tr>
<tr><td>Malignancy</td><td>0.837</td><td>0.837</td><td>Same as Y2</td><td>assumed 10%</td><td>Malignant tumours (excluding dysplasia) (Score 2 if > one site)</td></tr>
<tr><td>Musculo-skeletal</td><td>0.655</td><td>0.729</td><td>Increasing - See Appendix 9.26</td><td>assumed 10%</td><td>Weighted average of:<table><tr><th>Item</th><th>Utility Y1/Y2</th><th>Weight</th></tr><tr><td>Muscle atrophy / weakness</td><td>1 / 1</td><td>8%</td></tr></table></td></tr>
</table>

<table>
<tr><th rowspan="2">Organ Damage System</th><th colspan="3">Disutilities Year</th><th rowspan="2">SD</th><th rowspan="2">Assumption/justification</th></tr>
<tr><th>1</th><th>2</th><th>Subsequent</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>
<table>
<tr><td>Deforming or erosive arthritis (including reducible deformities, excluding avascular necrosis</td><td>0.645 / 0.662</td><td>19%</td></tr>
<tr><td>Osteoporosis with fracture or vertebral collapse (excluding avascular necrosis</td><td>0.80 / 0.91</td><td>35%</td></tr>
<tr><td>Avascular necrosis</td><td>0.57 / 0.63</td><td>26%</td></tr>
<tr><td>Avascular necrosis 2</td><td>0.57 / 0.63</td><td>2%</td></tr>
<tr><td>Osteomyelitis</td><td>1 / 1</td><td>2%</td></tr>
<tr><td>Ruptured tendon</td><td>1 / 1</td><td>8%*</td></tr>
</table>
</td></tr>
<tr><td>Neuro-psychiatric</td><td>0.713</td><td>0.772</td><td>Same as Y2</td><td>assumed 10%</td><td>Weighted average of:
<table>
<tr><th>Item</th><th>Utility Y1/Y2</th><th>Weight</th></tr>
<tr><td>"Cognitive impairment OR major psychosis"</td><td>0.850 / 0.866</td><td>23%</td></tr>
<tr><td>Seizures requiring therapy for 6 months</td><td>0.78 / 0.78</td><td>14%</td></tr>
<tr><td>Cerebral vascular accident ever or resection (for causes other than malignancy)</td><td>0.63 / 0.69</td><td>28%</td></tr>
<tr><td>Cerebral vascular accident ever or resection >1</td><td>0.57 / 0.62</td><td>1%</td></tr>
<tr><td>Cranial or peripheral neuropathy</td><td>0.867 / 0.929</td><td>31%</td></tr>
<tr><td>Transverse myelitis</td><td>0.427 / 0.741</td><td>3%*</td></tr>
</table>
</td></tr>
<tr><td>Ocular</td><td>0.97</td><td>0.99</td><td>Same as Y2</td><td>assumed 10%</td><td>Weighted average of :
<table>
<tr><th>Item</th><th>Utility Y1 / Y2</th><th>Weight</th></tr>
<tr><td>Cataract</td><td>0.98 / 1</td><td>78%</td></tr>
<tr><td>Retinal damage / optic and trophy</td><td>0.97 / 0.97</td><td>22%*</td></tr>
</table>
</td></tr>
<tr><td>Peripheral vascular</td><td>0.863</td><td>0.873</td><td>Same as Y2</td><td>assumed 10%</td><td>Weighted average of:
<table>
<tr><th>Item</th><th>Utility Y1 / Y2</th><th>Weight</th></tr>
<tr><td>Claudication x 6 months</td><td>0.714 / 0.748</td><td>26%</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><th rowspan="2">Organ Damage System</th><th colspan="3">Disutilities Year</th><th rowspan="2">SD</th><th rowspan="2">Assumption/justification</th></tr>
<tr><th>1</th><th>2</th><th>Subsequent</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>
<table>
<tr><td>Minor tissue loss (pulp space)</td><td>1 / 1</td><td>12%</td></tr>
<tr><td>Significant tissue loss ever (e.g. loss of digit or limb) (Score 2 if > one site)</td><td>0.787 / 0.787</td><td>17%</td></tr>
<tr><td>Significant tissue loss > 1 site</td><td>1 / 1</td><td>0%</td></tr>
<tr><td>Venous thrombosis with swelling, ulceration or venous stasis</td><td>0.99 / 0.99</td><td>46%”</td></tr>
</table>
</td></tr>
<tr><td>Premature gonadal failure</td><td>1</td><td>1</td><td>1</td><td></td><td>No disutility multiplier considered</td></tr>
<tr><td>Pulmonary</td><td>0.713</td><td>0.719</td><td>Same as Y2</td><td>assumed 10%</td><td>Weighted average of:
<table>
<tr><td>Item</td><td>Utility Y1 / Y2</td><td>Weight</td></tr>
<tr><td>Pulmonary hypertension</td><td>0.61 / 0.61</td><td>33%</td></tr>
<tr><td>Pulmonary fibrosis</td><td>0.748 / 0.748</td><td>42%</td></tr>
<tr><td>Shrinking lung (on chest radiograph</td><td>1 / 1</td><td>2%</td></tr>
<tr><td>Pleural fibrosis (on chest radiograph)</td><td>1 / 1</td><td>20%</td></tr>
<tr><td>Pulmonary infarction or resection</td><td>0.735 / 0.866</td><td>4%</td></tr>
</table>
</td></tr>
<tr><td>Renal</td><td>0.972</td><td>0.955</td><td>Over time, the proportion ESRD increases. However, also proportion (successful) transplant increases</td><td>assumed 10%</td><td>Renal consisted of:
Not in –ESRD: 1
Having ESRD
<table>
<tr><td>Utility</td><td></td></tr>
<tr><td>Dialysis</td><td>0.57</td></tr>
<tr><td>Graft transplant</td><td>0.81</td></tr>
<tr><td>Functioning graft (immunosuppression)</td><td>0.81</td></tr>
<tr><td>Graft rejection</td><td>0.57</td></tr>
</table>
</td></tr>
<tr><td>Skin</td><td>0.94</td><td>0.94</td><td>Same as Y2</td><td>assumed 10%</td><td>Weighted average of:
<table>
<tr><td>Item</td><td>Utility</td><td>Weight</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><th rowspan="2">Organ Damage System</th><th colspan="3">Disutilities Year</th><th rowspan="2">SD</th><th rowspan="2">Assumption/justification</th></tr>
<tr><th>1</th><th>2</th><th>Subsequent</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>
<table>
<tr><td></td><td>Y1/Y2</td><td></td></tr>
<tr><td>Scarring chronic alopecia</td><td>0.93/0.93</td><td>47%</td></tr>
<tr><td>Extensive scarring or panniculum other than scalp and pulp space</td><td>0.97/0.97</td><td>36%</td></tr>
<tr><td>Skin ulceration (not due to thrombosis) for more than 6 months</td><td>0.97/0.97</td><td>17%</td></tr>
</table>
</td></tr>
<tr><th>State</th><th colspan="3">Utility Value</th><th></th><th>Assumption/justification</th></tr>
<tr><td>Baseline Utility</td><td colspan="3">0.63 (example A)
0.67 (example B)</td><td>-</td><td>$U = 1.275 - 0.140 * \log e(AGE) - 0.036 * BLACK - 0.009 * SS$

For example: A: for a black African SLE patient, aged 40 years at entry with a SS score of 10

B: for a caucasian patient, aged 40 years at entry with a SS score of 10</td></tr>
<tr><td colspan="6">* Exponentiated to the average number of damage items for patients with damage in that system.
A light grey fill to cells is included to highlight where values have changed</td></tr>
</table>

Table 68. Overview of utility multipliers per organ system, 2010 (previous submission) and 2019 (current submission)

Organ system	Utility multiplier input Year 1		Utility multiplier input Year 2+	
	2010	2019	2010	2019
Cardiovascular	0.717	0.779	0.764	0.806
Diabetes*	0.910	0.910	0.910	0.910
Gastrointestinal*	0.786	0.786	0.906	0.906
Malignancy	0.919	0.837	0.919	0.837
Musculoskeletal	0.665	0.655	0.735	0.729
Neuropsychiatric	0.679	0.713	0.710	0.772
Ocular*	0.974	0.974	0.992	0.992
Peripheral vascular	0.856	0.863	0.919	0.873
Premature gonadal failure*	1	1	1	1
Pulmonary	0.693	0.713	0.693	0.719
Renal	0.972	0.972	0.958	0.955
Skin*	0.943	0.943	0.943	0.943

**These organ systems were not included in the search – cost changes are due to inflation*

A light grey fill to cells is included to highlight where values have changed

B.3.4.5 Adverse reactions

Consistent with the economic model provided as part of TA397, adverse events (AEs) continue not to be included in the IV and SC models. The rationale for this is detailed in Section 6.5.7 of the previous submission and consistent with the findings in BLISS-SC. Briefly, there was limited difference between treatment groups in the BLISS trials in the incidence of all reported adverse events or all serious events and hence there would not be important utility differentiation between the arms in the economic models to warrant inclusion.

B.3.4.6 Mortality

This submission uses the same approach to calculate patient mortality as in the economic analysis described in TA397. Please see Section 6.3.1 of the previous submission for details on mortality and its model implementation. To update the IV and SC models for the current submission mortality data has been updated to the most recent (2016-2018) UK values, as published by the Office for National Statistics (www.ons.gov.uk)[93].

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

The approach for costs and healthcare resource use identification, measurement and valuation in the current submission are consistent with the previous submission (TA397). Costs included in the economic analysis consist of:

- short-term disease activity related costs (based on Phase II trial (LBSL02) resource data)
  (see Section B.3.5.4 Disease related activity costs)
- long-term organ damage costs (derived from the literature) (see Section B.3.5.5)
- belimumab costs (see Section B.3.5.1.2).

Details of how relevant cost and healthcare resource use data for England were identified are provided in Section 6.5 of the company submission of TA397. Costs

related to disease activity were drawn from an analysis conducted in 2009 on the resource utilisation recorded in the one-year belimumab Phase 2 trial in which 2005/06 NHS reference costs were used. The methods used to calculate the disease activity costs from this study are in Section 6.5 of the company submission of TA397.

In this section, we briefly recap the previously used methodology to derive costs for the current economic analysis, methodology used to update costs for the current analysis, and how costs have changed. All costs identified are relevant to both the IV and SC models unless otherwise stated. Costs were updated based on the most recent data available and inflated 2018/2019 UK costs using the Hospital & Community Health Services (HCHS) Inflation Index where appropriate.

B.3.5.1 Intervention and comparators' costs and resource use

B.3.5.1.1 Standard Therapy

As both belimumab IV and SC are added to standard therapy, it is assumed that the costs for standard therapy treatments are negligible and will have little impact on the cost-effectiveness results. Therefore, the cost-effectiveness analysis for the base-case only considers the additional acquisition costs for belimumab.

B.3.5.1.2 Belimumab

Table 69 summarises the medicines acquisition cost and associated administration cost for Belimumab IV and Belimumab SC, respectively.

B.3.5.1.2.1. Cost of Belimumab IV administration

Patients require 14 infusions in Year 1 and 13 in Year 2 onwards. Two hours are required for the administration of belimumab IV via infusion. One hour is required for the actual infusion and another hour for patient preparation and monitoring post-infusion.

The administration cost for belimumab IV in the current appraisal, £154, is consistent with TA247 'Tocilizumab for the treatment of rheumatoid arthritis' and reflects the Committees preferences as per the Final Appraisal Determination of TA397.

B.3.5.1.2.2. Cost of Belimumab SC administration

It is assumed that patients require up to an hour with a specialist nurse within the first year of receiving belimumab SC in order to receive training and education on how to self-administer, and to monitor for adverse reactions that may occur.

According to Section 9 of the PSSRU Unit Costs of Health & Social Care 2019[94], the cost per hour of a Band 6 specialist hospital-based nurse is £113.00. Once a patient is trained in self-administrating with the pre-filled pen/auto-injector device, it is assumed there are no further costs associated with belimumab SC administration.

Table 69. Unit costs associated with the technology in the intravenous (IV) and subcutaneous (SC) economic model for the HDA-2 subgroup

	Belimumab 10mg/kg IV		Belimumab 200mg SC	
Items	**Cost**	**Further information**	**Technology cost**	**Further information**
Mean cost of technology treatment based on an average weight of 70.4 kg as seen in the BILAG biologics registry	Year 1 annual cost of belimumab = [redacted] Year 2 annual cost of belimumab = [redacted]	The list price for the vials are £121.50 and £405.00 for 120 mg and 400 mg, respectively. The PAS price for the vials are [redacted] and [redacted] for 120 mg and 400 mg, respectively. For each weight, the optimal vial combination is chosen and costs for waste are added. Weight distribution according to BILAG-BR is used to determine average yearly belimumab costs.	Year 1 annual cost of belimumab = [redacted] Year 2 annual cost of belimumab = [redacted]	The list price per prefilled pre-filled pen is [redacted]. The PAS price per prefilled device is [redacted] Unlike Belimumab IV, all patients receive a single belimumab 200mg subcutaneous pre-filled pen as this formulation is not dependent on patient weight for dosage.
Administration cost per infusion (IV) or injection (SC)	Year 1 infusion cost £2,156 Year 2+ infusion cost £2,002	Patients require 14 infusions in Year 1 and 13 in Year 2 onwards. The administration cost of for belimumab IV in the current appraisal, £154, is consistent with TA247 'Tocilizumab for the treatment of rheumatoid arthritis' and reflects the Committees preferences as per the Final Appraisal Determination of TA397.	Year 1 administration cost £113.00 Year 2+ administration cost £0	The model assumed that patients received 53 belimumab 200mg subcutaneous pre-filled pen in the first year and each year thereafter, with one self-administered each week by the patient. In the first year, it is assumed that patients receive up to an hour with a specialist nurse to receiving training on administration technique. Once a patient is trained in self-administering the SC pre-filled pen, it is assumed there are no further costs associated with Belimumab SC administration. Up to an hour with a specialist nurse within the first year of receiving belimumab SC to receive training and education on how to self-administer effectively, and how to assess and respond to any adverse reaction. 1-hour cost of specialist hospital-based nurse: £113.00 Self-administration per injection by a patient: £0

Monitoring and test cost	£0	No additional monitoring or tests are required for implementation of this technology	£0	No additional monitoring or tests are required for implementation of this technology
Total Year 1 costs	[redacted]		[redacted]	
Total Subsequent Year costs	[redacted]		[redacted]	

B.3.5.1.3 Patient access scheme

Mindful of NHS resources, GSK is proposing a patient access scheme (PAS) for both belimumab IV and SC, designed to support medicine access in England and Wales and reflect both the value GSK believes to be inherent in this technology and the data that supports it.

B.3.5.2 Health-state unit costs and resource use

As the IV model and the SC model do not include health states, costs have been presented in terms of short-term disease activity related costs and long-term organ damage costs.

B.3.5.3 Adverse reaction unit costs and resource use

Consistent with the economic model provided as part of TA397, adverse events (AEs) are not included in the IV and the SC models included with this submission. The rationale for this is detailed in Section 6.5.7 of the previous submission and is also consistent with findings from the BLISS SC study. Briefly, there was limited difference between treatment groups in the BLISS trials in the incidence of all reported adverse events or all serious events and hence we continue to expect there would not be an important cost differentiation between the arms in the health economic model with regards to adverse events.

B.3.5.4 Disease related activity costs

In the previous submission, costs related to disease activity were drawn from an analysis conducted in 2009 on the resource utilisation recorded in the belimumab Phase 2 trial (Please see Section 6.5.1 of the previous submission for further details). Resource utilisation items included in this study were:

- Number of surgeries or procedures
- Number of Accident and Emergency attendances
- Number of days in a nursing home or rehabilitation centre
- Number of overnight hospitalisations
- Length of stay in hospital

- Number of visits to health professionals
- Number of tests or diagnostic procedures

Each research item was costed based on 2005/2006 NHS reference costs and inflated to 2010 values in accordance with the year of the previous model. As part of the current update, these previous 2005/2006 values are inflated to 2018/2019 values, using the consumer price index for health as published by PSSRU in 2019[94]. These values are shown in Table 70.

Table 70. Overview of disease activity related costs, 2005/2006 and 2018/2019

SELENA-SLEDAI Score	Yearly Costs		
	2005/2006	2009/2010	2018/2019
0	£1030.12	£1,152.44	£1294.53
1	£1149.39	£1,285.87	£1444.42
2	£1268.66	£1,419.30	£1594.30
3	£1353.17	£1,513.84	£1700.50
4	£1402.87	£1,569.44	£1762.95
5	£1452.56	£1,625.04	£1825.40
6	£1502.26	£1,680.64	£1887.86
7	£1551.96	£1,736.23	£1950.31
8	£1601.65	£1,791.83	£2012.76
9	£1659.65	£1,856.72	£2085.65
10	£1725.91	£1,930.85	£2168.92
11	£1792.18	£2,004.98	£2252.19
12	£1858.44	£2,079.11	£2335.46
13-20	£1924.72	£2,153.26	£2418.76

B.3.5.5 Organ Damage Costs

Organ damage costs in the model in the previous submission were obtained from a targeted literature search of each of the 41 damage items over the twelve key organ systems in the SDI score. To calculate the average cost per patient for each of the organ damage systems, the frequency of each constituent medical condition of each organ system (as shown in Table 67) was multiplied by the full cost incurred for a single patient for year 1 and year 2 onwards separately.

In the current submission, organ damage costs were updated through a targeted literature search. Searches were restricted to key organ systems of the SDI which

were shown to contribute most to organ damage related loss in quality-adjusted life years (QALYs) and additional costs, based on initial modelling work. Due to restriction of the update searches to seven key organ systems, no searches for costs relating to diabetes, gastrointestinal, ocular organ systems were conducted. Costs for these organ systems were inflated to 2018/2019 values, using the consumer price index for health as published by PSSRU in 2019[94]. Details of the updated literature search are provided in Appendix H.

No searches were conducted for premature gonadal failure and skin organ systems, as these were also not searched previously, and costs remained zero. Furthermore, gastrointestinal and malignancy costs for year 2 onwards also remained zero.

Due to all costs, newly identified or previously identified, being inflated to 2018/19 as necessary, all costs for use in the current 2018/2019 analysis are higher than the costs used for the previous submission, except premature gonadal failure and skin which remain zero.

An overview of cost inputs per organ system is shown in Table 71.

Table 71. Overview of cost inputs per organ system, 2010 and 2018/2019

Organ system	Update search conducted for 2018/2019	Method applied	Cost input Year 1		Cost input Year 2+	
			2010	2018/2019	2010	2018/2019
Cardiovascular	YES	New reference for myocardial infection (MI) costs.	£3,440	£4,692	£505	£1,297
Diabetes	NO	-	£2,338	£2,658	£2,338	£2,658
Gastrointestinal	NO	-	£2,708	£3,097	£0	£0
Malignancy	YES	No new references identified.	£6,123	£7,096	£0	£0
Musculoskeletal	YES	New reference for erosive arthritis costs.	£5,431	£7,180	£1,903	£2,386
Neuropsychiatric	YES	Introduced cognitive impairment costs. New references for major psychosis and transverse myelitis costs.	£3,660	£6,821	£1,144	£2,786
Ocular	NO	-	£1,535	£1,810	£17	£24
Peripheral vascular	YES	Introduced claudication costs. New reference for significant tissue loss costs.	£2,988	£3,280	£598	£709
Premature gonadal failure	NO	-	£0	£0	£0	£0
Pulmonary	YES	Introduced pulmonary fibrosis costs.	£9,679	£14,888	£9,603	£14,937
Renal	YES	Renal costing model was updated with new transitional probabilities and new costing referencing	£1,765	£2,467	£2,453	£3,641
Skin	NO	-	£0	£0	£0	£0
Where appropriate, inflation has been applied to cost						

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

B.3.6.1 Summary of base-case analysis inputs

A summary of the input parameters of the economic models are shown in Table 71. All values are used in both the IV and SC economic models, unless stated otherwise.

Table 71. Summary of variables applied in the economic models for the HDA-2 population

	IV model with HDA-2 population			SC model with HDA-2 population		
Variable	**Value (reference to appropriate table or figure in submission)**	**Measurement of uncertainty and distribution: CI (distribution)**	**Reference to section in submission**	**Value (reference to appropriate table or figure in submission)**	**Measurement of uncertainty and distribution: CI (distribution)**	**Reference to section in submission**
Patient characteristics at baseline						
Age (years)	[redacted]	[redacted]	Table 58	[redacted]	[redacted]	Table 58
Percentage females (%)	[redacted]	[redacted]		[redacted]	[redacted]	
Percentage Black Ethnicity (%)	[redacted]	[redacted]		[redacted]	[redacted]	
SLE disease duration (yrs)	[redacted]	[redacted]		[redacted]	[redacted]	
SLICC/ACR damage index score	[redacted]	[redacted]		[redacted]	[redacted]	
Baseline disease activity parameters and steroid use simulated at baseline						
Baseline SLEDAI	[redacted]	[redacted]	Table 59	[redacted]	[redacted]	Table 59
Increased DNA binding	91.4%	0.914, Bernoulli		92.4%	0.924, Bernoulli	
Low Complement	83.1%	0.831, Bernoulli		66.6%	0.666, Bernoulli	
Vasculitis	11.8%	0.118, Bernoulli		10.5%	0.105, Bernoulli	
NP involvement	0.6%	0.006, Bernoulli		0.0%	0.000, Bernoulli	
Renal involvement	6.4%	0.064, Bernoulli		4.8%	0.048, Bernoulli	
Serositis involvement	1.1%	0.011, Bernoulli		5.7%	0.057, Bernoulli	
Haematological Involvement	6.4%	0.064, Bernoulli		1.8%	0.018, Bernoulli	
Skin Involvement	57.0%	0.57, Bernoulli		77.6%	0.776, Bernoulli	
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	
Linear regression explaining change in SELENA-SEDAI score after 52 weeks compared to ST						

SS^0 ST	[redacted]	[redacted]	Table 61	[redacted]	[redacted]	Table 61
SS^0 all belimumab	[redacted]	[redacted]		[redacted]	[redacted]	
SS^0 belimumab responders	[redacted]	[redacted]		[redacted]	[redacted]	
Summary of percentage belimumab continuations and natural discontinuation						
% belimumab patients satisfying treatment continuation rule at 24 weeks	[redacted]	[redacted]	Table 62	[redacted]	[redacted]	Table 62
Natural discontinuation rate for patients satisfying treatment continuation criteria at 24 weeks in Year 1	[redacted]	[redacted]		[redacted]	[redacted]	
Natural discontinuation rate for patients in year 2 and subsequent years	[redacted]	[redacted]		[redacted]	[redacted]	
Calibration factors						
Belimumab + ST	[redacted]	[redacted]	Table 66	[redacted]	[redacted]	Table 66
ST	[redacted]	[redacted]		[redacted]	[redacted]	Table 66

Utility multipliers per organ system (SD for all organ systems assumed at 10%)

Organ System	**Year 1**	**Year 2**	**Reference**	
Cardiovascular	0.717	0.779	Table 68	These utility multipliers are used in both the IV and SC models
Diabetes	0.910	0.910		
Gastrointestinal	0.786	0.786		
Malignancy	0.919	0.837		
Musculoskeletal	0.665	0.655		
Neuropsychiatric	0.679	0.713		
Ocular	0.974	0.974		
Peripheral vascular	0.856	0.863		
Premature gonadal failure	1	1		
Pulmonary	0.693	0.713		
Renal	0.972	0.972		
Skin	0.943	0.943		

Model cost inputs

Type	**Cost**	**Varied in PSA**	**Reference**	**Used in IV model**	**Used in SC model**	
Belimumab 120mg vial	[redacted]	No	Table 69	Yes	No	
Belimumab 400mg vial	[redacted]	No		Yes	No	

Belimumab 200mg subcutaneous prefilled pre-filled pen	[redacted]	No		No	Yes	
Admin cost per IV infusion	£154	No		Yes	No	
Cost of specialist hospital-based nurse per hour to deliver SC training	£113	No		No	Yes	
Other variables used in the model						
Item	**Value**	**Varied in PSA**	**Reference**	**Used in IV model**	**Used in SC model**	
Number of IV infusions in year 1	14	No	Table 69	Yes	No	
Number of IV infusions in year 2 onwards	13	No	Table 69	Yes	No	
Exposure to drug	100%	No	Table 73	Yes	Yes	
Vial sharing	Off	No	Table 73	Yes	No	
Average weight	70.4kg	No	B.3.2.3.1.1	Yes	No	
Discount rate for costs	3.5%	No	B.3.2	Yes	Yes	
Discount rate for effects	3.5%	No	B.3.2	Yes	Yes	
Disease activity related costs per year 2018/2019						
SELENA-SLEDAI Score	**Yearly cost**	**Reference**				
0	£1294.53	Table 70				
1	£1444.42					
2	£1594.30					
3	£1700.50					
4	£1762.95					
5	£1825.40					
6	£1887.86					
7	£1950.31					
8	£2012.76					
9	£2085.65					
10	£2168.92					
11	£2252.19					
12	£2335.46					
13 - 20	£2418.76					
Johns Hopkins cohort characteristics						
Item	**Value**	**Reference**				
Number of patients	1282					

Females	1,190 (92.8%)	Table 6.7 of the previous submission				
Black ethnicity	492 (38.4%)					
Caucasian	672 (52.4%)					
Age at diagnosis (mean (SD)	33.1 (13.0)					
Age at cohort entry (mean (SD)	38.2 (12.8)					
Disease duration at cohort entry (mean (SD)	5.15 (6.5)					
SLEDAI score at first visit (mean (SD)	3.32 (3.7)					
Steroid dose at first visit (mean (SD)	9.95 (15.3)					
Past smoker (%)	38.9%					
Hypertension (yearly risk)	15.8%					
Anti-cardiolipin antibodies positive (%)	3.0%					
Lupus anticoagulant positive (%)	9.6%					
Coefficient results for the linear regression model predicting change in mean SLEDAI – Johns Hopkins Cohort						
Item	**Coefficient**	**95% Confidence Intervals**		**Reference**		
Mean SLEDAI score in previous period	-0.4163	-0.4396	-0.3929	Table 6.9 of the previous submission		
Male gender	-0.0991	-0.2544	0.0562			
Black ethnicity	0.3524	0.2566	0.4482			
Log of age	-0.3586	-0.5072	-0.2100			
Constant	2.0577	1.4855	2.6299			
Sigma ui	0.4093					
Within R^2	0.3624					
Overall R^2	0.1668					
Linear regression model explaining average steroid dose per year (mg/day) based on SLEDAI score (model input) - Johns Hopkins cohort						
Regression parameter	**Coefficient (95% CIs)**	**P-value**	**Reference**			
Average SLEDAI score during current year	0.7199 (0.617, 0.823)	<0.001	Table 6.11 of the previous submission			
Constant	3.410 (3.073,3.747)	<0.001				
Weibull survival model explaining risk of death with AMS included and item involvement effects removed - JH cohort						

Covariates	Model coefficient	Reference				
Constant	-10.366	Table 6.12 of the previous submission				
Black ethnicity	0.7814					
Age at diagnosis	0.0321					
Cholesterol	0.0044					
AMS over lifetime	0.2135					
Cumulative Average Prednisone Dose (mg/month)	0.0012					
Renal damage	0.652					
Musculoskeletal damage at previous visit	0.415					
Peripheral vascular damage at previous visit	0.9783					
Gastrointestinal damage at previous visit	0.4684					
Diabetes at previous visit	0.6764					
Malignancy at previous visit	1.1489					
Any infection at time of death at current visit	0.7409					
Parametric distribution parameter for Weibull	1.6799					
Average SLICC scores per organ - based on all recordings in Johns Hopkins cohort						
Organ	**Score**	**Reference**				
Cardiovascular	1.42	Table 6.16 of the previous submission				
Diabetes	1.00					
Gastrointestinal	1.09					
Malignancy	1.00					
Musculoskeletal	1.41					
Neuropsychiatric	1.37					
Ocular	1.23					
Peripheral vascular	1.21					
Premature gonadal failure	1.00					
Pulmonary	1.31					
Renal	1.83					
Skin	1.14					

Abbreviations: CI, confidence interval; SD, standard deviation; SE, standard error; ST, Standard Therapy

B.3.6.2 Assumptions

All assumptions underlying the statistical methodology were described in Section 6.3.1 of the previous company submission provided as part of TA397. A summary of assumptions concerning the technology and its application in the economic analyses are presented in Table 73.

Table 73. Summary of assumptions used in the base-case economic model

Assumption	Implementation	Justification	Reference
Subgroup for primary economic analysis for each formulation is HDA-2 population	Subgroup set to HDA-2	Following advice from experts, it is understood there are SLE patients who are considered to be HDA with an SS score of ≥10 and at least one of the following serological markers: low complement or positive dsDNA.	Section B.2.7 Section B.3.2.1
Exposure to belimumab is assumed to be 100% for both IV and SC formulations	Exposure set to 100% in model	This assumption is reasonable for belimumab IV, as this formulation is delivered in a clinic setting. For the SC formulation, 100% exposure is also assumed to reflect the patient's motivation to prevent disease flares.	Section B.3.2.3.1.1 Section B.3.2.3.1.2
Patient weight distribution for the IV models is based data from the BILAG-BR registry	BILAG-BR data listing for patient weight are used as the data source	The use of patient weights from the BILAG-BR registry is appropriate as it reflects real-world usage in a UK-based population for whom this technology is being evaluated in.	Section B.B.2.3.4 and Appendix P
Vial sharing is assumed not to occur with belimumab IV	Vial wastage is 'on' in the model	Belimumab vials for IV infusion are provided on a named patient basis and are sensitive to light.	Section B.3.2.3.1.1
Patients on belimumab SC require three appointments with a specialist nurse in the first year of receiving the formulation only	Patients in the SC model receive an hour with a specialist nurse within the first year of receiving Belimumab SC in order to receive training and education on how to self-administer effectively, and	Patients are assumed to require training of how to self-administer the belimumab SC pre-filled pen. After these initial training sessions, it is assumed that the patients are competent to self-administer, as they will be doing this on a weekly basis.	Section B.3.2.3.1.2

	to assess any adverse reaction that are apparent on immediate administration.		
Patients receive belimumab and associated treatment effects for a lifetime horizon	'Maximum duration Belimumab treatment' and 'Duration maximum effect Belimumab' set to lifetime duration in the model.	There is no limitation in the product licence for how long patients may take belimumab assuming patients continue to receive benefit (beyond the 24 week continuation criterion). In the BLISS open-label extension study, real world treatment effect was measured, with no maximum treatment duration stipulated	SmPC for Belimumab IV and SC
The incorporation of the PSM findings through means of a calibration factor for long term organ damage is applied for 6 years and only to belimumab	PSM calibration factor for long term organ damage set to 6 years for belimumab only	The PSM analysis is applied (as a calibration factor) for long term organ damage for 6 years only to belimumab (based on observed data). It assumes equal impact on all components of the SDI sub-domains when the calibration factor is applied. The PSM analysis was based on approximately 1 year of RCT data and 5 years of long-term extension data. However, this is still considered a very conservative approach, as we would expect benefit to continue for as long as patients continue to take belimumab.	Section B.3.3.6
No treatment waning effect to the calibration factors is applied in the model	Duration of waning of calibration factors in the model set to zero	This assumption is conservative in nature and is in place due to no supporting evidence to show the contrary.	
Patients on belimumab must satisfy the treatment continuation criterion (demonstrating a SS score decrease of 4 or greater) at week 24 to continue treatment with belimumab. If patients do not satisfy the treatment continuation criteria, patients continue to receive ST treatments and assume the	Responder rule in the model set to 'SS reduction ≥ 4 at week 24'	The Final Appraisal Determination for TA397 stated that patients must have a SS reduction ≥ 4 at week 24 as a condition to remain on belimumab.	Section B.3.3.6.2 NICE TA397 Final Appraisal Determination

average ST level of disease activity for the remainder of the model horizon.			
Treatment discontinuation rates are taken from the BLISS trials	Data in the model reflects the HDA subgroup from the relevant BLISS formulation related trials.	It is reasonable to derive year 1 treatment discontinuation rates from the appropriate subgroups from the pivotal Phase 3 BLISS clinical trials for the IV and SC formulations. Annual discontinuation rates for year 2 onwards in HDA-1 and HDA-2 are assumed to be [redacted] and [redacted] respectively. These values are derived from an integrated P2 and P3 LTE analyses.	Section B.3.3.4
No treatment waning effect is applied in the model	Duration of waning of belimumab treatment effect is set to zero	A treatment waning effect is not applied as there is no evidence to date to support this assumption. This is consistent with TA397.	Table 57
The natural history model is based on Johns Hopkins cohort with SLEDAI involvement removed. However, adjusted mean SLEDAI (AMS) is added in.	Natural History Model is set to 'JH - AMS forced in, involvement removed'	The Johns Hopkins cohort was chosen to reflect the natural history model due to its availability, large size, and containing all the data required to conduct the appropriate level of analysis. There is no information on the natural history development of the SELENA-SLEDAI item involvement in the Johns Hopkins data. The AMS allows disease severity to be simulated over time.	Section 6.2.2 of the previous submission.
The long-term disease activity model is based on the Adjusted Natural History Model	The Adjusted Natural History Model is selected for the long-term disease activity model.	As detailed in the original submission, the constant value in the "disease activity model" constructed from the JH data to relate disease activity to risk of longer term organ damage was increased from 2.058 to 3.0 to better reflect the higher level of disease activity in the UK HDA sub-populations. The Johns Hopkins cohort was chosen to reflect the natural history model due to its availability, large size, and containing all the data required to conduct the appropriate level of analysis.	Section 6.3.1 of the previous submission

Steroid use is based on the John Hopkins linear regression l model	JH Linear regression model explaining average steroid dose per year(mg/day) based on SLEDAI score (model input)	As detailed in the original submission, due to the double-blind nature of the P3 studies, clinicians were reticent to make significant reductions in steroid doses hence any potential steroid sparing benefit with belimumab was likely under-estimated. Hence the relationship between disease activity level and steroid dose from the JH cohort was considered more appropriate for the model.	Section 6.3.1 of the previous submission.
Oral corticosteroid related adverse event costs are excluded in the base-case.	Oral corticosteroid related adverse event costs are set to zero	As both patients on belimumab and ST in the model receive ST, it is assumed that both groups of patients receive the same levels of oral corticosteroids (OCS). However, this assumption is very conservative, as belimumab has demonstrated steroid sparing effects.	Section B.2.6.4.1
Adverse events are excluded from the model	Not included in the model.	There was little difference between treatment groups in the BLISS trials in the incidence of all reported adverse events or all serious events and hence there would not be an important cost and utility differentiation between the arms in the health economic model with regards to adverse events.	Section 6.5.7 of the previous submission.
Disease flares are not simulated in the model	Not included in the model.	The SELENA-SLEDAI Flare Index (SFI) was not collected in the JH database. An alternative measure of flare could have been used however this may have caused problems due to the correlation between flare and SLEDAI score. As the model uses the adjusted mean SLEDAI, disease activity is 'smoothed' over time, and a flare or relapse of activity cannot be shown. A decrease in frequency of flares due to belimumab will however also decrease the AMS over the treatment period. However, the use of AMS may have underestimated the benefit of belimumab in reducing flares and therefore this methodology is considered conservative.	Section 6.3.1 of the previous submission.

B.3.7 *Base-case results*

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

The base-case results for the HDA-2 population show that with the proposed patient access scheme, belimumab IV and belimumab SC are both a cost-effective use of NHS resources and importantly, more clinically appropriate SLE patients with HDA could potentially have access to this licensed treatment. With the PAS offered by GSK, compared to ST alone, add-on belimumab IV leads to incremental costs of [redacted], [redacted] additional life years and [redacted] additional QALYs (discounted), resulting in an incremental cost effectiveness ratio (ICER) of £30,001 per QALY gained. Use of add-on belimumab SC leads to, incremental costs of [redacted], [redacted] added life year and [redacted] added QALYs (discounted), resulting in an ICER of £30,566 per QALY gained.

For completeness, the results for the HDA-1 population show that with the proposed patient access scheme, belimumab IV and belimumab SC are also both a cost-effective use of NHS resources. With the PAS offered by GSK, compared to ST alone, add-on belimumab IV leads to incremental costs of [redacted], [redacted] additional life years and [redacted] QALYs (discounted), resulting in an incremental cost effectiveness ratio (ICER) of £28,361 per QALY gained. Use of add-on Belimumab SC leads to, incremental costs of [redacted], [redacted] added life years and [redacted] added QALYs (discounted), resulting in ICER of £29,910 per QALY gained.

Fully incremental results for belimumab IV and SC relative to ST, with the proposed PAS price are shown in Table 74 and Table 75 for the HDA-2 and HDA-1 patient subgroups, respectively.

Table 74. Base-case results for HDA-2 population

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
IV model - [redacted]							
ST	£160,470	16.90	9.81				
Belimumab IV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£30,001
SC model - [redacted]							
ST	£151,999	17.12	10.06				
Belimumab SC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£30,566
All model outcomes presented are discounted. *Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years*							

Base-case results for HDA-1 are presented in Table 75. Please see Appendix Q for further results for HDA-1.

Table 75. Base-case results for HDA-1 population

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
IV model - [redacted]							
ST	£166,658	17.47	10.22				
Belimumab IV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£28,361
SC model - [redacted]							
ST	£156,692	17.68	10.48				
Belimumab SC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£29,910
All model outcomes presented are discounted. *Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years*							

B.3.7.2 Further IV model results for HDA-2

[redacted]

Figure 10. Belimumab discontinuation for the IV model - HDA-2

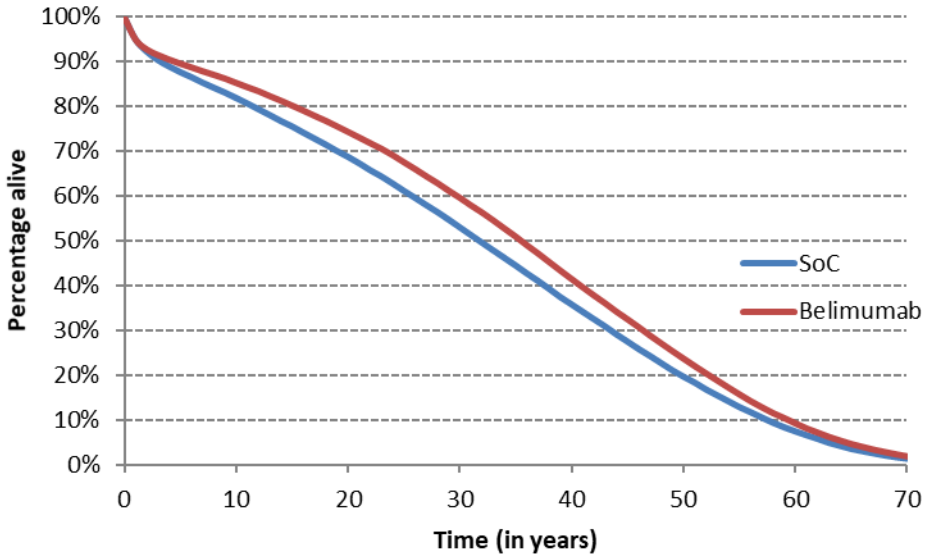


Figure 11. Survival for the IV model - HDA-2

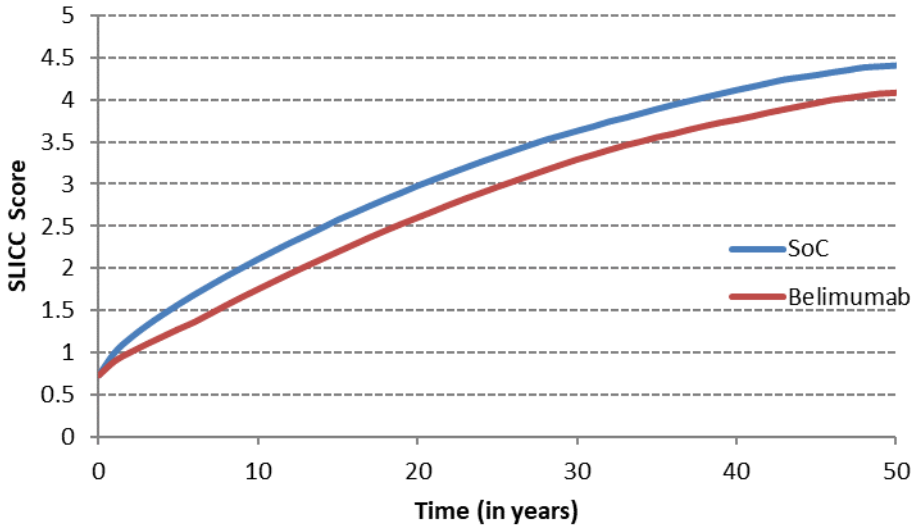


Figure 12.Average SLICC over time for the IV model - HDA-2

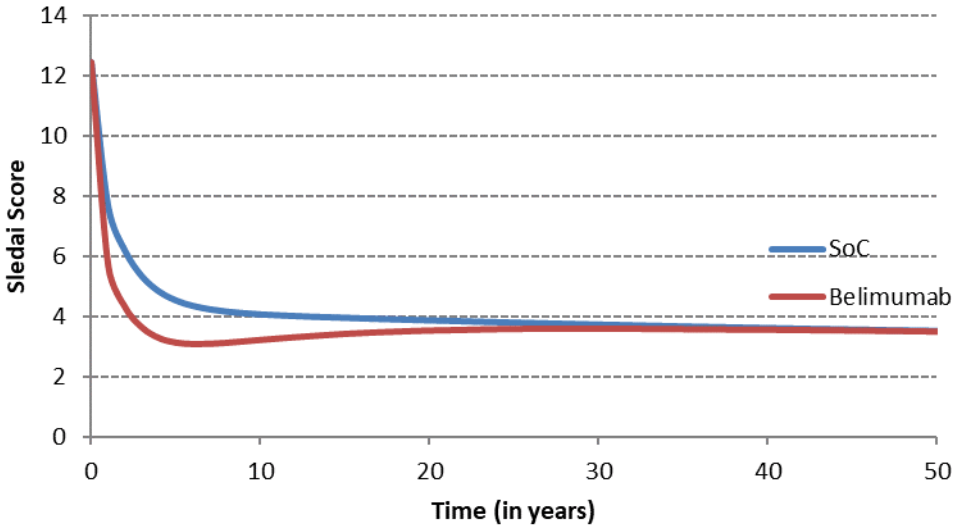


Figure 13. Average SELDAI over time for the IV model - HDA-2

B.3.7.3 Further SC model results for HDA-2

Figure 14. Belimumab discontinuation for the SC model - HDA-2

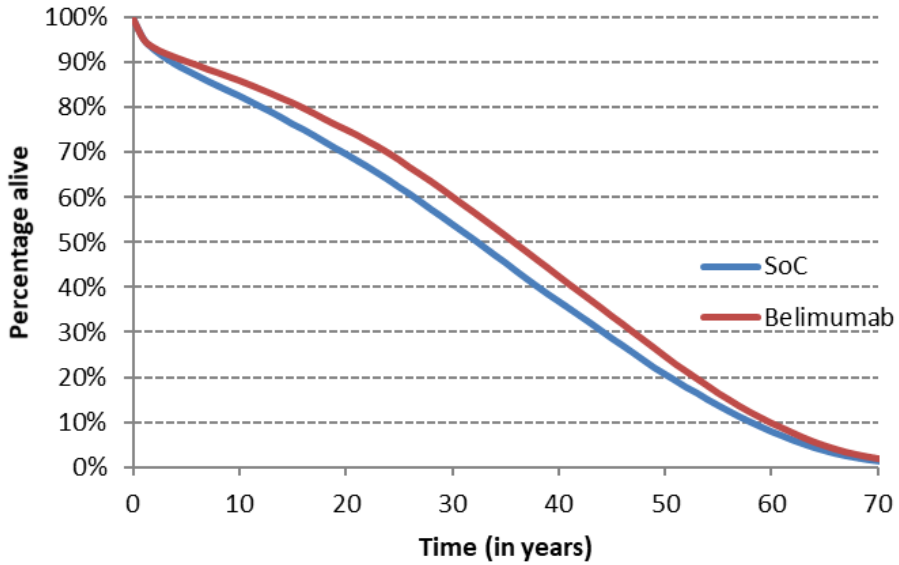


Figure 15. Survival for the SC model - HDA-2

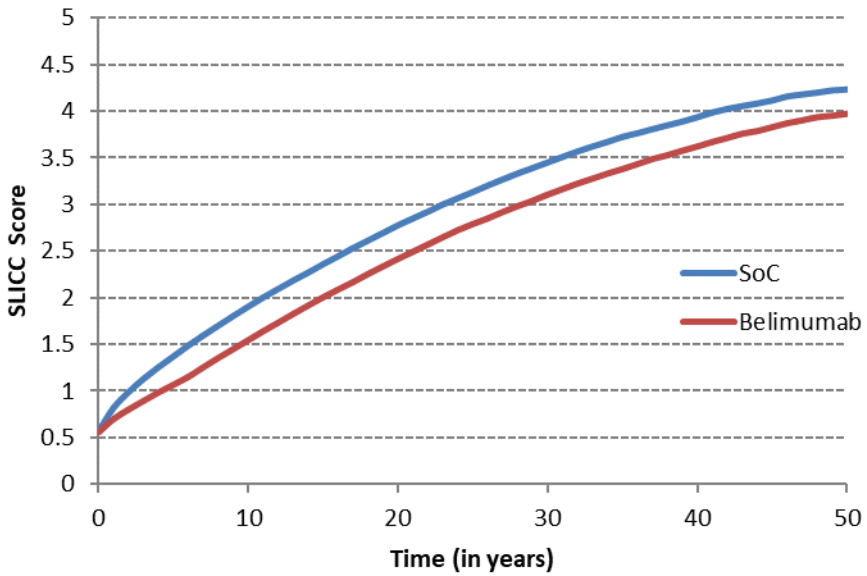


Figure 16.Average SLICC over time for the SC model - HDA-2

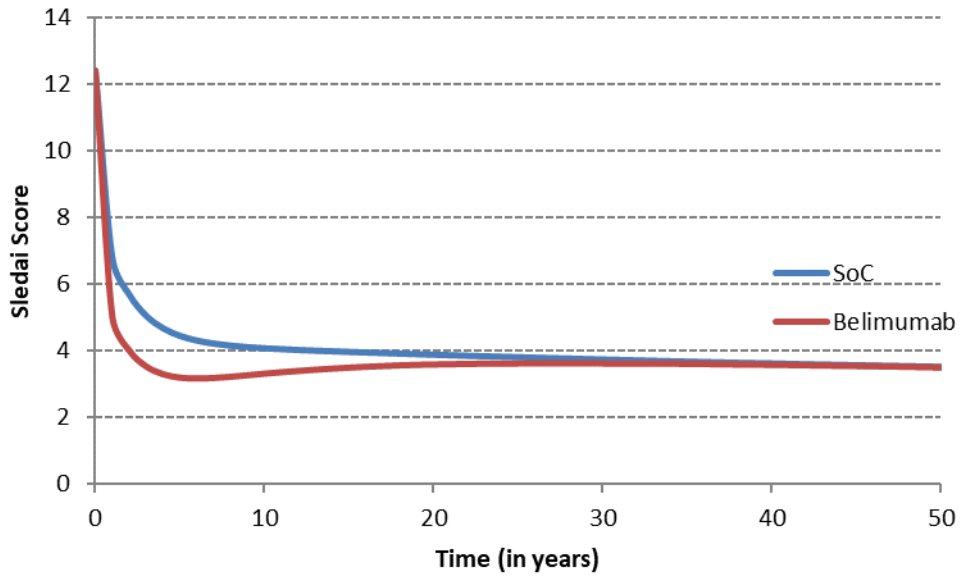


Figure 17. Average SELDAI over time for the SC model - HDA-2

B.3.8 *Sensitivity analyses*

Uncertainty around structural assumptions has been examined using both one-way and probabilistic sensitivity analysis within the populations of interest to this decision problem for both the IV and SC formulations of belimumab.

To test the robustness of model assumptions and parameters, the effect of changing parameters in one-way sensitivity analyses was examined. Effects of varying individual parameters was explored using 95% confidence intervals. Sensitivity results for each input were ranked from most sensitive to least sensitive and those that had the greatest effect were plotted on tornado diagrams. Analysed parameters, their base-case values, uncertainties and distributions are presented in Table 71.

B.3.8.1 Probabilistic sensitivity analysis

Probabilistic sensitivity analyses in the current submission were performed by varying the same sets of model parameters detailed for the univariate sensitivity analyses, simultaneously 1000 times to understand the impact on the cost per QALY results. There was a large amount of correlation between coefficients within each regression (i.e. parameters from the regressions numbered 1, 2, 4, and 6 for the sensitivity analyses in Section 6.6.2 of the previous submission). To account for this correlation the covariance matrices were generated and from these a set of PSA inputs were used. This process uses a multivariate normal distribution; a normal distribution was therefore assigned to these regressions in the PSA. The standardised mortality reported by Bernatsky et al (2006)[39] was assumed to follow a normal distribution. The costs were assigned a gamma distribution as recommended by Briggs et al. 2006[95].

Results of probabilistic sensitivity analyses for the HDA-2 subgroup are presented in Table 76. Scatter plots are shown in Figure 18 and Figure 20 for the IV and SC models respectively, and cost-effectiveness acceptability curves are shown in Figure 19 and Figure 21 for the IV and SC models respectively. Parameters included in the probabilistic sensitivity analysis, their base-case values, and their assumed distribution, are presented in Table 71.

Table 76. PSA for HDA-2 population

Technologies	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
IV model - [redacted]			
Belimumab IV vs ST	[redacted]	[redacted]	£31,629
SC model - [redacted]			
Belimumab SC vs ST	[redacted]	[redacted]	£29,264
All model outcomes presented are discounted. *Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years*			

B.3.8.1.1 Belimumab IV vs. ST alone: PSA results

Figure 18. Scatter plot of the PSA – HDA-2 Population in the IV model

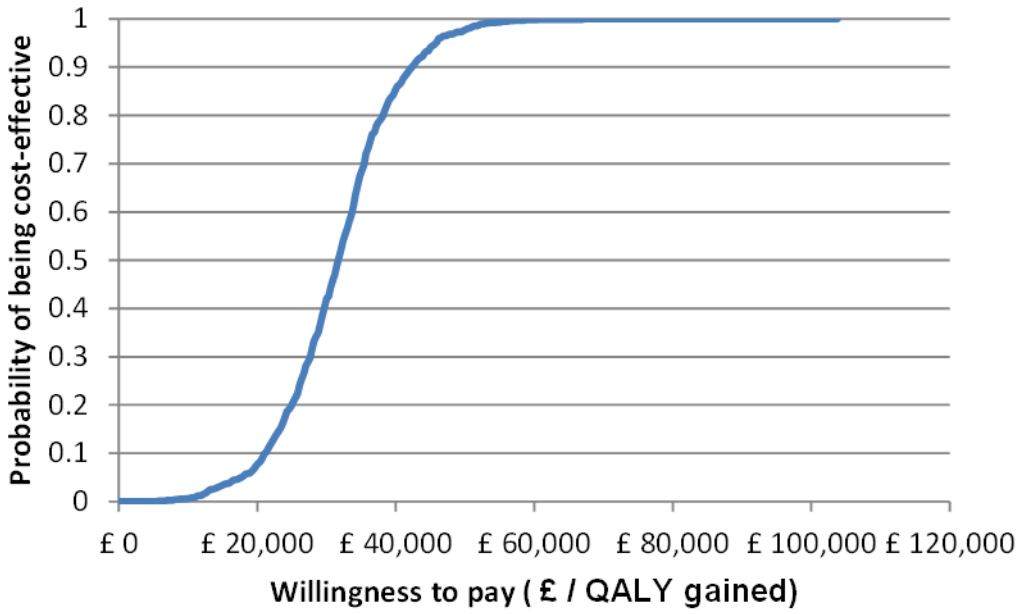


Figure 19. PSA Acceptability Curve – HDA-2 Population in the IV model

B.3.8.1.2 Belimumab SC vs. ST alone: PSA results

Figure 20. Scatter plot of the PSA – HDA-2 Population in the SC model

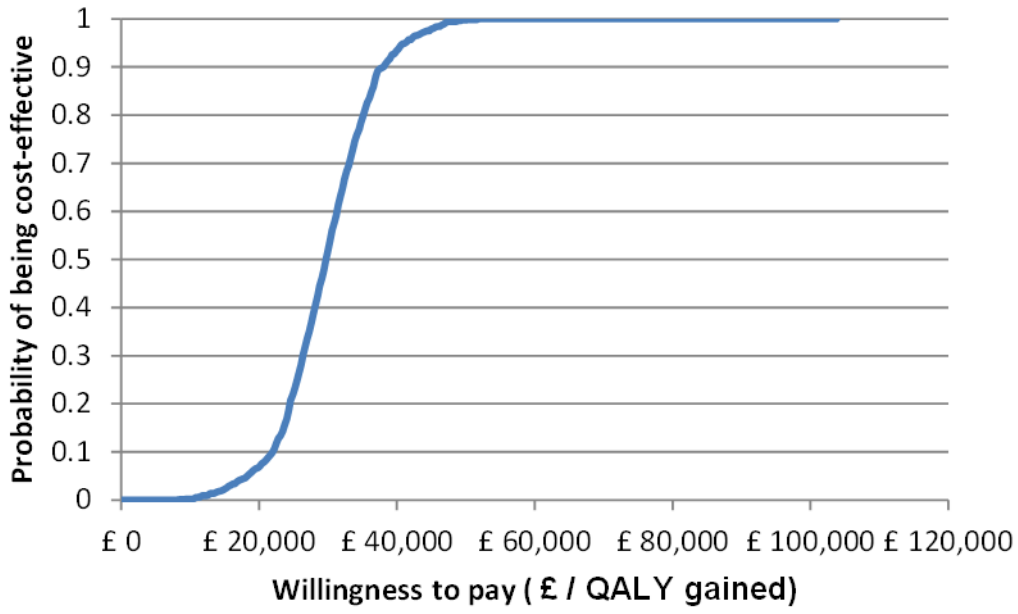


Figure 21. PSA Acceptability Curve – HDA-2 Population in the SC model

B.3.8.2 Deterministic sensitivity analysis

B.3.8.2.1 IV model results

Tornado diagrams for the ICER, incremental QALYs and incremental costs for the IV model are presented in Figure 22, Figure 23 and Figure 24 respectively.

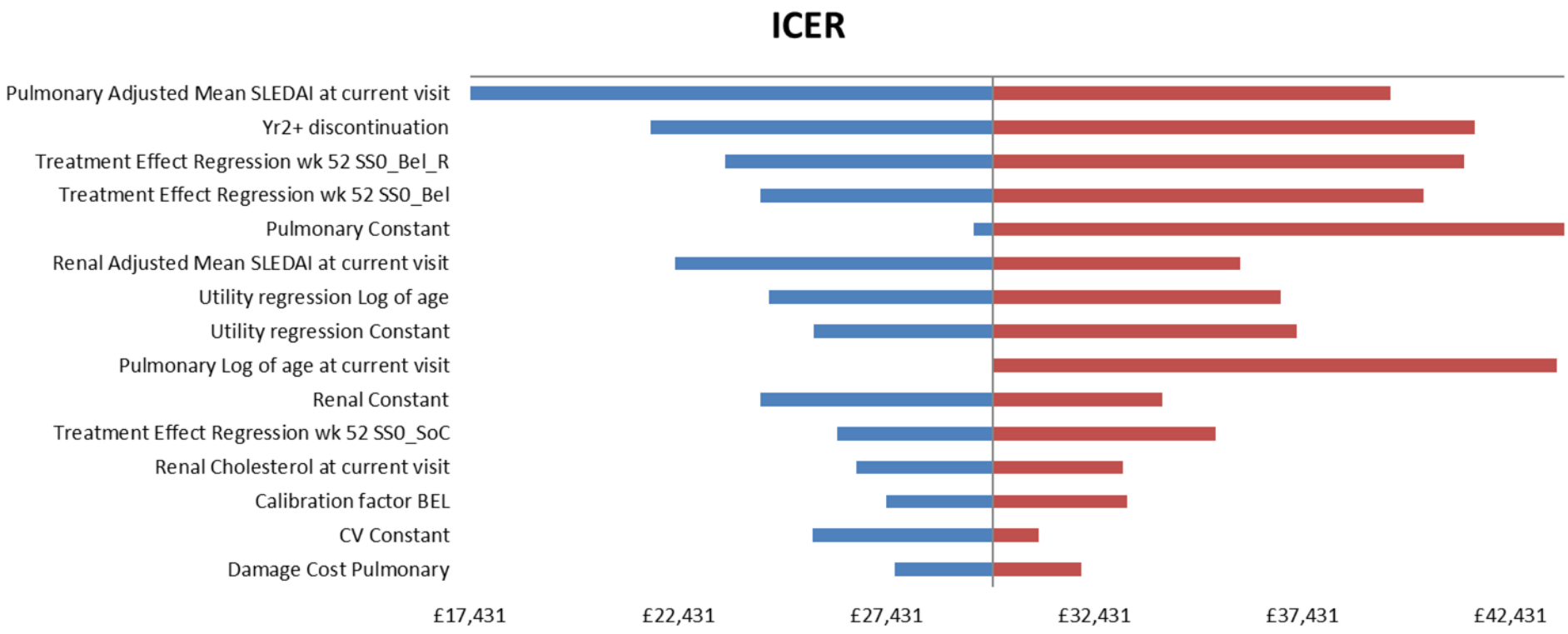


Figure 22. Tornado diagram for univariate sensitivity analyses on the ICER – HDA-2 Population in the IV model

Figure 23. Tornado diagram for univariate sensitivity analyses on the incremental costs (delta C) – HDA-2 Population in the IV model

Figure 24. Tornado diagram for univariate sensitivity analyses on the incremental QALYs (delta E) – HDA-2 Population in the IV model

B.3.8.2.2 SC model results

Tornado diagrams for the ICER, incremental QALYs and incremental costs for the SC model are presented in Figure 25, Figure 26 and Figure 27 respectively.

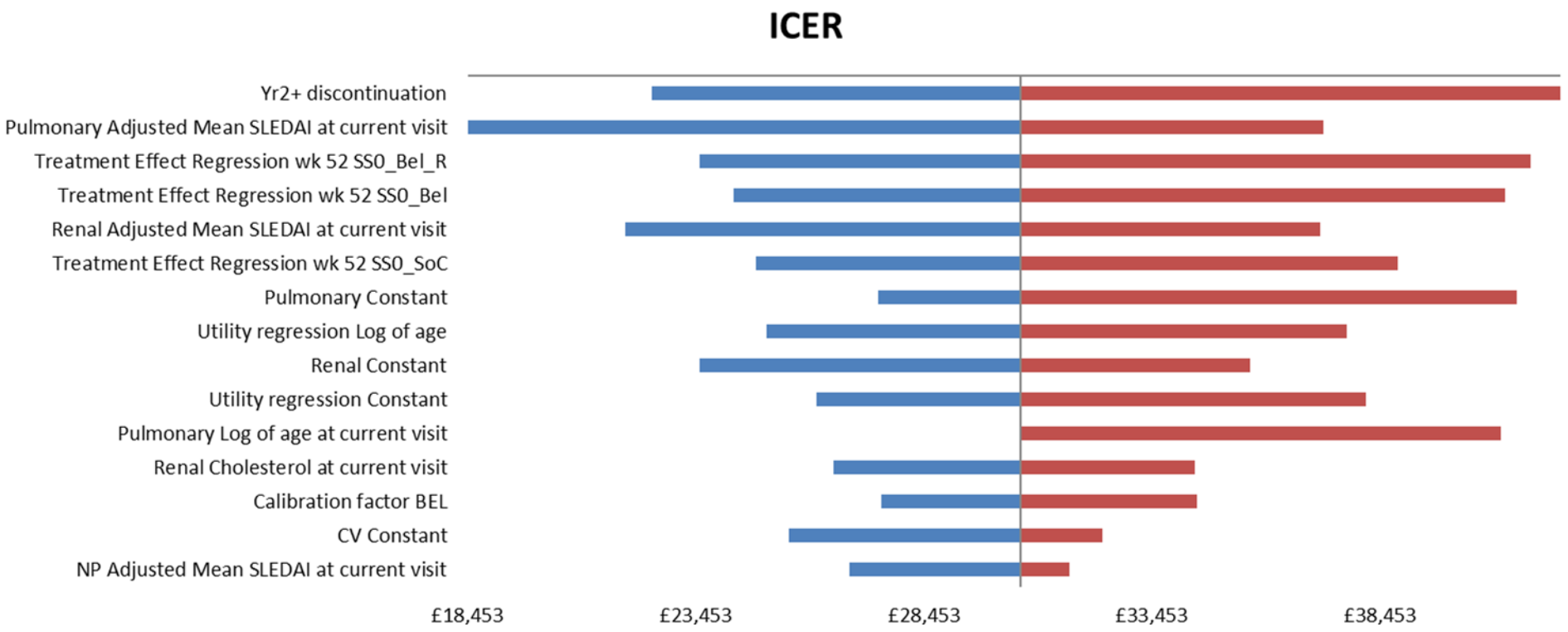


Figure 25. Tornado diagram for univariate sensitivity analyses on the ICER – HDA-2 Population in the SC model

Figure 26. Tornado diagram for univariate sensitivity analyses on the incremental costs (delta C) – HDA-2 Population in the SC model

Figure 27. Tornado diagram for univariate sensitivity analyses on the incremental QALYs (delta E) – HDA-2 Population in the SC model

B.3.8.3 Scenario analysis

A number of alternative scenario analyses among the HDA-2 population under consideration by the IV and SC models have also been conducted and these are detailed below.

1. A scenario analysis where patient weight for belimumab IV is based on clinical trial data. In our updated analysis, patient weight for belimumab IV, essential in the calculation for the dosage of belimumab IV treatment, is derived from the BILAG-BR. However, the clinical trial weights listings remain plausible in a patient population with SLE.

2. A scenario where both the treatment duration and effect of belimumab is restricted to 10 years. Patients are assumed to take belimumab for a lifetime (base-case assumption). This scenario explores the impact on a shorter treatment duration as a result of development of resistance to monoclonal antibodies[96].

3. A scenario where the calibration factors derived as part of the application of the PSM analysis are applied to both the belimumab and ST treatments in the models for 6 years. In the base-case, we take a conservative approach and apply the calibration factors derived from the PSM analysis to belimumab only for 6 years (i.e. the length of time for which there is observed data capture). However, a calibration factor was also derived for ST with the same methodology as for belimumab, and so its application should also be explored in the model.

4. A scenario where the calibration factors from the PSM analysis for belimumab is not limited to 6 years and is instead applied for a lifetime period to mirror the duration of treatment. In the base-case, a calibration factor is applied to belimumab for 6 years to reflect the length of time of data capture. However, it is reasonable to assume that if patients are taking belimumab longer than this time period, they will continue to benefit from belimumab in terms of long-term organ damage avoidance.

5. A scenario where the discount rates for costs and benefits are varied to 1.5% for benefits and 1.5% to costs. This scenario analysis is conducted to reflect the discount rate allowed by NICE where the treatment effects are both substantial in restoring health and sustained over a very long period. It also reflects the discount rate detailed in the Treasury Green Book updated in March 2018[97].
6. A scenario where the discount rates for costs and benefits are varied to 1.5% for benefits and 3.5% to costs. As above, this scenario is conducted to reflect differential discounting that may be applied if the Appraisal Committee deems that the treatment effects are both substantial in restoring health and sustained over a very long period.

Table 77 shows the results of the scenario analyses for the HDA-2 subpopulation, whilst the results for the HDA-1 subpopulation can be found in Table 5 in Appendix Q. Results for the IV model, ICERs range from £19,818 to £28,095 whilst for the SC model, ICERs range from £20,241 to £24,188.

Table 77. Results summary of the scenario analyses for HDA-2

	IV model				SC model			
Description of Scenario	**Incremental Cost Belimumab**	**Incremental LYs Belimumab**	**Incremental QALYs Belimumab**	**ICER**	**Incremental Cost Belimumab**	**Incremental LYs Belimumab**	**Incremental QALYs Belimumab**	**ICER**
Base case	[redacted]	[redacted]	[redacted]	£30,001	[redacted]	[redacted]	[redacted]	£30,566
1. Source of patient weight is BLISS trials	[redacted]	[redacted]	[redacted]	£28,095				
2. Belimumab treatment duration and effect restricted to 10 years	[redacted]	[redacted]	[redacted]	£20,485	[redacted]	[redacted]	[redacted]	£21,396
3. Calibration factors applied to both the belimumab and ST for 6 years	[redacted]	[redacted]	[redacted]	£23,419	[redacted]	[redacted]	[redacted]	£23,353
4. Calibration factors applied to belimumab only for patient lifetime	[redacted]	[redacted]	[redacted]	£24,187	[redacted]	[redacted]	[redacted]	£24,188
5. Discount rates 1.5% for both benefits and costs	[redacted]	[redacted]	[redacted]	£22,015	[redacted]	[redacted]	[redacted]	£22,556
6. Discount rates 1.5% for benefits and 3.5% for costs	[redacted]	[redacted]	[redacted]	£19,818	[redacted]	[redacted]	[redacted]	£20,241

All model outcomes presented are discounted.
ICER, incremental cost-effectiveness ratio; LY, life years; QALYs, quality-adjusted life years

B.3.9 *Subgroup analysis*

Please refer to Appendix Q for all details relating to the HDA-1 subgroup, including baseline patient characteristics, input parameters, and results.

B.3.10 *Validation*

B.3.10.1 Validation of cost-effectiveness analysis

Model convergence measures

The model convergence testing was undertaken for the original model of TA397 to minimise sampling error. The methodology is described in Section 6.8.1 in the previous submission.

Internal and external validation of long-term outcome predictions

The model accuracy of predicting long-term outcomes and mortality was tested through internal validation based on the Johns Hopkins Cohort. This is described in full in Section 6.8.1. in the previous submission.

Formula and functional error checking

Formula and functional error checking were undertaken by the supplier of the model and an independent health economic expert. Further, two independent academic health economists conducted reviews of the model suitability to address the decision problem and provided advice on how to improve the explanation of statistical methodology and assumptions used which were incorporated in the previous submission (Section 6.8.1 previous submission).

Validation undertaken for belimumab IV and SC models submitted for the current appraisal

A validation exercise by the supplier of the model (separate to the lead modeler) was undertaken on new model updates, functionality and formula for both the belimumab IV and SC models. This included, implementation of the new HDA-1 subpopulation,

updates to associated costs and utilities, application of UK-specific weight distribution and application of organ damage calibration. No important errors in the model formulae and functionality were identified.

B.3.11 *Interpretation and conclusions of economic evidence*

The economic analysis presents the cost-effectiveness of add-on belimumab (IV and SC) to standard therapy compared with standard therapy alone in SLE patients with high disease activity defined as a SS ≥ 10 with at least one of the following serological features: anti-dsDNA AND/OR low complement (HDA-2 population).

Since the commencement of the NICE appraisal of Benlysta in 2011 (TA397), there has been no new published de novo economic models undertaken in the SLE with high disease activity managed with belimumab or other human monoclonal antibody treatments. Three articles identified[87-89] are based on the same microsimulation model presented here. Therefore, the results from this cost-effectiveness analysis cannot be compared with studies, other than the previous economic analysis provided as part of TA397.

The economic evidence presented focuses on the HDA-2 population as the base-case due to reasons discussed in Section B.3.2.1. Whilst still a high disease activity population, HDA-2 population differs from the HDA-1 population upon which the current NICE guidance is based. As well as revisiting the efficacy and safety of belimumab IV including the long-term extension studies and real-world data observed through the OBSErve registry and BILAG-BR, the economic analysis presents the new belimumab SC formulation; a fixed dose weekly self-administered injection as an additional option to belimumab IV. Therefore, two models have been submitted, a belimumab IV model and a belimumab SC model which compare to standard therapy alone. As both belimumab IV and belimumab SC have an identical mechanism of action, both of these models share the same fundamental structure, differing only in how formulation specific administration costs are calculated.

The Committee, whilst acknowledging its complexities and limitations were broadly accepting of the presented model structure. For this reason, the model structure

presented here remains relatively unchanged; the microsimulation model most accurately captures the heterogeneity of SLE and the breadth of damage across multiple organ systems. There are four key changes to the model; all pertain to model parameter inputs.

Firstly, targeted updates to the costs and utilities in relation to organ damage systems were undertaken for those organ damage systems shown to contribute the most to organ damage related loss in QALYs and associated additional costs (neuropsychiatric, renal, pulmonary, cardiovascular, peripheral vascular, musculoskeletal and malignancy). For example, damage to the pulmonary organ system, and patients who experience this were associated with the worst outcomes in terms of utility score, and the highest organ damage system related management costs.

Secondly, year 2 onward assumed discontinuation rates were re-calculated based on an integrated analysis of Phase 2 and Phase 3 IV trials with over 10 years of data capture. This has reduced the assumed Year 2 onward discontinuation rate from 11.7% from the previously appraisal to [redacted].

Thirdly, the weight distribution of SLE patients currently on belimumab IV and captured in the BILAG-BR has been incorporated as a baseline characterisation in the IV model in order to provide UK specific data. This has increased the mean weight from 65.4 kg in TA397 to 70.4 kg in this submission. However, as there was a very high proportion of patients recruited into the BILAG-BR study from one centre in London this may not be totally representative of the whole UK SLE target population, so a sensitivity analysis was also provided using the weight seen in the P3 trials for our sub-populations.

Finally, and most importantly, since the last submission a propensity score matching comparative analysis of belimumab IV to a standard therapy cohort has provided insight into the comparative reduction in organ damage for SLE patients maintained on belimumab. This is a substantial step forward from the previous submission which only allowed the indirect estimate of the effect of belimumab on organ damage reduction to be captured as a result of an improved SLENAI-SLEDAI score. Through

a validation and calibration exercise the impact of belimumab on organ damage reduction has now been incorporated; an approach closer to a direct effect and as a result is now a key determinant of the presented ICERs in the scenario analyses undertaken. It should be noted that the application of the calibration factor has been made only to the belimumab arm, despite the fact that the original validation exercise appeared to demonstrate that the model simulated greater benefit for standard therapy than that observed for the PSM. The calibration factor to the belimumab arm was applied only for 6 years. It would be expected that benefits on belimumab would be experienced beyond this point in clinical practice.

In the clinical trials, belimumab significantly reduced disease activity after one year (measured by a composite endpoint which included SELENA-SLEDAI score). In the health economic model, QALYs seen for patients on belimumab were explained by lower disease activity scores associated with a decreased mortality risk and a higher quality of life. Lower disease activity was also associated with reduced risk of organ damage, resulting in fewer occurrences of cardiovascular, peripheral vascular, pulmonary, renal and skin damage.

As per TA397, the current economic analyses continue to represent a conservative approach in the cost-effective modelling of belimumab. Both the IV and SC economic models do not fully capture fatigue, flares, or the reduction in oral corticosteroids and potential for reduced OCS morbidity associated with belimumab. Therefore, this analysis may be understood to represent a conservative approach to the cost-effectiveness analysis of this technology.

Both the IV and SC formulations provide comparable cost-effectiveness. The deterministic base case in the HDA-2 population of add-on belimumab IV versus standard therapy only is associated with [redacted] additional costs, [redacted] LYs and [redacted] QALYs, and a resultant ICER of £30,001 per QALY gained. The largest drivers of incremental cost in the IV model were the costs associated with the acquisition of belimumab ([redacted]) and its administration ([redacted]), but this was offset by cost savings associated with the avoidance of [redacted] additional costs for long-term organ damage experienced by patients who received ST alone. The largest drivers for QALY gains

for patients treated with belimumab IV as compared to those treated with standard therapy alone were avoidance of damage in the cardiovascular and renal domains. For belimumab SC, the ICER is £30,566 and is associated with [redacted] additional costs, [redacted] LYs and [redacted] QALYs. The largest drivers of incremental cost in the SC model were the costs associated with the acquisition of belimumab ([redacted]), but this was offset by cost savings associated with the avoidance of [redacted] additional costs for long-term organ damage experienced by patients who received ST alone. Similar to the IV model, the largest drivers for QALY gains for patients treated with belimumab SC as compared to those treated with standard therapy alone were avoidance of damage in the cardiovascular and renal domains.

As discussed in section B.2.6.5.2, through an indirect treatment comparison, belimumab IV (10mg/kg) has demonstrated comparable effectiveness to belimumab SC (200mg). The differences seen in the reported deterministic ICER reflects the differences in response between belimumab and placebo arms of the BLISS trials and the consequential modelled reduction in SELENA-SELDAI, which is dependent on the size of the co-efficient in the regression responder analysis.

Probabilistic sensitivity analyses (PSA) results are largely consistent with the deterministic base case analyses and show that belimumab IV compared with standard therapy was associated with [redacted] additional QALYs and [redacted] additional costs, resulting in an ICER of £31,629, whereas belimumab SC compared with standard therapy was associated with [redacted] additional QALYs and [redacted] additional costs, resulting in an ICER of £29,264.

Results from the one-way deterministic sensitivity analyses showed that the parameters that were most sensitive to the ICER were the year two natural discontinuation rate and the regression coefficient associated with the ‘pulmonary adjusted mean SELDAI at current visit’. This is unsurprising, as patients who stop receiving benefit from belimumab and discontinue would stop incurring technology related costs. The pulmonary organ damage domain is associated with the lowest utility multiplier of all the organ damage systems and the highest ongoing cost. All

scenario analyses for the IV and SC models were associated with ICERs below £30,000 per QALY gained.

Analyses run for the HDA-2 subgroup were also performed for the HDA-1 subgroup for both belimumab formulations. Results across the base case, PSA, and scenario analysis showed that for the HDA-1 subgroup for both IV and SC models, all ICERS were below £30,000 per QALY gained.

The main strengths of this evaluation comprise:

- Short-term clinical efficacy was based on two well-designed RCTs for the IV formulation and one well-designed RCT for the SC formulation. Where possible, longer term evidence collection has been utilised.
- As per TA397, the natural history model for SLE was developed following an extensive analysis of the JH cohort, a large, long-term, observational dataset. The model is therefore able to accurately predict the long-term course of the disease and captures the heterogeneity and complexity of SLE, where a patient's history plays an important role in the future disease course. It enabled detailed examination of the relationships between various risk factors, organ damage and mortality.
- As per TA397, predictions of organ damage events over time and mortality were validated with a second longitudinal SLE database (Toronto Lupus Cohort) and showed good predictive accuracy for most disease organ systems and mortality.
- The results of the PSM analysis reporting the 5-year reduction in SDI was validated in the economic analysis from 1.5 years onwards. The subsequent PSM derived calibration factors that were applied to both the IV and SC models enable the incorporation of the longer-term benefit of belimumab for which we had previously been unable to demonstrate.
- A conservative assumption was made with respect to long-term effect of belimumab on disease activity levels (SS score). The difference in SS score may in fact increase over time whereas the assumption used in the analyses is that the difference observed at 52 weeks remains constant over time. As a result, beneficial effect on HRQoL related to long-term outcomes may be

underestimated in the model compared with what may be observed over the long-term in UK clinical practice, thus the ICER may be conservative.

- The time lag since the commencement of the previous submission and appraisal has enabled us to include all relevant long-term data and as such, for example, better estimate the likely long-term discontinuation rate.
- Comprehensive sensitivity and scenario analyses have been performed using all available data.

The main weaknesses of this evaluation comprise:

- The primary analysis of the PSM was conducted on the total population of the BLISS-76 US open label extension study, rather than the HDA subgroup under consideration by our economic analysis. This means that the model validation and subsequent derivation of calibration factors are assumed unchanged from an ITT population. We believe this to be a conservative approach. Belimumab has been shown to demonstrate a greater benefit in the two high disease activity subpopulations (HDA-2 and HDA-1).
- Disease flares, a common occurrence due to the relapsing and remitting nature of SLE, were not simulated in the model. Measures of flare were considered for inclusion in the disease activity model however the SELENA-SLEDAI Flare Index was not collected in the JH database. An alternative measure of flare could have been used however this may have caused problems due to the correlation between flare and SLEDAI score. As the model uses the adjusted mean SLEDAI (AMS), disease activity is 'smoothed' over time, and a flare or relapse of activity cannot be shown. A decrease in frequency of flares due to belimumab will however also decrease the AMS over the treatment period. Therefore, although flares are not directly simulated in the model, some effect of decreasing flares is incorporated. This "smoothed" effect may lead to underestimating the benefit of belimumab.
- During the internal validation exercise for TA397, it was seen that the predicted incidence of mortality was slightly underestimated. The reason for the lower incidence of death may arise because the organ systems were modelled

independently. Solutions to this problem were explored, however the complexity of statistical modelling required to account for this is considerable and would not have been possible within the timelines of this project. This approach is in line with what was presented as part of TA397 and may continue to lead to a conservative estimate of cost-effectiveness.

- The mortality model does not describe the rate of mortality for patients aged >65 years. Consequently, an adjustment is made in the cost-effectiveness model to allow the risk of mortality to increase in line with the general population at ages not represented in the JH data.
- Costs related to disease activity were modelled independently of the costs associated with organ damage. This approach could lead to double counting of some costs. It is unlikely that this will have an impact on the overall presented results as disease activity costs are minor. In addition, the cost of ST was not included in this model.
- A targeted literature review rather than a full systematic literature review was performed to update costs and utilities for the current submission. Where new costs were not located due to either not being searched or the targeted literature review not yielding any results, costs were updated using inflation indices. It may be argued that this approach underestimates costs associated with treating active disease and those associated with long-term organ damage.
- The EQ-5D is unlikely to be the most sensitive generic instrument to detect all aspects of SLE on patient HRQoL. For example, fatigue is one of the most frequently cited and most bothersome symptoms for SLE patients; the FACIT-fatigue instrument (Section 9.19, Appendix 19 of the previous submission) was collected during the BLISS-52 and BLISS-76 trials at 4, 8, 12, 24 and 52 week time-points. The FACIT-fatigue instrument was also used in in the BLISS-SC trial at week 52 and by visit. The results presented in Figures 5.14 to 5.17 in Section 5.5 of the previous submission demonstrate a significant improvement in fatigue scores with belimumab which was sustained over the trial period. This symptom will have a considerable impact on HRQoL. As the EQ-5D is not sufficiently sensitive to detect the impact of this symptom, the overall utility benefit with belimumab could be underestimated in the cost-effectiveness models.

- The updated model structure has still been unable to capture the potential benefit of reducing the exposure of patients to the cumulative effects of steroids. From the pooled analysis across the real-world study series, OBSErve, belimumab was shown to be steroid-sparing; most patients receiving steroids at the initiation of treatment with belimumab decreased their steroid dose after 6 months of treatment. Furthermore, of those receiving a dose >7.5 mg/day (78.4% of all patients receiving steroids at baseline), 54.5% had a dose reduction to <7.5 mg/day at Month 6. Further there is evidence to suggest that the healthcare resource utilisation of Lupus patients increases with higher doses of corticosteroids owing to increased emergency room and hospitalisation costs[98, 99] and costs associated with managing corticosteroid-related AEs. There is also evidence to suggest that SLE patients incur higher societal costs considered associated and most likely driven by corticosteroid use, fatigue and disease activity[100].
- Some tangible aspects of the disease considered important have not been included in the cost-effectiveness assessment. Particularly for the more severe SLE patients, their inability to work and their reliance on carers, carries both a financial burden and will impact significantly on their mental wellbeing.
- Low recruitment numbers to the BILAG-BR has resulted in limited data that can be incorporated into the current economic analyses.

The re-review by NICE has afforded GSK the opportunity to present new evidence since 2011, offer an additional formulation to the IV presentation to support equity of treatment access (especially important in COVID times) and present the cost-effectiveness of add-on belimumab in a more clinically relevant high disease activity population; HDA-2. Together these findings demonstrate that belimumab continues to offer the NHS a cost-effective use of resources and more importantly provides patients with this debilitating incurable condition, who continue to experience disease activity despite standard therapy, with access to a licensed therapy to better control their symptoms and reduce the rate of irreversible organ damage accrual.

B.4 References

1. Kalunian K, Merrill JT. New directions in the treatment of systemic lupus erythematosus. Curr Med Res Opin. 2009;25(6):1501-14.
2. Baker KP, Edwards BM, Main SH, Choi GH, Wager RE, Halpern WG, et al. Generation and characterization of LymphoStat-B, a human monoclonal antibody that antagonizes the bioactivities of B lymphocyte stimulator. Arthritis and rheumatism. 2003;48(11):3253-65.
3. Cancro MP, D'Cruz DP, Khamashta MA. The role of B lymphocyte stimulator (BLyS) in systemic lupus erythematosus. The Journal of clinical investigation. 2009;119(5):1066-73.
4. Zhang J, Roschke V, Baker KP, Wang Z, Alarcon GS, Fessler BJ, et al. Cutting edge: a role for B lymphocyte stimulator in systemic lupus erythematosus. Journal of immunology (Baltimore, Md : 1950). 2001;166(1):6-10.
5. Cheema GS, Roschke V, Hilbert DM, Stohl W. Elevated serum B lymphocyte stimulator levels in patients with systemic immune-based rheumatic diseases. Arthritis and rheumatism. 2001;44(6):1313-9.
6. GlaxoSmithKline. Benlysta (belimumab). Summary of Product Characteristics 30 September 2020 [Available from: https://www.ema.europa.eu/en/documents/product-information/benlysta-epar-product-information_en.pdf.
7. Manson JJ, Rahman A. Systemic lupus erythematosus. Orphanet J Rare Dis. 2006;1:6.
8. Rees F, Doherty M, Grainge M, Davenport G, Lanyon P, Zhang W. The incidence and prevalence of systemic lupus erythematosus in the UK, 1999–2012. Annals of the Rheumatic Diseases. 2016;75(1):136-41.
9. Johnson AE, Gordon C, Palmer RG, Bacon PA. The prevalence and incidence of systemic lupus erythematosus in Birmingham, England. Relationship to ethnicity and country of birth. Arthritis and rheumatism. 1995;38(4):551-8.
10. Somers EC, Thomas SL, Smeeth L, Schoonen WM, Hall AJ. Incidence of systemic lupus erythematosus in the United Kingdom, 1990-1999. Arthritis and rheumatism. 2007;57(4):612-8.
11. Rees F, Doherty M, Grainge MJ, Lanyon P, Davenport G, Zhang W. Mortality in systemic lupus erythematosus in the United Kingdom 1999–2012. Rheumatology. 2016;55(5):854-60.
12. Rahman A, Isenberg DA. Systemic Lupus Erythematosus. New England Journal of Medicine. 2008;358(9):929-39.
13. Morgan C, Bland AR, Maker C, Dunnage J, Bruce IN. Individuals living with lupus: findings from the LUPUS UK Members Survey 2014. Lupus. 2018;27(4):681-7.
14. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis and rheumatism. 1997;40(9):1725.
15. Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis and rheumatism. 1982;25(11):1271-7.

16. Lam NC, Ghetu MV, Bieniek ML. Systemic Lupus Erythematosus: Primary Care Approach to Diagnosis and Management. Am Fam Physician. 2016;94(4):284-94.
17. Petri M, Orbai A-M, Alarcón GS, Gordon C, Merrill JT, Fortin PR, et al. Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. Arthritis and rheumatism. 2012;64(8):2677-86.
18. Aringer M. EULAR/ACR classification criteria for SLE. Semin Arthritis Rheum. 2019;49(3s):S14-s7.
19. Fava A, Petri M. Systemic lupus erythematosus: Diagnosis and clinical management. J Autoimmun. 2019;96:1-13.
20. Kaul A, Gordon C, Crow MK, Touma Z, Urowitz MB, van Vollenhoven R, et al. Systemic lupus erythematosus. Nature Reviews Disease Primers. 2016;2(1):16039.
21. Farinha F, Freitas F, Águeda A, Cunha I, Barcelos A. Concerns of patients with systemic lupus erythematosus and adherence to therapy - a qualitative study. Patient Prefer Adherence. 2017;11:1213-9.
22. Zonana-Nacach A, Roseman JM, McGwin G, Jr., Friedman AW, Baethge BA, Reveille JD, et al. Systemic lupus erythematosus in three ethnic groups. VI: Factors associated with fatigue within 5 years of criteria diagnosis. LUMINA Study Group. LUpus in MInority populations: NAture vs Nurture. Lupus. 2000;9(2):101-9.
23. Lau CS, Mak A. The socioeconomic burden of SLE. Nat Rev Rheumatol. 2009;5(7):400-4.
24. Yelin E, Trupin L, Katz P, Criswell L, Yazdany J, Gillis J, et al. Work dynamics among persons with systemic lupus erythematosus. Arthritis and rheumatism. 2007;57(1):56-63.
25. Ugarte-Gil MF, Acevedo-Vásquez E, Alarcón GS, Pastor-Asurza CA, Alfaro-Lozano JL, Cucho-Venegas JM, et al. The number of flares patients experience impacts on damage accrual in systemic lupus erythematosus: data from a multiethnic Latin American cohort. Ann Rheum Dis. 2015;74(6):1019-23.
26. Gladman DD, Urowitz MB, Rahman P, Ibanez D, Tam LS. Accrual of organ damage over time in patients with systemic lupus erythematosus. J Rheumatol. 2003;30(9):1955-9.
27. Lopez R, Davidson JE, Beeby MD, Egger PJ, Isenberg DA. Lupus disease activity and the risk of subsequent organ damage and mortality in a large lupus cohort. Rheumatology (Oxford, England). 2012;51(3):491-8.
28. Gladman D, Ginzler E, Goldsmith C, Fortin P, Liang M, Urowitz M, et al. The development and initial validation of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology damage index for systemic lupus erythematosus. Arthritis and rheumatism. 1996;39(3):363-9.
29. American College of Rheumatology Ad Hoc Committee on Systemic Lupus Erythematosus. Guidelines for referral and management of systemic lupus erythematosus in adults. . Arthritis and rheumatism. 1999;42(9):1785-96.
30. Swaak AJ, van den Brink HG, Smeenk RJ, Manger K, Kalden JR, Tosi S, et al. Systemic lupus erythematosus: clinical features in patients with a disease duration of over 10 years, first evaluation. Rheumatology (Oxford, England). 1999;38(10):953-8.
31. Cervera R, Khamashta MA, Font J, Sebastiani GD, Gil A, Lavilla P, et al. Morbidity and mortality in systemic lupus erythematosus during a 10-year period: a

comparison of early and late manifestations in a cohort of 1,000 patients. Medicine (Baltimore). 2003;82(5):299-308.
32. Goldblatt F, Chambers S, Rahman A, Isenberg DA. Serious infections in British patients with systemic lupus erythematosus: hospitalisations and mortality. Lupus. 2009;18(8):682-9.
33. Campbell Jr R, Cooper GS, Gilkeson GS. Two aspects of the clinical and humanistic burden of systemic lupus erythematosus: Mortality risk and quality of life early in the course of disease. Arthritis Care & Research. 2008;59(4):458-64.
34. Roman MJ, Shanker BA, Davis A, Lockshin MD, Sammaritano L, Simantov R, et al. Prevalence and correlates of accelerated atherosclerosis in systemic lupus erythematosus. N Engl J Med. 2003;349(25):2399-406.
35. Urowitz MB, Gladman DD, Tom BD, Ibanez D, Farewell VT. Changing patterns in mortality and disease outcomes for patients with systemic lupus erythematosus. J Rheumatol. 2008;35(11):2152-8.
36. Bernatsky S, Boivin JF, Joseph L, Rajan R, Zoma A, Manzi S, et al. An international cohort study of cancer in systemic lupus erythematosus. Arthritis and rheumatism. 2005;52(5):1481-90.
37. Parikh-Patel A, White RH, Allen M, Cress R. Cancer risk in a cohort of patients with systemic lupus erythematosus (SLE) in California. Cancer Causes Control. 2008;19(8):887-94.
38. Bernatsky S, Ramsey-Goldman R, Isenberg D, Rahman A, Dooley MA, Sibley J, et al. Hodgkin's lymphoma in systemic lupus erythematosus. Rheumatology (Oxford, England). 2007;46(5):830-2.
39. Bernatsky S, Boivin JF, Joseph L, Manzi S, Ginzler E, Gladman DD, et al. Mortality in systemic lupus erythematosus. Arthritis & Rheumatism. 2006;54(8):2550-7.
40. Alarcón GS, McGwin JG, Bastian HM, Roseman J, Lisse J, Fessler BJ, et al. Systemic lupus erythematosus in three ethnic groups. VIII. Predictors of early mortality in the LUMINA cohort. Arthritis Care & Research. 2001;45(2):191-202.
41. Molad Y, Borkowski T, Monselise A, Ben-Haroush A, Sulkes J, Hod M, et al. Maternal and fetal outcome of lupus pregnancy: a prospective study of 29 pregnancies. Lupus. 2005;14(2):145-51.
42. Fanouriakis A, Kostopoulou M, Alunno A, Aringer M, Bajema I, Boletis JN, et al. 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. Annals of the Rheumatic Diseases. 2019;78(6):736-45.
43. Gordon C, Amissah-Arthur M-B, Gayed M, Brown S, Bruce IN, D'Cruz D, et al. The British Society for Rheumatology guideline for the management of systemic lupus erythematosus in adults. Rheumatology. 2017;57(1):e1-e45.
44. NHS England. Clinical Commissioning Policy Rituximab for refractory Systemic Lupus Erythematosus (SLE) in adults and post-pubescent children [200402P] 2020 [Available from: https://www.england.nhs.uk/wp-content/uploads/2020/07/Rituximab-for-refractory-Systemic-Lupus-Erythematosus-in-adults-and-post-pubescent-children-2.pdf.
45. GlaxoSmithKline. Clinical Study Report: A Phase 3, Multi-Center, Randomized, Double-Blind, Placebo-Controlled, 52-Week Study to Evaluate the Efficacy and Safety of Belimumab (HGS1006) Administered Subcutaneously (SC) to Subjects with Systemic Lupus Erythematosus (SLE) – Double-Blind Endpoint Analysis. Data on file. 8 March 2016.

46. GlaxoSmithKline. A Phase 3, Multi-Center, Randomized, Double-Blind, Placebo-Controlled, 52-Week Study to Evaluate the Efficacy and Safety of Belimumab (HGS1006, LymphoStat-B™), a Fully Human Monoclonal Anti-BLyS Antibody, in Subjects with Systemic Lupus Erythematosus (SLE). Data on file. 21 January 2010.
47. GlaxoSmithKline. A Phase 3, Multi-Center, Randomized, Double-Blind, Placebo-Controlled, 76-Week Study to Evaluate the Efficacy and Safety of Belimumab (HGS1006, LymphoStat-B™), a Fully Human Monoclonal Anti-BLyS Antibody, in Subjects with Systemic Lupus Erythematosus (SLE). Data on file. 28 September 2010.
48. Urowitz MB, Ohsfeldt RL, Wielage RC, Kelton KA, Asukai Y, Ramachandran S. Organ damage in patients treated with belimumab versus standard of care: a propensity score-matched comparative analysis. Annals of the Rheumatic Diseases. 2019;78(3):372-9.
49. GlaxoSmithKline. A Phase 3, Multi-Center, Randomized, Double-Blind, Placebo Controlled, 52-Week Study to Evaluate the Efficacy and Safety of Belimumab (HGS1006) Administered Subcutaneously (SC) to Subjects with Systemic Lupus Erythematosus (SLE) -Open-label phase - End of Study report. Data on File. 26 September 2016.
50. GlaxoSmithKline. A Multi-Center, Continuation Trial of Belimumab (HGS1006, LymphoStat-B), a Fully Human Monoclonal Anti-BLyS Antibody, in Subjects with Systemic Lupus Erythematosus (SLE) who Completed the Phase 3 Protocol HGS1006-C1056 or HGS1006-C1057. Data on file. 31 July 2017.
51. GlaxoSmithKline. A Multi-Center, Continuation Trial of Belimumab (HGS1006, LymphoStat-B), a Fully Human Monoclonal Anti-BLyS Antibody, in Subjects with Systemic Lupus Erythematosus (SLE) Who Completed the Phase 3 Protocol HGS1006-C1056 in the United States. Data on file. 2 December 2015.
52. Bruce IN, on behalf of the British Isles Lupus Assessment Group. BILAG Biologics Prospective Cohort: The Use of Novel Biological Therapies in the Treatment of Systemic Lupus Erythematosus (SLE). BILAG BR Protocol Version 1.4 07/03/2011. Data on file.
53. GlaxoSmithKline. Reporting and Analysis Plan (RAP) for the UK Benlysta▼ Prospective, Observational BILAG British Registry (BILAG-BR) Sub-Study. Data on file. 6 April 2020.
54. Collins C, Cortés-Hernández J, García M, Kempis Jv, Schwarting A, Touma Z, et al. FRI0189 Pooled analysis of the real-world effectiveness of belimumab in treatment of systemic lupus erythematosus using multi-country data from the observe studies. Annals of the Rheumatic Diseases. 2019;78(Suppl 2):771-2.
55. Collins CE, Dall'Era M, Kan H, Macahilig C, Molta C, Koscielny V, et al. Response to belimumab among patients with systemic lupus erythematosus in clinical practice settings: 24-month results from the OBSErve study in the USA. Lupus science & medicine. 2016;3(1):e000118.
56. Babini A, García MA, Barreira JC, Pons-Estel B, Iglesias M, Streger G. Evaluation of Use of Belimumab in Clinical Practice Settings: Results in Argentina. Abstract number: 775. 2016 ACR/ARHP Annual Meeting; Washington, DC, USNovember 11-16, 2016.
57. Schwarting A, Schroeder JO, Alexander T, Schmalzing M, Fiehn C, Specker C, et al. First Real-World Insights into Belimumab Use and Outcomes in Routine

Clinical Care of Systemic Lupus Erythematosus in Germany: Results from the OBSErve Germany Study. Rheumatology and therapy. 2016;3(2):271-90.
58. Cortes J, Marras C, Andreu JL, Calvo-Alen J, Garcia-Aparicio AM, Diez Alvarez E, et al. Evolution of Patients with Systemic Lupus Erythematous Treated with Belimumab in Clinical Practice Settings. Abstract number: 668. 2014 ACR/ARHP Annual Meeting; Boston, MA, USNovember 14-19, 2014.
59. Touma Z, Sayani A, Pineau CA, Fortin I, Matsos M, Ecker GA, et al. Belimumab use, clinical outcomes and glucocorticoid reduction in patients with systemic lupus erythematosus receiving belimumab in clinical practice settings: results from the OBSErve Canada Study. Rheumatology international. 2017;37(6):865-73.
60. von Kempis J, Duetsch S, Reuschling N, Villiger R, Villiger PM, Vallelian F, et al. Clinical outcomes in patients with systemic lupus erythematosus treated with belimumab in clinical practice settings: a retrospective analysis of results from the OBSErve study in Switzerland. Swiss medical weekly. 2019;149:w20022.
61. Brunner HI A-MC, Viola DO, et al. . Safety and efficacy of intravenous belimumab in children with systemic lupus erythematosus: results from a randomised, placebo-controlled trial. Annals of the Rheumatic Diseases. 2020.
62. Wallace DJ, Stohl W, Furie RA, Lisse JR, McKay JD, Merrill JT, et al. A phase II, randomized, double-blind, placebo-controlled, dose-ranging study of belimumab in patients with active systemic lupus erythematosus. Arthritis and rheumatism. 2009;61(9):1168-78.
63. Wallace DJ, Ginzler EM, Merrill JT, Furie RA, Stohl W, Chatham WW, et al. Safety and Efficacy of Belimumab Plus Standard Therapy for Up to Thirteen Years in Patients With Systemic Lupus Erythematosus. Arthritis & rheumatology (Hoboken, NJ). 2019;71(7):1125-34.
64. Sheikh S, Scheinberg M, Cheng-Chung Wei J, Tegzova D, Stohl W, Mucenic T, et al. LB0012 Headline results for a phase 4, 52-week, randomised, double-blind, placebo-controlled study to assess adverse events of special interest (aesi) in adults with active, autoantibody-positive systemic lupus erythematosus (sle) receiving belimumab. Annals of the Rheumatic Diseases. 2019;78(Suppl 2):266-.
65. Bae S, Dimelow R, Ji B, Kurrasch R, Muzaffar S, Punwaney R, et al. Results of the Open-label, Non-randomized 52-Week Study to Evaluate Treatment Holidays and Rebound Phenomenon After Treatment with Belimumab in Patients with SLE. Arthritis & rheumatology (Hoboken, NJ). 2019;71((suppl 10)).
66. Sergio Guedes Barbosa L, Saxena A, Ginzler E, Furie R, D'Cruz D, Burriss S, et al. Efficacy of Belimumab in Patients of Black Race with Systemic Lupus Erythematosus and High Disease Activity or Renal Manifestations. Arthritis & rheumatology (Hoboken, NJ). 2019;71(suppl 10).
67. GlaxoSmithKline. Clinical Protocol HGS1006-C1115. Protocol Amendment: 05. EudraCT Number: 2011-003814-18. A Phase 3, Multi-Center, Randomized, Double-Blind, Placebo-Controlled, 52-Week Study to Evaluate the Efficacy and Safety of Belimumab (HGS1006) Administered Subcutaneously (SC) to Subjects with Systemic Lupus Erythematosus (SLE). 12 June 2014.
68. GlaxoSmithKline. Pooled analysis of the real-world effectiveness of belimumab using multi-country data from the OBSErve studies. 19 March 2020.

69. GlaxoSmithKline. An Open-Label Single-Arm Study to Evaluate the Reliability of an Autoinjector that Administers Belimumab Subcutaneously in Subjects with Systemic Lupus Erythematosus (SLE). 23 October 2015.
70. Ramachandran S, Parks D, Kurtinecz M, Roth DA, Alfonso-Cristancho R. An indirect comparison of intravenous and subcutaneous belimumab efficacy in patients with SLE and high disease activity. Journal of Comparative Effectiveness Research. 2018;7(6):581-93.
71. GlaxoSmithKline. Reporting and Analysis Plan (RAP) for BELl 12341, A Phase 3, Multi-Center, Randomized, Double-Blind, Placebo-Controlled, 52-Week Study to Evaluate the Efficacy and Safety of Belimumab (HOS 1006) Administered Subcutaneously (SC) to Subjects with Systemic Lupus Erythematosus (SLE). Data on file. 16 April 2015.
72. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health. 1998;52(6):377-84.
73. Stohl W, Schwarting A, Okada M, Scheinberg M, Doria A, Hammer AE, et al. Efficacy and Safety of Subcutaneous Belimumab in Systemic Lupus Erythematosus: A Fifty-Two-Week Randomized, Double-Blind, Placebo-Controlled Study. Arthritis & rheumatology (Hoboken, NJ). 2017;69(5):1016-27.
74. Babini A, Cappuccio AM, Caprarulo C, Casado G, Eimon A, Figueredo H, et al. Evaluation of belimumab treatment in patients with systemic lupus erythematosus in a clinical practice setting: Results from a 24-month OBSErve study in Argentina. Lupus. 2020:961203320947814.
75. Collins CE, Cortes-Hernández J, Garcia MA, von Kempis J, Schwarting A, Touma Z, et al. Real-World Effectiveness of Belimumab in the Treatment of Systemic Lupus Erythematosus: Pooled Analysis of Multi-Country Data from the OBSErve Studies. Rheumatology and therapy. 2020.
76. GlaxoSmithKline. Follow up Questionnaires and Interviews Assessing Patient Experience with Belimumab and Delivery Device Preferences in Systemic Lupus Erythematosus (SLE). 7 October 2015.
77. Dashiell-Aje E, Harding G, Pascoe K, DeVries J, Berry P, Ramachandran S. Patient Evaluation of Satisfaction and Outcomes with an Autoinjector for Self-Administration of Subcutaneous Belimumab in Patients with Systemic Lupus Erythematosus. The Patient - Patient-Centered Outcomes Research. 2018;11(1):119-29.
78. GlaxoSmithKline. An Open-label, Non-randomized, 52-Week Study to Evaluate Treatment Holidays and Rebound Phenomenon After Treatment With Belimumab 10 mg/kg in Systemic Lupus Erythematosus Subjects. 31 July 2019.
79. Doria A, Stohl W, Schwarting A, Okada M, Scheinberg M, van Vollenhoven R, et al. Efficacy and Safety of Subcutaneous Belimumab in Anti-Double-Stranded DNA-Positive, Hypocomplementemic Patients With Systemic Lupus Erythematosus. Arthritis & rheumatology (Hoboken, NJ). 2018;70(8):1256-64.
80. van Vollenhoven RF, Petri MA, Cervera R, Roth DA, Ji BN, Kleoudis CS, et al. Belimumab in the treatment of systemic lupus erythematosus: high disease activity predictors of response. Annals of the rheumatic diseases. 2012;71(8):1343-9.
81. Petri M, Singh S, Tesfasyone H, Malik A. Prevalence of flare and influence of demographic and serologic factors on flare risk in systemic lupus erythematosus: a prospective study. J Rheumatol. 2009;36(11):2476-80.

82. McElhone K, Abbott J, Sutton C, Mullen M, Lanyon P, Rahman A, et al. Sensitivity to Change and Minimal Important Differences of the LupusQoL in Patients With Systemic Lupus Erythematosus. Arthritis Care Res (Hoboken). 2016;68(10):1505-13.
83. Merrill JT, Neuwelt CM, Wallace DJ, Shanahan JC, Latinis KM, Oates JC, et al. Efficacy and safety of rituximab in moderately-to-severely active systemic lupus erythematosus: the randomized, double-blind, phase II/III systemic lupus erythematosus evaluation of rituximab trial. Arthritis and rheumatism. 2010;62(1):222-33.
84. Touma Z, Gladman DD. Current and future therapies for SLE: obstacles and recommendations for the development of novel treatments. Lupus science & medicine. 2017;4(1):e000239-e.
85. Vukelic M, Li Y, Kyttaris VC. Novel Treatments in Lupus. Front Immunol. 2018;9:2658-.
86. National Health Service. Clinical guide for the management of Rheumatology patients during the coronavirus pandemic. Version 2 8 April 2020 [Available from: https://www.england.nhs.uk/coronavirus/wp-content/uploads/sites/52/2020/03/clinical-guide-rheumatology-patients-v2-08-april-2020.pdf.
87. Specchia ML, de Waure C, Gualano MR, Doria A, Turchetti G, Pippo L, et al. Health technology assessment of belimumab: a new monoclonal antibody for the treatment of systemic lupus erythematosus. BioMed research international. 2014;2014:704207.
88. Pierotti F, Palla I, Treur M, Pippo L, Turchetti G. Assessment of the Economic Impact of Belimumab for the Treatment of Systemic Lupus Erythematosus in the Italian Setting: A Cost-Effectiveness Analysis. PloS one. 2015;10(10):e0140843.
89. Canadian Agency for Drugs and Technologies in Health. CDEC FINAL RECOMMENDATION. BELIMUMAB (Benlysta – GlaxoSmithKline Inc.) Indication: Systemic Lupus Erythematosus April 25, 2012 [Available from: https://www.cadth.ca/media/cdr/complete/cdr_complete_Benlysta_April-27-12_e.pdf.
90. Ibañez D, Urowitz MB, Gladman DD. Summarizing disease features over time: I. Adjusted mean SLEDAI derivation and application to an index of disease activity in lupus. The Journal of Rheumatology. 2003;30(9):1977.
91. Medical Decision Modeling Inc. (MDM). Clinical Study Report HO-16-16611; eTrack: 206347. A propensity score-matched study of systemic lupus erythematosus related organ damage in the BLISS long term extension trials (BEL112233 and BEL112234) and the Toronto Lupus Cohort. Data on file. January 10, 2018.
92. Dolan P. Modeling valuations for EuroQol health states. Med Care. 1997;35(11):1095-108.
93. Office for National Statistics. National life tables – life expectancy in the UK: 2017 to 2019 [Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/nationallifetablesunitedkingdom/latest
94. Curtis L, Burns A. Unit Costs of Health and Social Care 2019, Personal Social Services Research Unit, University of Kent, Canterbury 2019 [
95. Briggs A, Sculpher M, Claxton K. Decision modelling for health economic evaluation. Oxford: Oxford University Press; 2006.

96. Torka P, Barth M, Ferdman R, Hernandez-Ilizaliturri FJ. Mechanisms of Resistance to Monoclonal Antibodies (mAbs) in Lymphoid Malignancies. Curr Hematol Malig Rep. 2019;14(5):426-38.
97. HM Treasury. The green book: Central government guidance on appraisal and evaluation London: HM Treasury; 2018 [Available from: www.gov.uk/government/publications.
98. Kabadi S, Yeaw J, Bacani AK, Tafesse E, Bos K, Karkare S, et al. Healthcare resource utilization and costs associated with long-term corticosteroid exposure in patients with systemic lupus erythematosus. Lupus. 2018;27(11):1799-809.
99. Petri M, Bechtel B, Dennis G, Shah M, McLaughlin T, Kan H, et al. Burden of corticosteroid use in patients with systemic lupus erythematosus: results from a Delphi panel. Lupus. 2014;23(10):1006-13.
100. Bexelius C, Wachtmeister K, Skare P, Jönsson L, Vollenhoven R. Drivers of cost and health-related quality of life in patients with systemic lupus erythematosus (SLE): a Swedish nationwide study based on patient reports. Lupus. 2013;22(8):793-801.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (Review of TA397) [ID1591]

Clarification questions

December 2020

File name	Version	Contains confidential information	Date
ID1591 Belimumab clarification letter to PM	**V4.0**	**Yes**	**10th February 2021**

Tables

Figures

Section A: Clarification on effectiveness data

Literature Searching

A1. The search methods state that searches were limited to articles published in English, however this limit is not included any of the search strategies. Please confirm whether the limit was applied during the searching or inclusion screening.

Response: The limit was applied during the inclusion screening stage.

Clinical effectiveness searches

A2. Please confirm which resources were used to identify recent systematic literature reviews (SLRs), practice guidelines and conference abstracts (appendix D; page 3), and explain why trials registers were not included in this update search when they were included in the original 2010 searches.

Response: Separate targeted searches were conducted to identify recent SLRs and practice guidelines published in the previous 2 years for reference checking. Conference abstracts from the American College of Rheumatology, European League Against Rheumatism, British Society for Rheumatology, American Society of Nephrology Kidney Week, and the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) conferences were searched on conference websites or via abstract book (ASN Kidney week was only available as a PDF book). We acknowledge that trial registers were not searched.

A3. Please confirm whether appendix D; Table 4 was used to identify non-RCT studies (rather than RCTs, as stated in the company submission).

Response: That is correct, the searches shown in Table 4 of Appendix D were used to identify non-RCTs. The table title has been updated to reflect this change.

A4. Please provide details of the numbers of references retrieved from the individual databases in the Cochrane Library (CDSR and CENTRAL), if this information is available when the Cochrane Library is searched via the Ovid interface.

Response: These were previously run together in Ovid. We re-ran to get the results by individual source and there were 252 hits for CENTRAL and 26 for CDSR.

Cost-effectiveness searches

A5. Please provide details of the hosts used to search Embase and EconLit, and the date ranges searched for all databases.

Response: Both Embase and EconLit were searched via OVID on 28th January 2020 and 31st January 2020 respectively. No date range filters were applied to any database searches.

A6. Please provide correct details of the search terms used in line #28 of the PubMed search (appendix G; page 4).

Response: The correct search terms are 'Search (((costs AND cost analysis)))'

A7. Please provide the name of the 'CRD Database' searched (appendix G; page 5).

Response: This stands for Centre for Reviews and Dissemination (please see https://www.crd.york.ac.uk/CRDWeb/). Databases searched included DARE, NHS EED and HTA.

A8**.** Please supply the full name of the 'ACS' resource used for the cost effectiveness searches (appendix G; page 5).

Response: This should read 'ACR - American College of Rheumatology'

A9. Please provide full details of all conference and grey literature searches, and information on how SLR studies were identified (company submission; pages 121-122).

Response: Conference abstracts were identified by searches of all eight bibliographic databases but primarily were mainly identified through the ISPOR database. Grey literature was identified by examining studies reported in systematically identified literature and internet searches.

SLRs were identified through searches of all eight bibliographic databases identified in Appendix G, using the search strategies described for each database. All search results were pooled into a single database and screened for inclusion against the PICOS criteria identified in Table 54 of Document B. A table with a full list of included and excluded studies along with reasons may be found on pages 10 and 6 respectively of Appendix G.

Health-related quality of life/ Cost and healthcare resource use searches

A10. Please provide full details (including search terms) of searches carried out in PubMed or on the NICE website (appendix H and I; page 3).

Response: Please see Appendix H and I for details of how searches were carried out. In the first instance, the NICE website was searched for relevant HTA's reporting on costs and utilities. If no relevant data was identified, then further data was collected using PubMed by searching for relevant keywords and Medical Subject Headings (MeSH) under each organ system. Please also see the included file, 'ID1591 A10. Clarification' for details on how the NICE website was searched.

Decision problem

A11. Patients with severe active lupus nephritis or CNS lupus were excluded from the BLISS trials. Could the company confirm that these patients would not be eligible for belimumab in UK clinical practice?

Response: Patients with severe active CNS lupus were excluded from the pivotal BLISS trials and no evidence to support the use of belimumab is available in this population. Patients with lupus nephritis (LN) were also excluded from the pivotal BLISS trials; however, a clinical trial in this population, BLISS-LN, has recently been published[(1)]. BLISS-LN met its primary endpoint, with significantly more patients in the belimumab group than the placebo group achieving primary efficacy renal response (a ratio of urinary protein to creatinine of ≤0.7, an estimated glomerular filtration rate [eGFR] that was ≤20% below the pre-flare value or ≥60 ml/minute/1.73 m^2 of body-surface area, and no use of rescue therapy)[(1)]. A Type II variation for an indication extension has been submitted to the EMA on 24th June 2020 and an outcome is anticipated in H1 2021. Therefore, SLE patients with lupus nephritis may be eligible for treatment with belimumab in the future.

A12. Priority question. The company present evidence from NHS England to argue that rituximab is not a comparator because it can only be prescribed at a later line of therapy to that for belimumab (company submission, page 103). However, given that several criteria are listed could the company explain

precisely which of those criteria indicate that rituximab can only be prescribed at a later line of therapy?

Response: The patient pathway presented in the Clinical Commissioning Policy Rituximab for refractory Systemic Lupus Erythematosus (SLE) in adults and post-pubescent children [200402P][(2)] clearly shows that patients who are eligible for belimumab should be considered for treatment with belimumab before rituximab. This pathway is reproduced below. Please note the text which has been marked with a red box, clearly stating belimumab should be considered first as a licensed and NICE-approved treatment. In addition, the left-hand side part of the pathway, describing belimumab-eligible patients, clearly positions rituximab as an option only in patients who do not respond to belimumab.

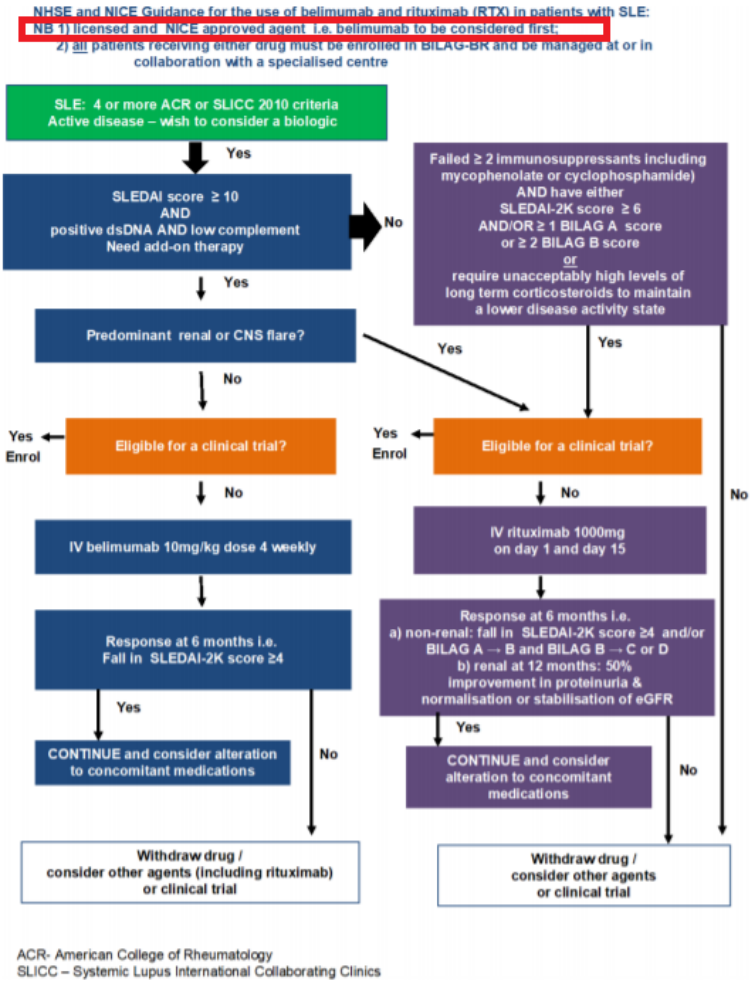


Figure 1. patient pathway presented in the Clinical Commissioning Policy Rituximab for refractory Systemic Lupus Erythematosus (SLE) in adults and post-pubescent children [200402P]

Furthermore, the eligibility criteria for rituximab on page 6 of the Clinical Commissioning Policy[2] (quoted on page 103 of the company submission) state that patients should have been assessed as ineligible for belimumab in order to be considered for treatment with rituximab. The relevant text is highlighted below:

Eligibility criteria:

Rituximab should be considered for adults and post-pubescent children with moderate or severe refractory SLE with active disease, who have failed to respond or have had adverse events to 2 or more immunosuppressive therapies (one of which must be either mycophenolate or cyclophosphamide, unless contraindicated) and have:

EITHER

- Disease activity with at least one BILAG A and/or two B scores or a SLEDAI-2K score ≥ 6

Or

- Requiring unacceptably high levels of oral glucocorticoids e.g. more than 7.5mg prednisolone in an adult per day, to maintain a lower disease activity state

AND

- been assessed as not eligible for clinical trials or belimumab.

Figure 2. Rituximab eligibility criteria

GSK understand that prior to the availability of belimumab under the Managed Access Agreement rituximab was the only available biologic within England for patients who had continued disease activity despite standard therapy. As an additional and licenced biologic has been available (i.e. belimumab) since October 2016, the updated Clinical Commissioning Policy[2] very clearly positions rituximab as a later line of therapy in patients who are eligible for treatment with belimumab, rather than as a direct alternative.

A13. **Belimumab is now available as a SC formulation which the company submission (CS) states "reduces the burden on NHS resources as regular clinic time is not required for administration." What proportion of patients do you anticipate will use the SC formulation and what proportion will use the IV formulation?**

Response: Based on the [redacted], we expect that there will be [redacted], resulting in approximately [redacted] of patients on [redacted] by the end of Year 1 [redacted] and [redacted] of patients on [redacted] by the end of year two. We do not expect the [redacted]

A14**. According to the NICE scope, SC injection is available for use in adults only. Could SC injection be extended to children and adolescents given appropriate parental training?**

Response: There is currently no data supporting the use of the SC formulation in children and the marketing authorisation for the SC formulation includes adult patients only. As outlined on page 114 of the CS, there is an on-going Phase 2 PK-PD study assessing belimumab SC in paediatric patients with SLE, which is expected to complete in September 2023.

A15**. The NICE scope specifies that patients eligible for belimumab have a "high degree of disease activity despite standard therapy". According to Figure 1 in the CS, do you anticipate that belimumab will mainly be offered as a third line treatment?**

Response: The company anticipates that belimumab will be offered primarily as third-line treatment after failure of antimalarials and immunosuppressants (both of which may be supplemented with corticosteroids). A minority of patients may receive belimumab at second line, in line with the EULAR guidelines which state that "*Belimumab should be considered in extrarenal disease with inadequate control (ongoing disease activity or frequent flares) to first-line treatments (typically including combination of an antimalarial and prednisone with or without immunosuppressive agents), and inability to taper glucocorticoids daily dose to acceptable levels (i.e., maximum 7.5 mg/day*)"[(3)]. It should also be noted that pivotal trials of belimumab did not require patients to have received a certain number of prior therapies, or prior therapy with immunosuppressants.

A16**. Why were rate and duration of remission not considered suitable outcomes in this submission?**

Response: The rate and duration of remission was not directly measured in clinical trials of belimumab and at present, there is no universally accepted, validated definition of remission in SLE. This is despite the fact that a consensus framework for development of such definitions exists[4]. The EULAR guidelines define remission as the "*absence of clinical activity with no use of glucocorticoids and immunosuppressive drugs*", and at the same time acknowledge that remission defined in such a way is infrequent[(3)]. In parallel, clinicians have sought to define a

"low disease activity state", which could be a more attainable target than remission for treat-to-target approaches; however, the proposed definitions[5,6] also vary.

Systematic literature review

A17. According to the inclusion criteria of your systematic review (appendix D, Table 6) the population was 'Adults (≥18 years) diagnosed with SLE'. However, according to the decision problem (CS, Table 1) the population included is 'People aged 5 years or more with SLE'. Please clarify what searches were performed to identify studies in children between 5 and 18 years.

Response: No searches were performed in people over the age of 5 as the CS focuses on an adult population with SLE as does the economic modelling. The majority of clinical effectiveness data available on belimumab is for adult patients aged 18 years and older with SLE.

A18. Priority question. On page 11 of appendix D (Table 7), the numbers of included studies are reported: 65 unique studies in total (48 RCTs, and 17 non-RCTs). Please provide a full list of these included studies split by RCTs and non-RCTs with all related references for each unique study. And please clarify how each included study was used in the submission.

Response: A SLR was commissioned to include all treatment options in use for patients with SLE globally for markets beyond the UK. Please see the document named 'ID1591_Belimumab_Appendix D EVA-26527 SLE SLR Report for NICE' (which was previously embedded in Appendix D) for comprehensive details about all studies included in the submission. This document describes a break down by RCT and non-RCT, along with details of how each study was used in the submission.

A19. The SLR included studies of adults (≥18 years) with SLE, but not if the main participant population included ≥15% with lupus nephritis. How was this percentage arrived at? How many studies were excluded by imposing this criterion?

Response: The '≥15% with lupus nephritis' exclusion criteria was based on internal discussion as no studies or selection of studies were identified to guide this criterion. It was considered that populations that included '≥15% with lupus nephritis' may impact on the reported treatment effect for the population under consideration for this

Decision Problem. Publications which included a lupus nephritis population were also searched to see if the treatment effect was reported in a sub-group excluding these patients. This filter resulted in the exclusion of 18 studies.

A20. In Table 6 of appendix D, outcomes are listed in the comparators section of the SLR. Could you confirm the eligible comparators for the SLR?

Response: Please find the correct list of comparators:

Rituximab
Cellcept® (mycophenolate mofetil)
Prednisolone and other steroids
Hydroxychloroquine and other anti-malarials
Azathioprine
Cyclophosphamide
Methotrexate
Placebo and mixed routine care (i.e., combination treatments)

A21. The systematic review appeared to be limited to studies published in English only. How many relevant studies were omitted due to this language restriction?

Response: A single study was excluded due to not being published in English. Please see the file 'ID1591_Belimumab_Appendix D EVA-26527 SLE SLR Report for NICE' for details of all excluded studies along with rationale.

A22. Priority question. Could you provide a list of studies excluded from the review with reasons for their exclusion?

Response: Please see the attached Excel file 'ID1591_Belimumab_Appendix D EVA-26527 SLE SLR Report for NICE'.

Belimumab trials

A23. Priority question. Please provide full baseline characteristics (as in Table 37 of the CS) from the two BLISS trials (BLISS-52 and BLISS-76) separately for the relevant population (HDA-2).

Response: Please see the attached document 'ID1591 A23. Clarification'.

A24. Priority question. Please provide full final results from the two BLISS trials (BLISS-52 and BLISS-76) separately for the relevant population (HDA-2). Please present these results using the same format as reported for the pooled results in Tables 38 and 39 of Section B 2.7.1 of the CS, and in appendix L.

Response: Please see the attached document 'ID1591 A24. Clarification'.

A25. How many patients aged between 5 and 18 years were included in each trial of the two BLISS trials (BLISS-52 and BLISS-76). Please provide numbers per arm and for the relevant population (HDA-2).

Response: Patients were required to be ≥ 18 years of age in both the BLISS-52 and BLISS-76 trials which therefore did not include any paediatric patients. Because of this, the HDA-2 subgroup analysis conducted on the BLISS trial data also only included adults ≥ 18 years of age. The paediatric population has been studied in a randomised, double-blind, Phase 2 study (PLUTO) evaluating IV belimumab 10mg/kg plus standard therapy in patients aged 5 to 17 years with active SLE. Part A of this study (52-week double-blind phase) has been completed and provided in Appendix O of CS. Ongoing, open-label safety continuation studies are included in section B.2.11 (Ongoing studies) of the CS.

A26. Priority question. Please provide details of the method used to pool data from the BLISS-52 and 76 trials.

Response: BLISS-52 and BLISS-76 data were pooled to achieve a more stable estimate of belimumab's global treatment effect. This was considered appropriate as the studies were essentially identical in design and the effects of belimumab on the endpoints of interest were similar between the studies. With reference to the principles outlined in ICH E9 (Statistical Principles of Clinical Trials) when pooling the data across these studies, we considered study design, inclusion and exclusion criteria relative to disease severity, and whether the studies were run contemporarily such that the SoC treatment options were similar. These studies followed very similar protocols, were of nearly identical design, had identical inclusion and exclusion criteria, and were conducted over the same time period. Nevertheless, given the heterogeneous presentation of SLE disease and the fact that the Phase 3 program was run globally, one should expect to have variation in the patient population, both within the studies (e.g. between different centres) and between the

studies (analogous to differences between centres within the same study). While pooling is not necessary to establish the effectiveness of belimumab, it was considered appropriate in order to evaluate treatment effects in high disease activity subgroups of interest, given that the individual studies were not designed to provide sufficient power to demonstrate effectiveness within subgroups.

Patient-level data from BLISS-52 and BLISS-76 were simply combined into one aggregated patient-level data set. Where the size and composition of the subpopulation allowed, analyses were performed controlling for baseline stratification factors and study; otherwise, some of the covariates were omitted or unadjusted analyses were performed.

When the two Phase 3 studies were pooled, a test for a treatment-by-study interaction was undertaken for the SRI analysis and the treatment-by-study interaction was >0.5, suggesting that the effect of belimumab was not significantly different between BLISS-52 and BLISS-76.

A27. Please provide UK patient numbers by trial and treatment group for BLISS-52 and BLISS-76. Please also provide the same numbers for the HDA-2 populations in each trial by trial arm.

Response: There were no UK patients enrolled in BLISS-52. In BLISS-76, a total of 11 patients from the UK were enrolled, constituting 1.3% of the total trial population. Of those, 6 patients were randomised to placebo, 4 to the unlicensed 1 mg/kg belimumab dose and 1 to the licensed 10 mg/kg dose. The numbers for the HDA-2 population cannot be readily provided within the time frame for responding to these clarification questions but are likely to be small considering the total size of the UK patient sample in BLISS-76.

A28. How generalisable to UK clinical practice are the patients in the BLISS trials? This includes features of the disease, patient characteristics and concomitant medication.

Response: The population enrolled in the BLISS trials is representative of patients with moderate to severe, active SLE in the UK. As shown in the table below, the baseline demographics of patients taking part in the BLISS trials were similar to those receiving belimumab in the UK and enrolled in BILAG-BR. Patients in both

BLISS trials and BILAG-BR were predominantly females of working age, which is also consistent with the patterns of lupus incidence in general. The racial distribution was slightly different, with the BLISS trials including fewer African Heritage patients compared with the BILAG BR. In addition, the proportion of White/Caucasian patients was similar between the UK BILAG-BR study and BLISS-SC but was higher in the pooled IV population. The proportion of Asian patients was similar.
Disease activity (as described by SELENA SLEDAI/SLEDAI-2K and BILAG scores) and daily steroid dose both appeared higher in the BILAG-BR than the BLISS trials; however, this is likely to be a reflection of the BILAG-BR only collecting data on a subgroup of patients with high disease activity (HDA-1), whereas a broader population was enrolled in the BLISS trials.

With regards to medication use, the BILAG-BR identified antimalarials, immunosuppressants, and steroids as standard of care. These are the same medication classes that were considered standard therapy in the BLISS trials. However, despite the fact the BILAG-BR patients appeared to have higher disease activity than those in BLISS trials, a smaller proportion of patients were reported to receive standard therapy at baseline. This is likely a result of incomplete reporting in the real-world BILAG-BR study than a true difference. Concomitant medications were entered into the BILAG-BR data through a free-text field, so that the lower-than-anticipated proportion of patients receiving standard of care could stem from both errors in data entry and missing data.

Table 1. Baseline patient characteristics for BLISS patients

Baseline Characteristics	BILAG-BR (belimumab-treated patients only) [redacted]	BLISS-SC		Pooled BLISS-52 and BLISS-76 data	
		Belimumab 200 mg SC N=556	Placebo N=280	Belimumab 10 mg/kg IV N=563	Placebo N=562
Patient characteristics					
Female, N (%)	[redacted]	521 (93.7)	268 (95.7)	539 (95.7)	522 (92.9)
Age (years), mean (SD)	[redacted]	38.1 (12.10)	39.6 (12.61)	37.9 (11.3)	38.1 (12.0)
Race, N (%)					
White/Caucasian	[redacted]	336 (60.4)	166 (59.3)	260 (46.2)	270 (48.0)
Asian	[redacted]	119 (21.4)	63 (22.5)	127 (22.6)	116 (20.6)
African American/African Heritage	[redacted]	56 (10.1)	30 (10.7)	50 (8.9)	50 (8.9)
American Indian/Native American	[redacted]	43 (7.7)	21 (7.5)	126 (22.4)	125 (22.2)
Native Hawaiian or Other Pacific Islander	[redacted]	2 (0.4)	0	0	1 (0.2)
Multiracial	[redacted]	6 (1.1)	3 (1.1)	4 (0.7)	3 (0.5)
Other	[redacted]	NR	NR	NR	NR
Missing (n)	[redacted]	-	-	-	-
Disease characteristics					
SLE disease duration (years), mean (SD)	[redacted]	6.4 (6.60)	6.8 (6.83)	6.08 (6.42)	6.66 (6.48)
Missing, n/N (%)	[redacted]	-	-	-	-
SELENA-SLEDAI at baseline, mean (SD)	[redacted]	10.5 (3.19)	10.3 (3.04)	9.75 (3.77)	9.75 (3.79)
At least BILAG 1A or 2B, N (%)	[redacted]	388 (69.8)	210 (75.0)	332 (59.0)	353 (62.8)
Medication usage					
Average daily prednisone dose (mg/day), mean (SD)	[redacted]	10.8 (8.21)	11.2 (9.09)	10.9 (9.1)	10.7 (8.5)
Number (%) of patients taking:					
Antimalarials	[redacted]	391 (70.3)	189 (67.5)	353 (62.7)	381 (67.8)
Immunosuppressants	[redacted]	244 (43.9)	137 (48.9)	271 (48.1)	276 (49.1)

Baseline Characteristics	BILAG-BR (belimumab-treated patients only) [redacted]	BLISS-SC		Pooled BLISS-52 and BLISS-76 data	
		Belimumab 200 mg SC N=556	Placebo N=280	Belimumab 10 mg/kg IV N=563	Placebo N=562
Steroids	[redacted]	481 (86.5)	241 (86.1)	478 (84.9)	488 (86.8)

**at baseline of any treatment round.*

†SLEDAI-2K was used in the BILAG-BR

‡Regular steroid dose is presented.

[redacted]

BILAG: British Isles Lupus Assessment Group; BR: Biologics Registry; SD: standard deviation; SELENA: Safety of Estrogen in Lupus Erythematosus National Assessment; SLE: systemic lupus erythematosus; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index.

To further confirm generalisability of the BLISS trials to UK clinical practice, we asked for expert opinion from several consultant rheumatologists in England These experts concluded that the patients in the BLISS trials are, in their opinion, generalisable to the UK population. They were global studies with a large number of patients, and so they are very representative of the mixed ethnicity populations in the UK.

A29. The majority of patients across the BLISS trials are women (over 90%):

a) Was any analysis conducted with the subgroup of men?

Response: The female predominance of SLE as a condition was also reflected in the enrolment into the BLISS trials. Pre-specified exploratory subgroup analyses of the primary efficacy endpoint by gender was conducted for BLISS-SC and for pooled data from BLISS-52 and BLISS-76. The results are presented below.

b) Please present the results of a subgroup analysis for men.

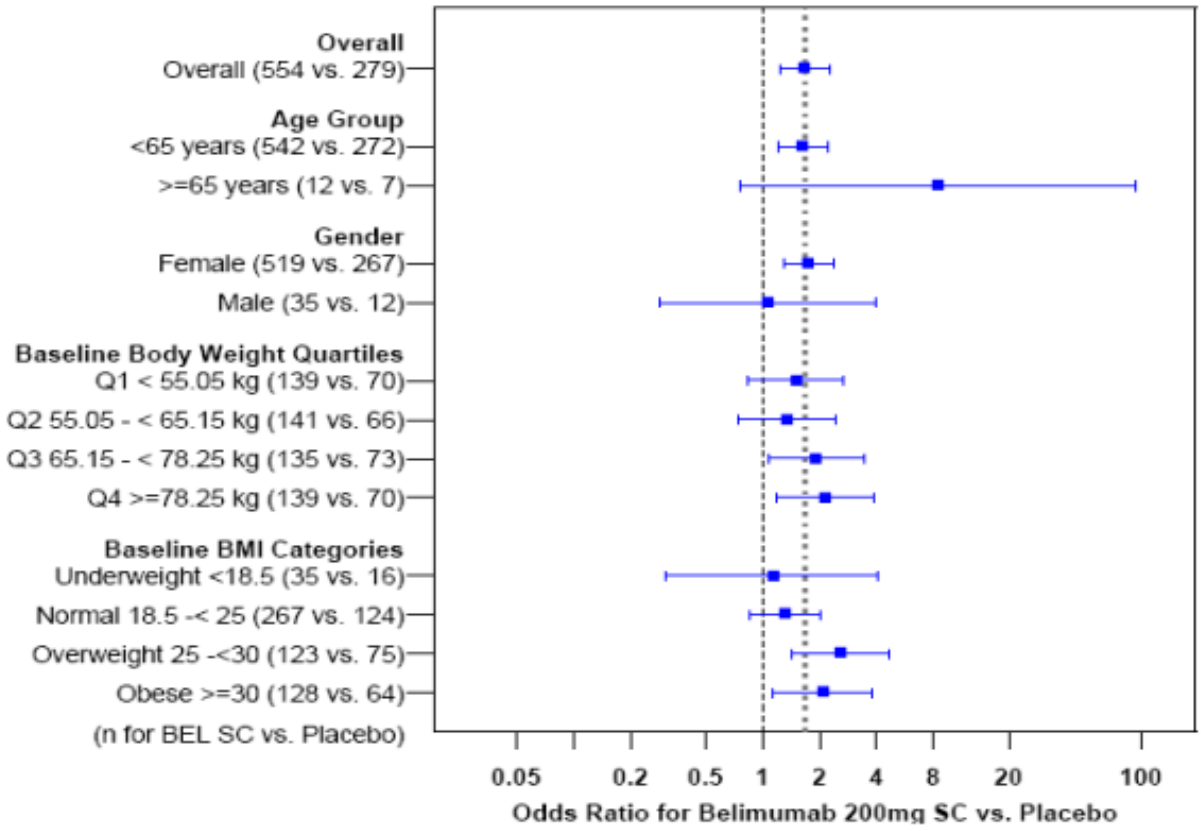
Response: BLISS-SC: No significant treatment-by-subgroup interaction was observed when exploring gender (see Table 2.78 from BLISS-SC CSR reproduced below and the forest plot that includes subgroup analysis by gender), but the relatively small number of male patients enrolled (n=47) limit meaningful interpretation.

SRI Response at Week 52 by Gender

	Placebo (N=280)	Belimumab 200mg (N=556)
Gender: Male		
n	12	35
Response	6 (50.0%)	18 (51.4%)
Observed difference vs. Placebo		1.43%
OR (95% CI) [1] vs. Placebo		1.06 (0.29, 3.93)
P-value [1]		0.9319
Gender: Female		
n	267	519
Response	129 (48.3%)	322 (62.0%)
Observed difference vs. Placebo		13.73%
OR (95% CI) [1] vs. Placebo		1.75 (1.30, 2.36)
P-value [1]		0.0002
Interaction p-value [2]		0.4648

Table 2. SRI Response at Week 52 weeks by Gender

Figure 3. Odds Ratio for Belimumab 200mg SC vs Placebo



BLISS-52 and BLISS-76 pooled analysis: When exploring gender subgroups in the pooled phase 3 IV data, treatment-by-subgroup interactions were not statistically significant, with belimumab treatment offering numerical benefits in each subgroup relative to placebo. Please see the results table below. The forest plot including subgroup analysis by gender for the licensed 10 mg/kg dose is also presented.

Figure 4. Primary response at Week 52 by subgroup (dropout = failure), gender

T95 Primary response at Week 52 by subgroup (dropout = failure), gender

	Male			Female		
	Placebo N=40	1 mg/kg N=35	10 mg/kg N=24	Placebo N=522	1 mg/kg N=524	10 mg/kg N=539
n in the subgroup	40	35	24	522	524	539
No.(%) Response	15 (37.5%)	15 (42.9%)	13 (54.2%)	203 (38.9%)	243 (46.4%)	272 (50.5%)
Observed difference vs. Placebo		5.36	16.67		7.49	11.57
OR (95% CI)[1] vs. placebo		1.47 (0.52, 4.12)	2.40 (0.78, 7.37)		1.40 (1.09, 1.81)	1.65 (1.29, 2.13)
P–value[2]		0.8989	0.5507		NA	NA

[1]From logistic regression for the comparison between each belimumab dose and placebo in pooled data. Independent variables will include treatment group, baseline SELENA SLEDAI score (<= 9 vs. >= 10), baseline proteinuria level (< 2 g/24 hour vs. >= 2 g/24 hour equivalent), race (African descent or indigenous–American descent vs. other) and study.
[2]For treatment by subgroup interaction effect from a logistic regression model by adding the subgroup and interaction effect to the above model.

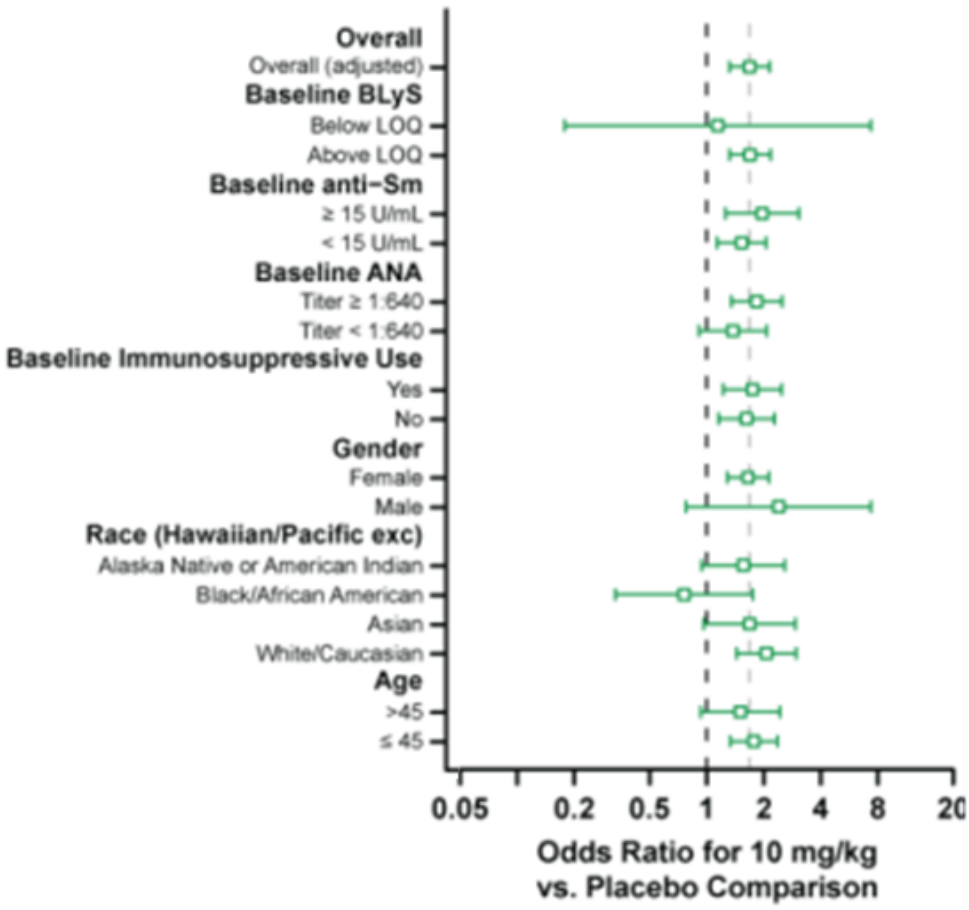


Figure 5. Odds Ratio for 10 mg/kg vs. Placebo Comparison

c) Do you have evidence that belimumab will work similarly in men?

Response: The response rates observed in men were generally consistent with those observed in the overall population, as evidenced by the 95% CIs that overlap the odds ratio from the primary analysis. However, the number of males enrolled was small compared with the females, and the associated CIs for the subgroup of men are very wide for both the SC and pooled IV populations. Overall, there is no clinical trial evidence to suggest that belimumab may be less efficacious in men.

A30. Belimumab had a possible link with depression in the BASE trial.

a) Were patients with depression excluded from the BLISS trials?

Response: Patients with depression were not specifically excluded from the pivotal IV BLISS trials, BLISS-52 and BLISS-76. The BLISS-SC trial was initiated after the IV trial results were available. This trial excluded patients at high risk of suicide. Specifically, the exclusion criterion specified in the BLISS-SC protocol was:

“Subjects who have evidence of serious suicide risk including any history of suicidal behaviour in the last 6 months and/or any suicidal ideation of type 4 or 5 on the CSSRS in the last 2 months or who in the investigator's opinion, pose a significant suicide risk.”

b) Would it be necessary to screen for depression before prescribing belimumab in practice?

Response: In line with the summary of product characteristics, the risk of depression and suicide should be assessed before initiating treatment with belimumab and monitored during treatment. Specifically, the summary of product characteristics states that “*Physicians should assess the risk of depression and suicide considering the patient's medical history and current psychiatric status before treatment with Benlysta and continue to monitor patients during treatment. Physicians should advise patients (and caregivers where appropriate) to contact their healthcare provider about new or worsening psychiatric symptoms. In patients who experience such symptoms, treatment discontinuation should be considered*.”

A31. In Table 13 of the CS, length of treatment/extent of discontinuations over time is listed as “Available upon request” for the BLISS LTEs. Please could you provide this information?

Response: Please find below the integrated analysis of the Phase 2 and Phase 3 LTE IV studies for length of treatment/discontinuations.

Table 3. Subject Completion Status by Study Year – Year 0 to 3

Table 4. Subject Completion Status by Study Year – Year 3 to 7

Table 5. Subject Completion Status by Study Year – Year 7 to 10+

A32. Table 15 in the CS lists concomitant therapies taken by patients in BLISS. Are you able to provide data on the numbers of prior therapies taken by patients? Please provide the same numbers for the total trial populations and for the HDA-2 populations in each trial by trial arm.

Response: BLISS-SC, BLISS-52 and BLISS-76 required patients to be on a stable treatment regimen for at least 30 days prior to the first dose of belimumab. Data on concomitant medications received at baseline was provided in the company submission (Table 15 page 47 for the ITT population and Table 37, page 88 for HDA-2). Medication history as such was not collected in the BLISS trials, so that the number of prior therapies, or lines of treatment received, is not available.

A33. Priority question. PLUTO does not appear to be mentioned in Table 12 of the CS, the summary of presented evidence. Does it present any new evidence that could be mentioned in Table 13?

Response: The PLUTO trial provides information on efficacy, safety, and pharmacokinetics of belimumab in the paediatric SLE population, as well as on the effects of belimumab on quality of life in this population. Results from the Phase 2 randomised, placebo-controlled, double-blind 52-week treatment phase of PLUTO (Part A) have been summarised in Appendix O in the CS, but the trial was accidentally omitted from Table 12, which provided an overview of the presented evidence. PLUTO was, however, deliberately excluded from Table 13, which outlined how the evidence collected since the previous submission addresses the key areas of uncertainty identified during TA397. These areas of uncertainty pertained to the adult population only, as belimumab had no marketing authorisation for the treatment of paediatric patients at the time, and no evidence in the paediatric population was presented to NICE as part of TA397. Because no adult patients were included in PLUTO, there is no overlap in patient populations between PLUTO and TA387, and the trial provides no information that could address the areas of uncertainty arising from that appraisal.

A34. Were organ system-specific HRQoL scales utilised for patients when appropriate? If not, why not? If so, please list all utilised scales, the number of patients who used these scales, and the findings of these scales.

Response: No organ-system-specific HRQoL scales were used in the BLISS pivotal trials and, to the best of our knowledge, there are no robust organ system-specific scales that have been validated for SLE.

A35. Please clearly define the pre-planned subgroups for the BLISS-SC trial. On page 46 of the CS, the pre-planned subgroups for the BLISS-SC trial by region includes US/Canada, Europe/Australia/Israel, Asia/Americas, but then also excludes the US/Canada.

Response: The pre-planned subgroups for BLISS-SC were listed correctly on page 46 of the CS. It is the subgroup nomenclature that perhaps was not clear. The analysis by region included the following subgroups:

1. US and Canada
2. Europe, Australia, and Israel
3. Asia and Americas, with Americas excluding US and Canada, which were analysed as a separate subgroup (1).

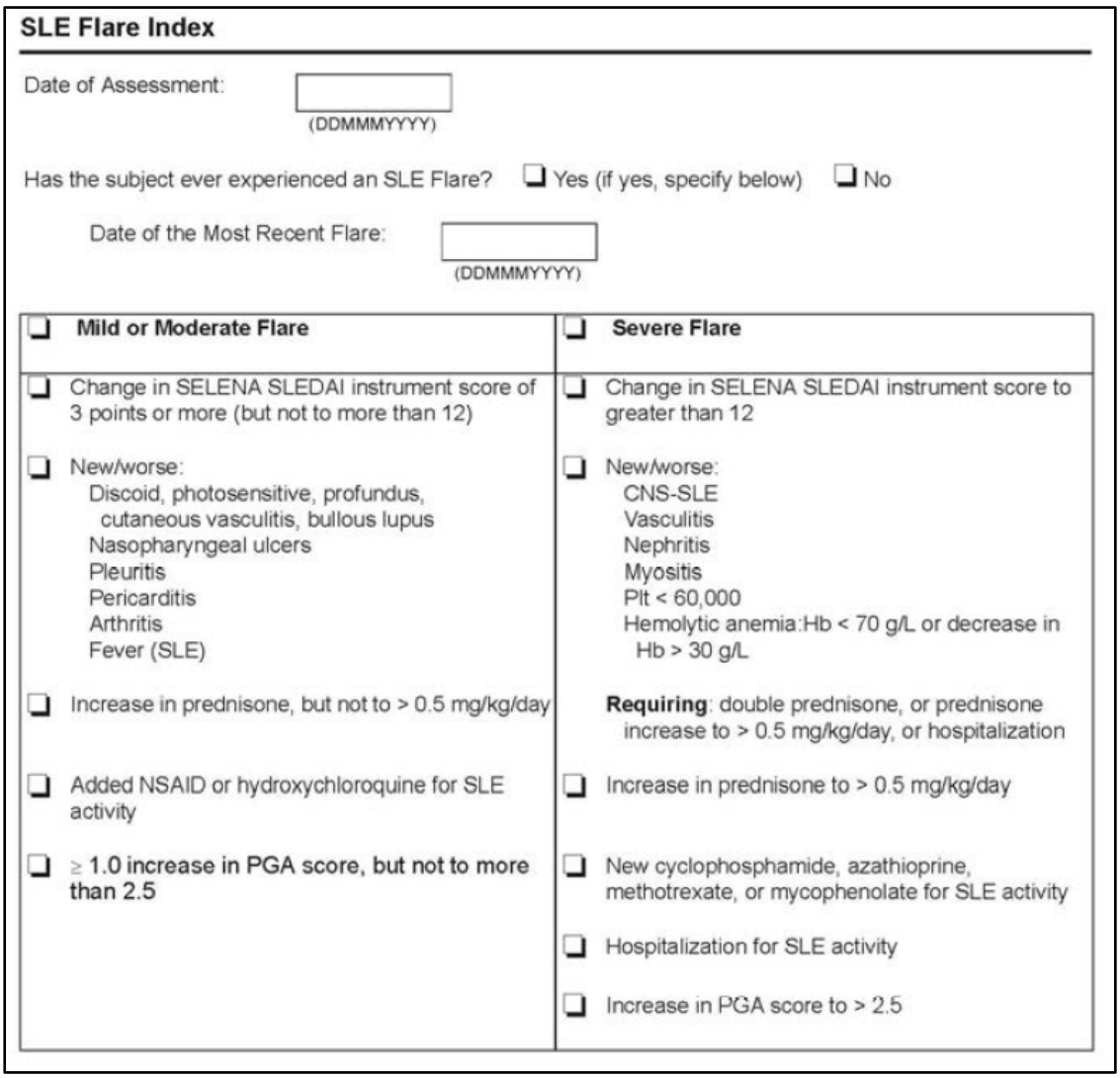
A36. Please define severe flare and how it is differentiated from mild/moderate flare.

Response: Flares were categorised using the SFI (SELENA-SLEDAI Flare Index), as used in the SELENA trials[7,8]. The SFI categorises flares as “mild or moderate” or “severe” based on 5 variables:

- Change in SELENA-SLEDAI score from the most recent assessment to current.
- Change in signs or symptoms of disease activity.
- Change in prednisone dosage.
- Use of new medications for disease activity or hospitalization.
- Change in PGA score.

The composites of mild/moderate and severe flares are provided in the image below. The presence of ≥1 criterion is sufficient to define an SFI flare (e.g. the appearance or worsening of CNS SLE alone would trigger a severe flare). However, in the BLISS trials, a modification was applied to exclude severe flares triggered by an increase in SELENA-SLEDAI score to >12, since this may only represent a modest increase in

disease activity given that the trials were open to patients with high disease activity (inclusion criterion of SELENA-SLEDAI score ≥6 at screening for BLISS-52 and BLISS-76 and ≥8 for BLISS-SC). Therefore, any flares triggered by a ≥3-point increase in SELENA-SLEDAI were classed as mild/moderate.



SLE Flare Index

Date of Assessment: (DDMMMYYYY)

Has the subject ever experienced an SLE Flare? ❑ Yes (if yes, specify below) ❑ No

Date of the Most Recent Flare: (DDMMMYYYY)

❑ **Mild or Moderate Flare**	❑ **Severe Flare**
❑ Change in SELENA SLEDAI instrument score of 3 points or more (but not to more than 12)	❑ Change in SELENA SLEDAI instrument score to greater than 12
❑ New/worse: Discoid, photosensitive, profundus, cutaneous vasculitis, bullous lupus Nasopharyngeal ulcers Pleuritis Pericarditis Arthritis Fever (SLE)	❑ New/worse: CNS-SLE Vasculitis Nephritis Myositis Plt < 60,000 Hemolytic anemia:Hb < 70 g/L or decrease in Hb > 30 g/L
❑ Increase in prednisone, but not to > 0.5 mg/kg/day	**Requiring**: double prednisone, or prednisone increase to > 0.5 mg/kg/day, or hospitalization
❑ Added NSAID or hydroxychloroquine for SLE activity	❑ Increase in prednisone to > 0.5 mg/kg/day
❑ ≥ 1.0 increase in PGA score, but not to more than 2.5	❑ New cyclophosphamide, azathioprine, methotrexate, or mycophenolate for SLE activity
	❑ Hospitalization for SLE activity
	❑ Increase in PGA score to > 2.5

Figure 6. SLE Flare Index

Adverse Events

A37. Priority question. All adverse events are reported for the whole trial populations in Section B.2.10 and Appendix F. Please provide the same data for the HDA-2 population, including all data as reported in Tables 49-50 of the CS and all Tables in Appendix F.

Response: Please see the attached document 'ID1591 A37. Clarification'.

Non-RCT evidence

A38. Priority question.

a) **Please justify the choice of the BLISS-76 US LTE (rather than the other BLISS LTEs) results for belimumab used in the indirect cohort study (Section B.2.6.3.3)?**

Response: The primary analysis for the propensity scored matching comparative analysis was based on the BLISS-76 US LTE study. This was because the BLISS-76 US LTE/Toronto Lupus Cohort (TLC) allowed matching on 14 clinical predictors of organ damage (17 operationalised variables), whereas the pooled BLISS LTE/TLC dataset allowed matching on fewer clinical predictors (12; 16 operationalised variables [smoking status excluded due to unexpected large differences]). Exploratory analyses were conducted for the same end points using the pooled LTE/TLC dataset. The results of this are provided in the study report (provided with the CS) and have also just been published (Urowitz et al., 2020; https://lupus.bmj.com/content/7/1/e000412)
Results are provided below:

Table 6. Change in 5-year SDI

Change in 5-year SDI	Belimumab Mean (95% CI)	ST Mean (95% CI)	Mean Treatment difference (95% CI)
Based on BLISS-76 US LTE/TLC [Primary end point]	0.283 (95% CI 0.166 to 0.400)	0.717 (95% CI 0.500 to 0.934)	−0.434 (95% CI −0.667 to −0.20) p<0.001
Based on Pooled LTE/TLC [Exploratory post-hoc analysis]	0.265 (95% CI 0.180 to 0.350)	0.718 (95% CI 0.548 to 0.889)	–0.453 (95% CI –0.646 to –0.260) p<0.001

b) **Please also provide further justification for the choice of the Toronto Lupus Cohort for the source of standard of care data in the matched analysis.**

Response: A systematic review was conducted and aimed to identify SLE cohorts and registries (see response to d) below). Key cohort characteristics were identified for scoring the SLE cohorts. These included the compatibility of the cohort outcomes and the BLISS LTE trial outcomes, the size of the cohort and the severity of the disease found in the cohort. A score was calculated as the square root of the product of the number of patients, mean SLEDAI score, and mean SDI score. The square root was used only to reduce the scale of the score. Other characteristics were less able to be quantified and included racial make-up of the cohort, and how similar it was to the BLISS LTE trials. Analyses of these characteristics were left for qualitative comparison.

A summary of the cohorts and their respective score is shown in the table below. TLC received the highest score, due not only to the size of the cohort, but also to the high disease activity (SLEDAI) and organ damage (SDI) in the cohort. RELESSER, the largest of the cohorts, was scored much lower than TLC due to its particularly low SLEDAI and SDI scores.

The scores for several cohorts were zero due to cohorts reporting disease activity or organ damage using a measure different than those used in the BLISS trials. This was particularly true for disease activity, for which there are a number of other commonly used measures. In other cases, for instance the Danish and Swedish databases, SLE disease activity and organ damage have not been reported and may well not be recorded in any form in the databases. Cohorts with few publications frequently lacked disease activity or organ damage measures. This may be due to sparse reporting or lack of recorded data. Other cohorts were simply too small to warrant consideration.

A cohort used as a SoC comparator to a BLISS LTE trial should have a similar racial makeup due to the racial differences in SLE disease progression. The far right columns of table below show the proportions of black and Asian participation in the cohorts. None of the cohorts matched the racial makeup of the BLISS US LTE trial. Two large cohorts were disqualified because of their focus on specific minorities,

e.g., the GLADEL and RELESSER cohorts. The SLICC cohort matched BLISS black participation best while the JHLC matched Asian participation best. TCL had half the rate of black participation and twice the rate of Asian participation as the BLISS US LTE. However, these percentages were based on the entire cohort; the racial mix of patients with a SLEDAI score ≥ 6 could be quite different.

Table 7. Identification of TLC and cohort scoring by size, disease severity and BLISS US LTE compatibility

Cohort/Register	Publications	n	SLEDAI (0/mean)*	SDI (0/mean)*	Score†	% Black (BLISS = 21.3)	% Asian (BLISS = 4.9)
Toronto Lupus Clinic Prospective Cohort	39	1,781	8.7	1.6	**157**	10	9
Grupo Latinoamericano de Estudio del Lupus (GLADEL) Cohort	12	1,480	13.6	0.74	**122**	0	0
1000 Canadian Faces of Lupus Cohort	8	1,724	4.3	1.6	**109**	9.5	14.7
Johns Hopkins Lupus Cohort	43	2,265	3.5	1.2	**98**	37.6	3.4
Registry of Systemic Lupus Erythematosus Patients of the Spanish Society of Rheumatology (RELESSER)	5	3,658	2	1	**86**	0.2	0.6
Montreal General Hospital / McGill University Health Centre Cohort	9	600	2.8	2.5	**65**	10	6
Systemic Lupus International Collaborating Clinics (SLICC) Inception Cohort	25	1,837	5.3	0.3	**54**	16.8	13.8
Tromsø Lupus cohort	3	158	6	1.26	**35**	≥ 3.2	≥ 3.2
Lupus-Cruces Cohort	3	284	2	1.13	**25**	≥ 0.7	≥ 0.7
Danish National Registry of Patients	3	2,211	0	0	**0**	≥ 4	≥ 4
Duke University Medical Center	3	408	0	0	**0**	48	0
Euro-Lupus Cohort	4	1,000	0	0	**0**	1.9	≥ 1
Genetic Profile Predicting the Phenotype (PROFILE) Cohort	7	2,228	0	1.07	**0**	35.7	≥ 1.8
Georgians Organized Against Lupus (GOAL) Cohort	5	751	0[a]	0[b]	**0**	79.2	≥ 2.2
Instituto Nacional de la Nutricion Salvador Zubiran	3	667	0	0	**0**	Unknown	Unknown
Lupus in Minorities: Nature vs. Nurture (LUMINA) Cohort	79	643	0[c]	0.71	**0**	37.3	0
Lupus Outcomes Study Cohort	27	1,204	0[a]	0[b]	**0**	11.6	11.2
Pittsburgh Lupus Registry	4	983	0[c]	1.2	**0**	≥ 15.6	Near 0
Swedish National Databases (including MigMed database, Swedish National Patient Register and Hospital Discharge Register)	8	7,624	0	0	**0**	Unknown	Unknown
University College Hospital London Lupus Clinic	15	600	0[d]	0	**0**	20.5	4.6
University of Pittsburgh Lupus Cohort	3	1,327	2	0[e]	**0**	16.3	Near 0

SDI = Systemic Lupus International Collaborating Clinics SLE Damage Index; SLE = systemic lupus erythematosus; SLEDAI = Systemic Lupus Disease Activity Index

* Cohort receives 0 if the BLISS-compatible outcome is not used.

† Square root of the product of the *n*, SLEDAI and SDI values for the cohort.

[a] Uses Systemic Lupus Activity Questionnaire (SLAQ) instrument.

[b] Uses Brief Index of Lupus Damage (BILD) instrument.

[c] Uses Systemic Lupus Erythematosus Activity Measure (SLAM) instrument.

[d] Uses British Isles Lupus Assessment Group disease activity index (BILAG) instrument.

[e] SDI first recorded in 2008; not yet reported.

c) **As this analysis used data from participants from the US and Canada please explain how the results can be applied to UK clinical practice.**

Response: The BLISS-76 US LTE was a multicentre continuation trial of the Phase 3 BLISS-76 trial in the US. Patients who completed BLISS-76 were eligible. The baseline characteristics of the population in BLISS US LTE are presented in the CS in Table 17. The results of the BLISS-76 LTE are still considered generalisable to UK practice as part of complete body of evidence supporting the use of belimumab in UK.

Prior to matching, the selection of the TLC was based on the size of the cohort, the severity of the cohort and compatibility of cohort outcomes compared with the BLISS LTE outcomes. Prior to matching, the LTE and TLC samples were not well balanced; the percent bias is more than 10% for most variables (mean bias of 40%). After matching, the propensity-matched samples of 99 LTE and 99 TLC patients were well balanced; bias is less than 5% for 9 of the 17 variables and less than 10% for all variables (the mean bias is 4.6%).

The matched cohort represents the balance between the treated and untreated patient groups on those observed clinical characteristics. It cannot stochastically balance unknown variables. A strength of using the TLC was that patients who were otherwise indicated for treatment did not receive belimumab simply because it was not available, and not due to any other clinical considerations. A further reinforcement of the generalisability to a UK population, is that when the BLISS pooled cohort (BLISS 76 and 56 i.e. US and non-US) / TLC PSM comparative analysis was conducted the results were similar to those based on BLISS US LTE / TLC cohort (Urowitz et al., 2020).

d) **Please provide further details of the systematic review methods used to inform the PSM analysis (as specified in Urowitz 2019) to include the search strategy (search terms, databases and search dates), details of the excluded studies and why they were not eligible for the matching analysis.**

Response: A systematic review of the literature to identify SLE cohorts and registries was conducted. The search was performed with uncommon breadth because the nature of the SLE research was not considered relevant, only the

characteristics of the cohort on which it was conducted. Focus areas for the search included: systemic lupus erythematosus, the word “cohort”, the word “registry”, human research only and English language.

Although databases might possibly have been of interest, databases are primarily retrospective in nature and therefore they were not specifically sought out.

Data Sources

The search was limited to PubMed due to the breadth of the search. PubMed is a service of the U.S. National Library of Medicine that includes over 21 million citations from MEDLINE and other life science journals for biomedical articles back to 1948. Meta-analyses and systematic reviews found by the PubMed search were reviewed for additional publications.

Literature Search Process

Management of Acquired References: Results of the PubMed search were exported to an XML file and imported into Excel using PubMed2XL. This process produced an Excel file with the following columns: PubMed ID, Title, Authors, Year, Journal, Volume, Issue, Page and Abstract. Additional columns were added for management purposes: Sequential publication number, Reason for exclusion (1st level), Cohort name, Reason for exclusion (2nd level) and Comment.

Inclusion and Exclusion Criteria

The purpose of the literature search was to identify cohorts that might have sufficient, long-term data to act as a ST comparator to the active treatment arm of belimumab LTE trials as well as provide data to develop a natural history of SLE.

Table 8. Inclusion and exclusion criteria for the SLR

Inclusion Criteria	Exclusion criteria

When reviewing the literature search results, the following inclusion criteria were used to identify relevant articles for inclusion:
• Adult SLE
• Longitudinal study
• Known cohort or at least 400 patients in the study
	When reviewing the literature search results, the following exclusion criteria were used to determine that articles were irrelevant:
	• Opinion-based – articles including, but not limited to, editorials, comments, non-systematic reviews, and letters
	• Not SLE
	• Perinatal, pediatric or juvenile SLE
	• Not only or primarily SLE (i.e., included other autoimmune diseases)
	• Predominately Asian patients (research taking place in eastern Asia)
	• Adhoc or short-term data collection
	  ○ Cross-sectional study
	  ○ Case-control study
	  ○ Epidemiological study
	• Narrow focus of data
	  ○ Single-focus study
	  ○ Genetics/biomarker study/cohort
	• Small sample
	  ○ Small sample
	  ○ Single hospital/clinic

Reference Processing

References processing was performed in two passes. In the first pass the titles and abstracts of all publications were read from Excel. Those found relevant were coded green, retrieved, entered into Zotero and the cohort name (if given) was entered into the Excel sheet. Those whose relevance could not be determined from the title and abstract were coded yellow (possibly relevant), retrieved and entered into Zotero. Those found irrelevant were coded red and the exclusion reason was entered into the Excel sheet. However, studies that we found irrelevant but with patient populations of 400 or more were still treated as possibly relevant when the exclusion reason was not one of the following: opinion based; not SLE; not only or primarily SLE; prenatal, paediatric or juvenile SLE; and predominately Asian patients.

In the second processing pass relevant publications lacking a cohort name and possibly relevant publications were read to find the cohort/institution name. Cohort/institution names were recorded in the Excel sheet if relevant. If a possibly relevant publication was found relevant it was coded green and the name of the cohort/institution was recorded. If it was found irrelevant is was coded red and the exclusion reason was recorded. In addition, the reference lists of meta-analyses and systematic reviews were checked for publications missing from the PubMed results. Any additional publications were retrieved, read and their relevance determined. Relevant additional publications were appended to the list of publications in Excel, entered into Zotero, coded blue and the cohort/institution name was recorded. The Excel list was then sorted by cohort name. A list of cohorts was created with the number of publications found for each cohort.

Data Extraction

A data extraction form was created and reviewers were assigned cohorts to retrieve information from the identified cohort publications. Data were extracted for all cohorts with 3 or more publications.

Summary of search terms and hits

Table 9. PubMed search terms and numbers of hits.

Search	Query	Hits
#1	Search **systemic lupus erythematosus [MeSH]**	60,408
#2	Search **cohort**	381,286
#3	Search **registry**	109,966
#4	Search **#1 and (#2 or #3)**	2,603
#5	Search **#1 and (#2 or #3)** Filters: **Other Animals**	57
#6	Search **#1 and (#2 or #3)** Filters: **Randomized Controlled Trial**	20
#7	Search **#4 not #5**	2,546
#8	Search **#7 not #6** Filters: **Publication date from 1995/01/01 to 2016/12/31; English**	2,362

[MeSH]: medical subject headings database, the National Library of Medicine controlled vocabulary thesaurus used for indexing PubMed citations.

Processing

The PRISMA diagram documents the processing of the publications. In all, 21 cohorts/databases received further attention.

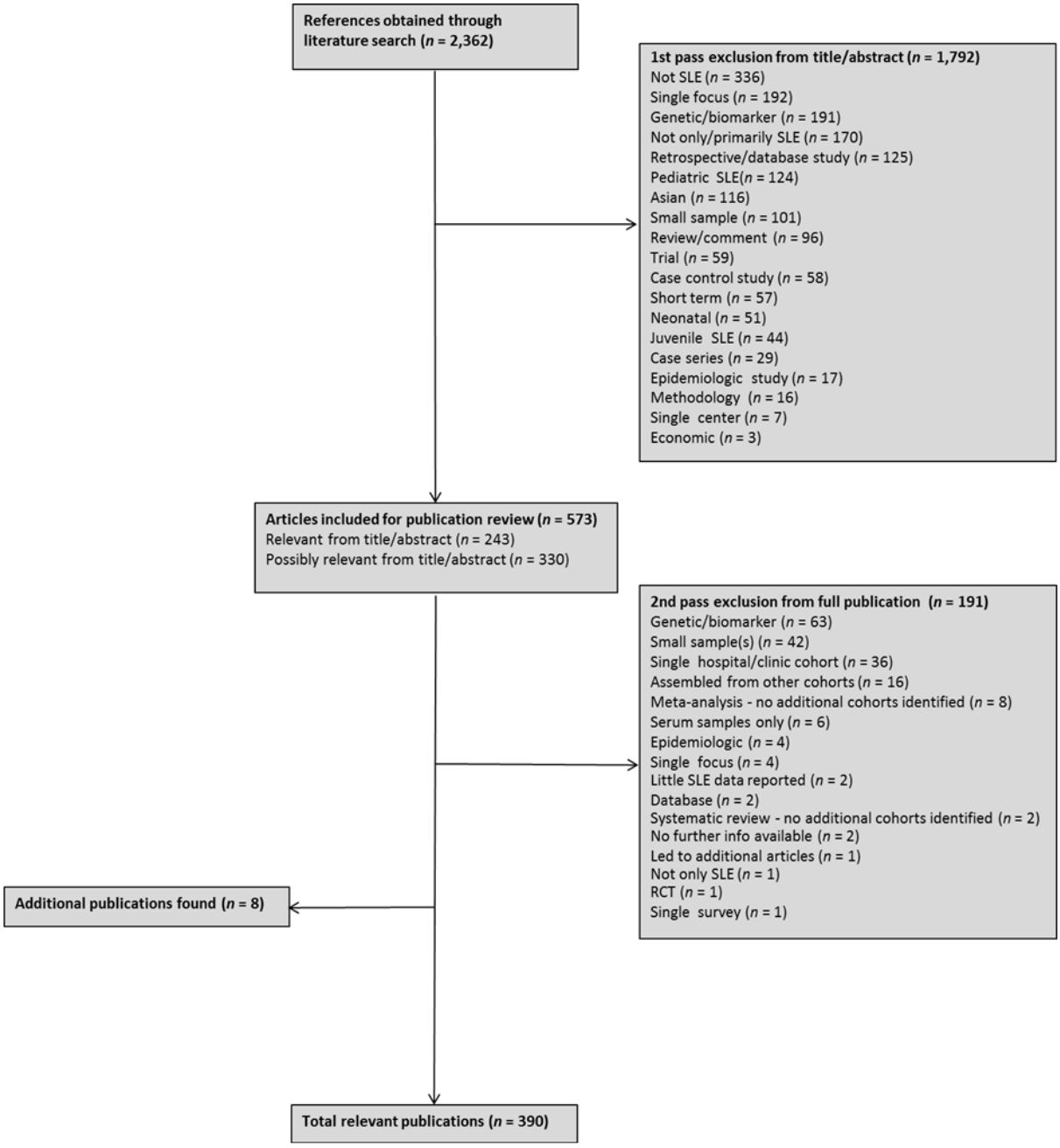


Figure 7. PRISMA diagram

The identified cohorts and number of publications are provided as an appendix to this document (IB1591 A38-e. Clarifications - Identification and Selection of cohort for PSM).

e) **Please explain how the results of the PSM analysis apply to the HDA-1 and HDA-2 populations.**

Response: The PSM comparative analysis shows that over 5 years the increase in SDI from baseline is statistically significantly lower in those SLE patients maintained on add-on belimumab IV compared with ST only. The results are consistent based on the primary analysis undertaken with BLISS-76 US LTE and the exploratory analysis which includes the pooled LTE studies (see earlier response to A38). The analysis is conducted on a matched population of the ITT population and therefore generally has less active disease than the HDA-1 and HDA-2 populations.

A PSM-equivalent analysis is not available for the HDA-1 and HDA-2 populations. It is likely that the matched populations would be of limited size based on the proportion of BLISS ITT patients meeting the criteria for HDA-1 and HDA-2. We have consulted with two Consultant Rheumatologists (and lupus experts) who have advised that they would expect a high disease activity subgroup to sustain greater damage over time, and hence see a greater impact of belimumab in preventing damage accumulation.

From a sub-analysis of the BLISS pivotal trials across IV and SC, moving from the ITT population to a more active disease state as defined by HDA-1 and HDA-2, shows that add-on belimumab has a greater treatment effect. The size of the treatment effect, measured by the SELENA-SLEDAI score is consistent with the, albeit limited, UK real world usage of belimumab as captured in the BILAG-BR.

We recognise the limitations of using the results of the PSM analysis to calibrate the model to inform on the longer-term organ damage accrual. These limitations have been taken into consideration in the way in which we have applied this benefit:

- The improvement in organ damage accrual is applied to the belimumab arm only, despite the model validation suggesting that that organ damage accrual for those on ST was underestimated.

- The improvement in organ damage accrual is applied to the belimumab arm for 5 years only, from 1.5 years to 6.5 years, despite there being longer-term

data to support continued treatment effect out to 13 years (Wallace et al., 2019).

A39. The BILAG-BR is stated to include data over a 3-5-year period, while the OBSErve registry was based over the course of 2 years. Please specify the years these periods cover from these registries.

Response: The BILAG-BR was set up in March 2010, but data for belimumab have only been collected since it became available on the NHS in June 2016. The data cut for the BILAG-BR analysis was 7th July 2020; therefore, the data collection period for belimumab from BILAG-BR was June 2016–July 2020.

For the OBSErve study series, data collection periods were as follows:

- US: primary data collection period was February 2012–May 2014. The augment laboratory test component of data collection occurred between January and August 2014.
- Spain: December 2013–February 2014
- Canada: December 2012–January 2015
- Germany: April 2013–November 2013
- Switzerland: December 2014–February 2017
- Argentina: March 2014–March 2016

Ongoing trials

A40. In Section B.2.11 of the CS (Ongoing studies) it is stated that primary analysis of the BASE study is now complete. Could you supply these data in full? Are data from any other ongoing trials available?

The BASE clinical study report has been provided in answer to question C1. Similarly, Part A of the PLUTO trial (52 week double-blind phase) has also been provided. All other studies are on-going, and the reports are expected in line with the dates presented in section B.2.11.

A41. Have all data collected so far from the PLUTO trials been provided in the CS? When will final results be available for the SC and IV trials?

Response: Available data from PLUTO (Part A) has been summarised in Appendix O and the CSR is now provided as part of the reference pack for Appendices. As stated in the CS, the final study report from PLUTO is expected in [redacted] and final study report from PLUTO-SC in [redacted].

Indirect comparison

A42. Priority question. The company rule out the use of the EXPLORER trial to perform an indirect comparison with rituximab at least partly on the basis of the population of the EXPLORER trial being of more severe disease. However, a comparison of baseline characteristics between the BLISS and EXPLORER trials reveals that, although the percentage with at least one A score is higher in EXPLORER, a substantial percentage of BLISS patients had experienced at least one A score (at least 13.9%, which was the value for belimumab in the pooled BLISS-52 and BLISS-76 data). The company also report that SELENA-SLEDAI was not employed in EXPLORER. However, BILAG was employed in both EXPLORER and the BLISS trials. Given that rituximab is a comparator in the scope and the high risk of selection bias in the BILAG-BR registry:

a) **Could the company provide a more detailed comparison of baseline characteristics and outcomes between the EXPLORER and each of the BLISS trials.**

Response: Baseline characteristics in both EXPLORER and the BLISS trials were consistent with the patterns of SLE incidence, in that all trials included predominantly females of working age. The ethnic distribution differed somewhat between EXPLORER and the BLISS trials – while white patients predominated in all trials, the proportion of Asian patients was higher and of black patients was lower in the BLISS trials than in EXPLORER.

In terms of disease characteristics, mean SLE disease duration was longer in EXPLORER than in the BLISS trials. The patients in the EXPLORER trial had significant and acute disease activity at entry to the study; 53% had at least one BILAG A score (severe disease activity) and a further 28% had at least 3 BILAG B

scores (please note that although a BILAG B score represents moderate disease activity, the presence of 3 BILAG B scores in some organs indicates more severe disease activity). Initially, patients were receiving very high daily doses of prednisone (mean 45.9 mg ±16.4 mg) to treat the significant level of disease activity and this dose was to be tapered where possible during the trial. In addition, all patients were receiving one immunosuppressant at study entry. In contrast, the patients in the BLISS studies were a broader population and not all patients were experiencing major disease flares, receiving an immunosuppressant or requiring high doses of steroids at baseline as seen in the EXPLORER trial. Key baseline characteristics that are available for both the BLISS trials and EXPLORER are tabulated below. The proportion of patients with at least 1 BILAG A score was approximately 3–4 times higher and the average prednisone dose at baseline was approximately 4 times higher in EXPLORER than the BLISS trials.

	BLISS-SC		Pooled BLISS-52 and BLISS-76 data		EXPLORER[9]	
	Belimumab 200 mg SC N=556	Placebo N=280	Belimumab 10 mg/kg IV N=563	Placebo N=562	Rituximab (n = 169)	Placebo (n = 88)
Baseline characteristics						
Age (years), mean (SD)	38.1 (12.10)	39.6 (12.61)	37.9 (11.3)	38.1 (12.0)	40.2 (11.4)	40.5 (12.8)
Female, N (%)	521 (93.7)	268 (95.7)	539 (95.7)	522 (92.9)	82 (89.9)	152 (93.2)
Race, N (%)*						
White	336 (60.4)	166 (59.3)	260 (46.2)	270 (48.0)	95 (56.2)	49 (55.7)
African American	56 (10.1)	30 (10.7)	50 (8.9)	50 (8.9)	40 (23.7)	24 (27.3)
Asian	119 (21.4)	63 (22.5)	127 (22.6)	116 (20.6)	6 (3.6)	5 (5.7)
SLE disease duration (years), mean (SD)	6.4 (6.60)	6.8 (6.83)	6.08 (6.42)	6.66 (6.48)	8.5 (7.2)	8.7 (7.6)
BILAG ≥1A, n (%)	87 (15.6)	51 (18.2)	78 (13.9)	89 (15.8)	86 (51.0)	49 (56.0)
Average daily prednisone dose (mg/day), mean (SD)	10.8 (8.21)	11.2 (9.09)	10.9 (9.1)	10.7 (8.5)	45 (16.4) across all patients, not available by trial arm	
Immunosuppressant use, n (%)	244 (43.9)	137 (48.9)	271 (48.1)	276 (49.1)	ALL patients (inclusion criterion)	
*Only the most common race groups across all trials are presented so that the percentages do not add up to 100%						

Table 10. Baseline patient characteristics for BLISS-SC, Pooled BLISS-52 and BLISS-76 data, and the EXPLORER trial

No outcomes were tabulated because although both the BLISS trials and EXPLORER employed the BILAG tool, the specific endpoints assessed and the way in which the tool was applied differed. Further details on the lack of compatibility between the trials are provided below.

b) Could the company please perform an indirect comparison with any outcomes at any time point that are common to any of and all BLISS trials and EXPLORER.

Response: No indirect treatment comparison has been conducted. As per the previous submission (TA397) we believe the differences in the endpoints and trial design preclude any meaningful indirect comparison between rituximab and belimumab. For example:

- In the BLISS studies, the primary efficacy end point was the SRI-4 response rate at week 52 which was a composite endpoint defined by a reduction of at

least 4 points in SS, no new BILAG A domain score, no more than 1 new BILAG B organ domain score and no worsening in PGA compared to baseline. The primary endpoint in the EXPLORER trial was the effect of placebo versus rituximab in achieving and maintaining a major clinical response, a partial clinical response, or no clinical response at week 52 was assessed using each of the 8 BILAG index organ system scores. Whilst GSK acknowledges BILAG index has been used in both trials, the application across the studies was different i.e. the BLISS trials used the BILAG index as a worsening endpoint whereas the EXPLORER trial used it as an improvement endpoint.

- In the EXPLORER trial patients were randomised at a 2:1 ratio to receive intravenous rituximab or placebo, which was added to prednisone and to the baseline immunosuppressive regimen. After screening, eligible patients continued their immunosuppressant therapy and received additional daily oral prednisone, based on the BILAG score at entry and the amount of steroids already being taken at the time of entry. Steroids were tapered beginning on day 16, with the goal of reaching a dosage of ≤10 mg/day over 10 weeks and ≤5 mg/day by week 52 (see Figure 1A from Merrill et al.,[9] presenting EXPLORER study design, reproduced below)

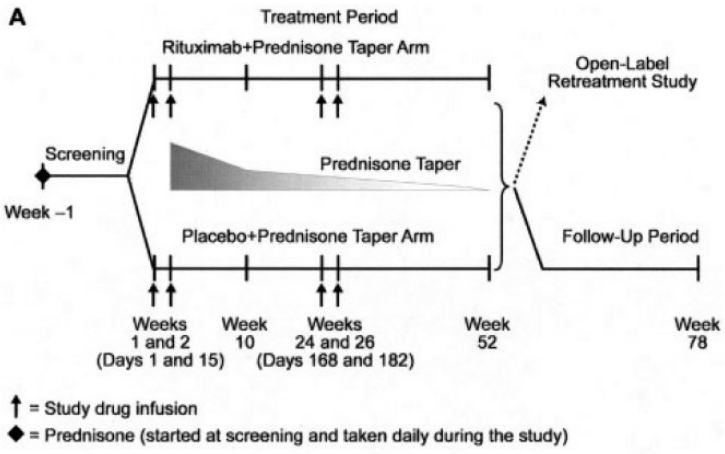


Figure 8. EXPLORER study design

- In the BLISS trials, patients were randomised either
    - in a 1:1:1 ratio to placebo, or belimumab 1 mg/kg or 10 mg/kg, plus standard therapy (BLISS-52 and BLISS-76)

or

- in a 1:2 ratio to placebo or belimumab 200mg SC (BLISS-SC).

At entry patients were required to have a stable treatment regimen with fixed doses of prednisone (0–40 mg/day). Therefore, we believe that the differences in the use of steroids to manage disease activity between the EXPLORER and BLISS trials and consequently the differences in the type of response observed in the placebo arms do not support an indirect comparison; in the BLISS studies, changes in the total dose of systemic steroids was only permitted during the first 6 months of the trial with no defined treatment tapering regimen through to week 52, as was employed in the EXPLORER trial.

- More importantly, despite differences in study design and trial endpoints such as those outlined above, the EXPLORER trial did not meet its primary or secondary efficacy endpoints

c) In order to mitigate differences in patient characteristics between EXPLORER and the BLISS trials, could the company consider 1 or more of the following methods: subgroup analysis, statistical adjustment as described in the NICE Decision Support Unit Technical Support Document (TSD) 17.

Response: Please see responses to parts a) and b) of this question to understand why a comparison was not conducted.

Section B: Clarification on cost-effectiveness data

B1. Literature searches:

a) Appendix G describes the search strategy applied to identify published cost-effectiveness studies for belimumab. From the document it is unclear how search terms were identified and which rationale was used to identify the selected papers given the large numbers of citations found. Please reflect on the rationale behind the search strategy?

Response: A broad search strategy was employed across all searched databases to ensure that any relevant cost-effectiveness studies would be appropriately captured. No published cost-effectiveness studies were identified as part of searches included in TA397.

b) Please describe whether and how identified studies were used.

Response: Of the three identified published belimumab cost-effectiveness studies, two reported on the same model and modelled population used as part of a health technology assessment (HTA) performed for Italian HTA authorities. The microsimulation model used in these studies was derived from the same model structure provided to NICE as part of TA397. The final included publication was a decision summary authored by the Canadian Agency for Drugs and Technologies in Health (CADTH). This HTA decision presented only limited details of how the economic analyses were performed. No new data or information relevant to the current NICE appraisal was identified from any of the three identified published belimumab cost-effectiveness studies and so were not included any further in our current economic analysis.

B2**.** Population:

a) Could the company clarify whether the population modelled in the cost effectiveness analyses (including baseline characteristics, efficacy estimates from the BLISS trial programme and other evidence sources) is based solely on an adult population or whether they include paediatric patients?

Response: This was communicated at both the Scoping process and Decision Problem Meeting. The marketing authorisation for belimumab IV formulation includes

patients aged 5 - 17 years old, based on one Phase 2 RCT, the PLUTO Study (BEL 114055). Due to the rarity of paediatric SLE, a statistically powered study was not considered feasible. The Phase 2 study was therefore exploratory and not powered to show a statistical difference between treatment groups. It was designed to descriptively evaluate the efficacy and safety of belimumab in paediatric SLE subjects (a total of 93 subjects; 40 in the placebo group and 53 in the belimumab group). Therefore, due to the limited belimumab data in paediatric SLE patients, particularly in our HDA subgroups, GSK focuses only on the adult population for this appraisal. The population modelled in the cost effectiveness analysis is based solely on an adult population and did not include paediatric patients.

b) Please perform a full economic analysis in children using results from the PLUTO trial.

Response: This has not been conducted for reasons provided above in part a).

Comparator

B3. Priority question. The scope includes rituximab as a relevant comparator. Observational data is available for rituximab and trial data from EXPLORER. Please provide a cost effectiveness analysis including rituximab as a comparator (as well as standard therapy), possibly based on analyses presented in Appendix P or using an indirect comparison with EXPLORER, as requested in question A42. If long-term outcomes are not available for rituximab, these can be explored by making assumptions, such as equivalent outcomes to belimumab (given that the company considers short-term effectiveness estimates to be similar according to the CS Section B.2.9) Indirect and mixed treatment comparisons.

Response: Please see the company response to question A42. We have not provided a cost-effectiveness analysis which includes rituximab as a comparator.

As shown in the response to question A2, the patient pathway presented in the Clinical Commissioning Policy Rituximab for refractory Systemic Lupus Erythematosus (SLE) in adults and post-pubescent children [200402P][(2)] clearly

shows that patients who are eligible for belimumab should be considered for treatment with belimumab before rituximab.

Furthermore, the eligibility criteria for rituximab on page 6 of the Clinical Commissioning Policy[2] (quoted on page 103 of the company submission) state that patients should have been assessed as ineligible for belimumab in order to be considered for treatment with rituximab.

The GSK economic model relies on the SELENA-SELEDAI (SS), a measure captured within all the pivotal BLISS trials, as the main driver for tracking disease progression and enforcing NICE mandated treatment discontinuation rules. As SS is not captured within the EXPLORER trial, the outcomes would not conform to the current model structure.

Rituximab remains an unlicensed treatment in SLE and there is a lack of regimen standardisation. Therefore, understanding the timing of a re-treatment is unknown and unable to be modelled for. The reason for treating with belimumab and its impact on the disease is different to rituximab and this also makes comparison in a microsimulation model challenging. Belimumab is provided as an ongoing maintenance treatment (and the model reflects this) impacting on disease progression whereas rituximab is generally prescribed on a flare of disease activity and aims to supress B cells rather than impacting disease progression.

It is unknown whether reactive (based on B-cell depletion) treatment with rituximab i.e. treating flares as and when they occur, has a detrimental effect on long-term organ damage, disease progression, or may cause other health problems in SLE patients, especially if treatment with rituximab is accompanied by high-dose corticosteroids. In addition, elevated serum Blood B-Cell Activating Factor (BAFF) levels are associated with rising anti-double-stranded DNA antibody levels and may drive a disease flare after B cell repopulation following B cell depletion therapy in SLE[10] . Therefore, GSK believes that making equivalence assumptions about the long-term outcomes of rituximab is inappropriate and would approximate to a simple cost-minimisation exercise.

Model validation

B4. Priority question. It is stated that a validation of the base case results against the predicted results (e.g. comparison of mortality to mortality observed in JH cohort) was performed.

a) **In addition to above, please provide external validation of modelled events and outcomes (mortality, long-term disease activity with belimumab and ST, and treatment continuation) against data from other sources (LTE studies or observational studies).**

Response:

The original health economic model, beyond 1 year, was developed on the Johns Hopkins cohort. In order to assess the internal and external validity of predicted organ damage occurrences, two analyses were performed; the internal validation compared the model outcomes with the JH cohort, and the external validation with the Toronto cohort.

Baseline demographics and baseline organ damage from the two cohorts were imputed in the health economic model before 50,000 patients were simulated. Organ damage occurrence was presented as Kaplan-Meier (K-M) estimates at 1, 5 and 10 years and compared to the actual damage prevalence in the cohorts in scatter plots.

The comparison between model prediction and JH damage occurrence is presented in Figure 1. Most predicted values are very close to the actual occurrences, with some slight variation at higher ages and larger occurrences of damage.

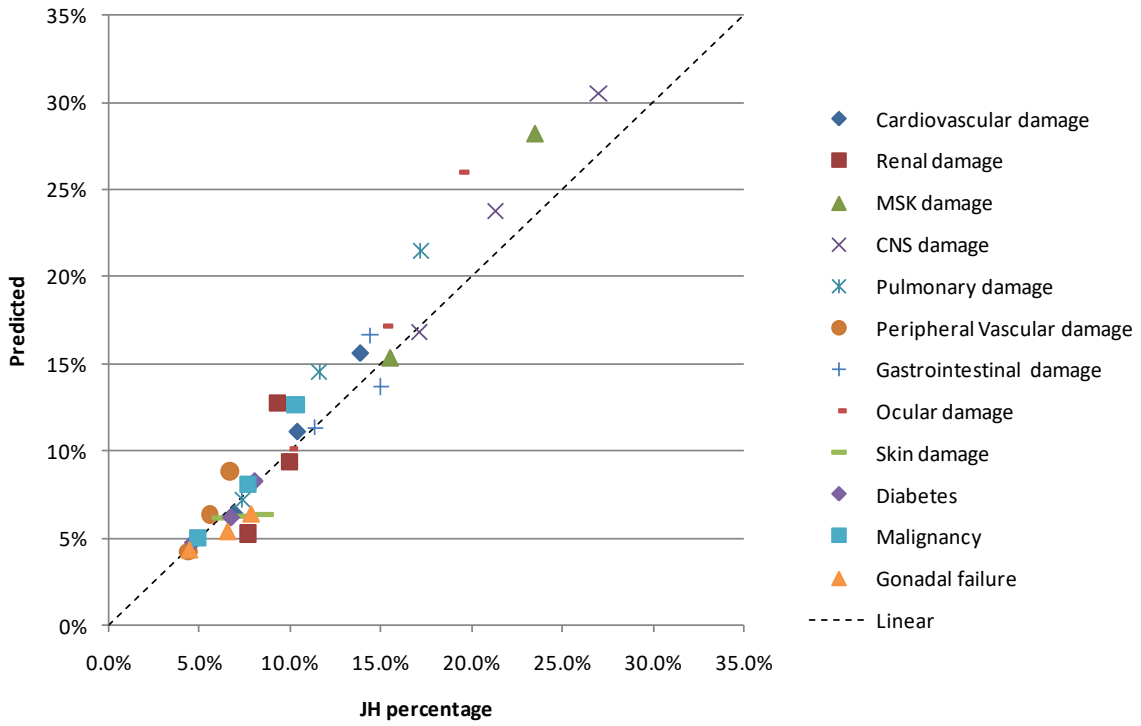


Figure 9. Scatter plot of organ damage in JH cohort vs predicted damage at 1, 5 and 10 years (internal validation)

The external validation is presented in Figure 2 and shows reasonable predictions for most organs. However, pulmonary malignancy and CNS damage, tend to be over predicted, whereas skin damage is slightly under predicted.

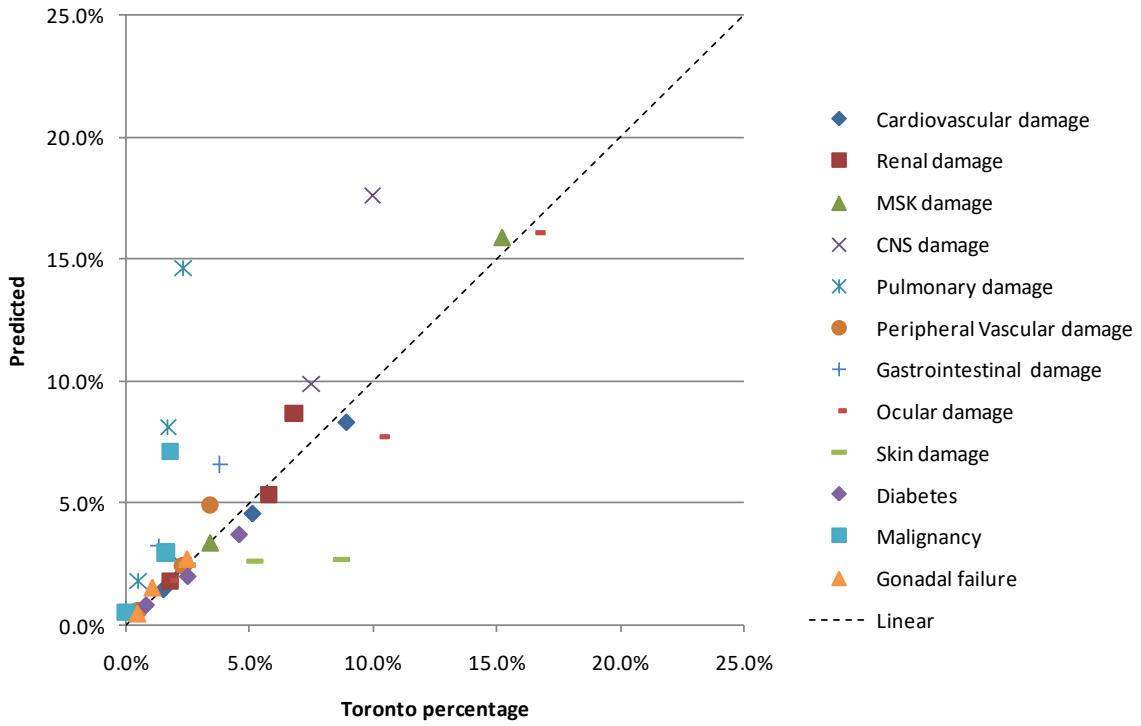


Figure 10. Scatter plot of organ damage in Toronto cohort vs predicted damage at 1,5 and 10 years (external validation)

The internal validity of the model predictions is positive. Predicted values are generally close to the actual values observed in the JH cohort. The external validation showed more errors in organ damage occurrences than the internal validation, especially in CNS damage, pulmonary damage and malignancy.

The model also underwent review by three external reviewers. The trial analysis has been validated both by an external vendor and internally by GSK.

As the current model is adapted from the original model and remains fundamentally unchanged, no further validation exercises were undertaken on unchanged elements of the original model.

b) To ensure internal validity of the model, please complete the TECH-VER checklist which is a verification checklist to reduce errors in models and improve their credibility (see: *Büyükkaramikli, N. C., Rutten-van Mölken, M. P., Severens, J. L., & Al, M. (2019). TECH-VER:*

***A verification checklist to reduce errors in models and improve their credibility. Pharmacoeconomics, 37(11), 1391-1408*).**

Response: Other than the addition of calibration factors to align the model with further collected data on long-term organ damage (between only 1.5-6.5 years for belimumab only), the model remains identical to the model that was previously submitted to NICE as part of TA397. Please note the previous ERG comment with regards to the model in the Final Appraisal Document provided as part of TA397:

"The ERG considered that the manufacturer's model was complex, though generally well constructed. It noted that the model conformed to the NICE reference case and that the longer-term effects of systemic lupus erythematosus had been modelled well, using the Johns Hopkins cohort".

c) **In addition, please provide a model file that enables creation of the following output items for each simulated patient: summary of baseline characteristics, treatment duration, disease activity over time, time at which events happen, or alternatively, provide a table overview of this detail from 1 model run.**

Response: Please see the 'PatientLog1' and 'PatientLog2' sheets of the file 'ID1591 B04-c. Clarification - Patient Profiles model file'. Please note a tab is included called 'Plausibility check' for checking the plausibility of generated patient profiles.

B5. Priority question. Considering differences between this submission and the previous TA397 submission, please provide:

a) **a detailed overview regarding model structure**

Response: The microsimulation model structure fundamentally remains unchanged since TA397 and is presented in the figure below.

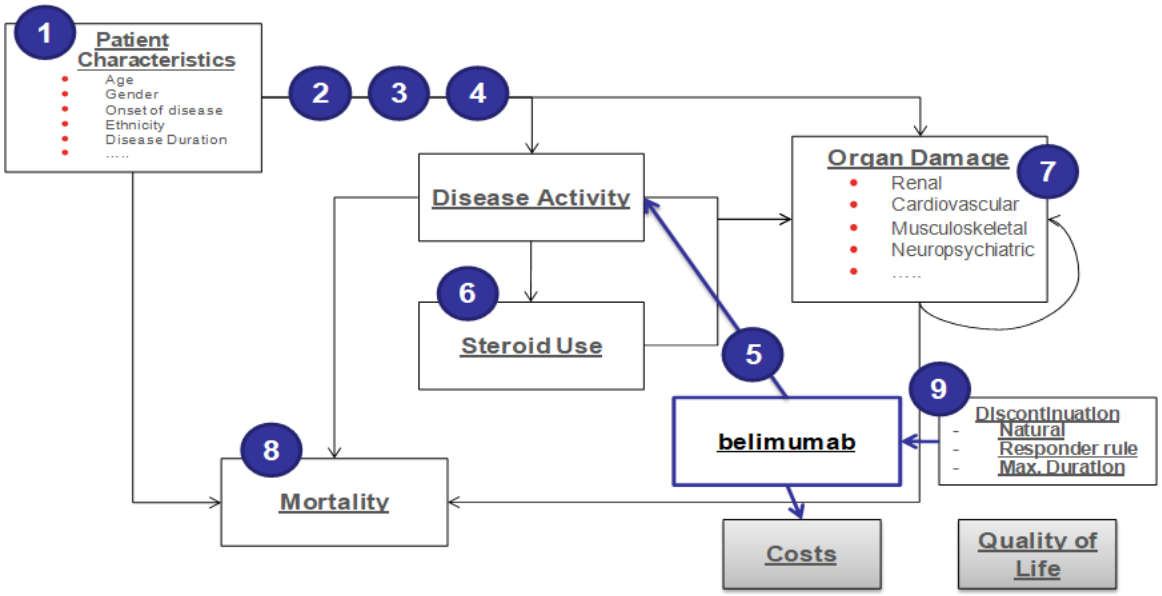


Figure 11. Figure showing Schematic overview of interdependencies between baseline characteristics, treatment and outcomes in the micro-simulation model (presented in TA397)

1. Simulation of a patient: Baseline characteristics are sampled from the baseline characteristics of the relevant population in the BLISS trials (by formulation).
2. Response at 24 weeks (defined as a decrease in SS score of ≥4 points after 24 weeks): Determined from the probabilities of response in the BLISS trials, stratified by baseline SS score.
3. Disease activity in the first year: A regression model produced from BLISS trial data to explain the change in SS score after 52 weeks, based on treatment, baseline SS score and SS score response at 24 weeks (yes or no).
4. Disease activity over time: SS score over time (after the first year) for a standard therapy (ST) patient is determined with a statistical model developed using the Johns Hopkins cohort longitudinal database.
5. Effect of belimumab on SS score: The regression model for SS score at 52 weeks (3) is used to determine the difference between a ST and a belimumab patient. This is subtracted from the disease activity over time (4). A patient discontinuing belimumab treatment returns to ST disease activity levels.
6. Steroid use: determined by a model developed on the Johns Hopkins cohort. The model explains steroid use at a time point based on the average disease activity in the last year.
7. Organ damage: The Johns Hopkins database was also used to estimate the time to organ damage outcomes. Yearly organ damage probabilities are calculated based on patient characteristics, disease activity (adjusted [average] mean SLEDAI [AMS]) and steroid use. A propensity score matched comparative analysis has since provided an estimate of the long-term reduction in SDI for patients on add-on belimumab compared with a matched cohort (Toronto Lupus Cohort) on ST. In the current appraisal, we incorporate the findings from the propensity score matched analysis to model the long term organ damage reduction treatment effect shown by belimumab (see Section B.3.3.6), by means of a calibration factor.
8. Mortality: yearly mortality risk is calculated by combining average population life tables with an increased mortality in SLE patients and a statistical model

explaining the influence of patient characteristics, disease activity and organ damage on mortality.

b) a detailed description of any change of inputs and how these changes affected outcomes

Response: The following table shows the impact of reverting 8 key parameter input updates on the Benlysta IV model for the HDA-1 patient subgroup provided as part of the current submission, to values used in final base case of TA397. To demonstrate the effects of each parameter update, only a single parameter at a time was reverted to values used in the model provided as part of TA397, and all other parameters were held as they are used in the current submission.

The key findings of this exercise are that removal of the calibration factor and reverting to annual long-term organ damage costs provided as part of the previous submission result in worse ICER's for belimumab as compared to standard therapy than the base case presented in this submission. In other words, if we were to start with the previous model and update only these parameters to values from the current submission, the ICER would be improved for belimumab. Conversely, when parameters in the current model for SLEDAI annual costs, patient mortality table, organ damage disutilities, patient weight source, natural discontinuation rates and drug cost are reverted to values used in the previous model, the ICER's for belimumab as compared to standard therapy are improved.

Table 11. Impact of reverting 8 key parameter input updates on the Benlysta IV model for the HDA-1 patient subgroup provided as part of the current submission, to values used in final base case of TA397

Parameters	Base Case Value In Current Model	Previous Value applied from TA397	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
IV HDA-1 model (base)	-	-	[redacted]	[redacted]	[redacted]	£28,362
Calibration factors	Added	None	[redacted]	[redacted]	[redacted]	£48,604
Annual Long-Term Organ Damage costs	Updated literature search	Previous literature search	[redacted]	[redacted]	[redacted]	£37,219
SLEDAI Annual costs	HCHS inflation 2018/19 values	2009/10	[redacted]	[redacted]	[redacted]	£28,307
Patient mortality table	ONS 2016-18 values	ONS 2007-09 values	[redacted]	[redacted]	[redacted]	£28,235
Organ damage disutilities	Updated literature search	Previous literature search	[redacted]	[redacted]	[redacted]	£28,007
Patient weight source	BILAG Registry	Trial	[redacted]	[redacted]	[redacted]	£25,984
Natural discontinuation rates	[redacted]	Y1 8.0% / Y2 11.7%	[redacted]	[redacted]	[redacted]	£23,521
Drug cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

c) **a detailed overview of the TA397 committee's preferred assumptions and whether and how these were incorporated in this submission. Please provide justification in case of deviation from committee's preferred assumptions.**

Response: Where possible, the current model has incorporated the committee preferred assumptions from the TA397 Final Appraisal Determination.

- The committee preferred the use of an administration cost of £154 for Benlysta IV. This was incorporated into the final economic model submitted as part of TA397 continues to be used in the current IV model.

- The committee did not consider that a maximum treatment duration of 6 years could be considered robust for decision making. A maximum treatment duration of patient lifetime was incorporated into the final economic model submitted as part of TA397 continues to be used in the current IV and SC models.

- The committee preferred the use of natural discontinuation rates of 8% in the first year and 11.7% in subsequent years. These values have been superseded by an integrated analysis captured over 13 years (unavailable for the previous submission), which forms part of our current base case.

d) **a model version in which it is possible to reproduce the original TA397 base-case analysis (or a description of how to do this in the current model file).**

Response: All changes required to reproduce the original TA397 base-case analysis from the IV model supplied as part of the current submission are identified in the company's response in section b) for this question. The cost of belimumab, patient weight source, and the use of calibration factors may be altered in the 'Scenario' tab of the economic model. Please see an included Excel file (ID1591 B05-c. Clarification - Parameters to recreate original model) that include previous parameters for annual Long-Term Organ Damage costs, SLEDAI Annual costs,

patient mortality tables, organ damage disutilities and natural discontinuation rates. Values from this Excel file may be used to overwrite values in the current model. All values used concurrently will reproduce results associated with the previous appraisal.

Model structure

B6. Given that the cycle length is annual:

a) please elaborate on how the 24-week response assessment and treatment discontinuation at this time point were incorporated in the model

Response: In the first year of the model, an additional assessment of a patient's status is conducted at week 24, to determine whether they are a responder (defined as a reduction of SELENA-SLEDAI score greater than or equal to 4). This determines whether a patient discontinues or not. For those who discontinue due to this rule, corresponding belimumab costs and effects are only calculated for the first 24 weeks of that year.

b) please also provide justification for the annual cycle length and reflect on the appropriateness and any potential biases introduced by this modelling choice.

Response: SLE is a long-term chronic disease. The changes in overall disease activity and the accumulation of organ damage are believed to be adequately captured with a yearly cycle over a lifetime horizon. If long-term data on the incidence and severity of flares had been available, a shorter cycle length may have been more appropriate to capture the pattern of flares over time.

B7. Priority question. The SELENA-SLEDAI (SS) score component of the composite SRI-4 endpoint at week 24 was used to model response.

a) Please comment on the appropriateness of the SS score as the only outcome determining response.

Response: Belimumab IV 10 mg/kg demonstrated superiority to SoC for the SRI-4 composite endpoint in both the BLISS-52 and BLISS-76 trials. In this composite endpoint, SS score is the measure of efficacy in terms of disease activity reduction whilst both BILAG and PGA are measured to ensure any observed improvement in

SS score is not reported as a response if accompanied by a worsening of the disease in another organ system or in the general well-being of the patient. Since there was no long-term cohort data in which all the three measures of the composite endpoint were recorded, determining the long-term effects of the SRI were not possible. The disease activity score itself (i.e. SS score) however, has been shown to be predictive of organ damage and mortality (Ibanez et al. 2003). As such, for the purpose of the health-economic model, the SS score alone was deemed more appropriate to link with long-term outcomes; it was part of the composite SRI endpoint; is the measure of efficacy within that endpoint; and is the primary driver of the SRI response in the BLISS trials.

b) Please clarify what SS score patients who discontinue revert to (for each reason for discontinuation and for each comparator separately).

Response: There is no discontinuation on standard therapy alone. Patients may only discontinue belimumab in the economic model due to no longer deriving treatment benefit. Patients who discontinue Benlysta due to any reason revert to being treated with standard therapy alone, and therefore assume the average SELENA-SELDAI score associated with this comparator.

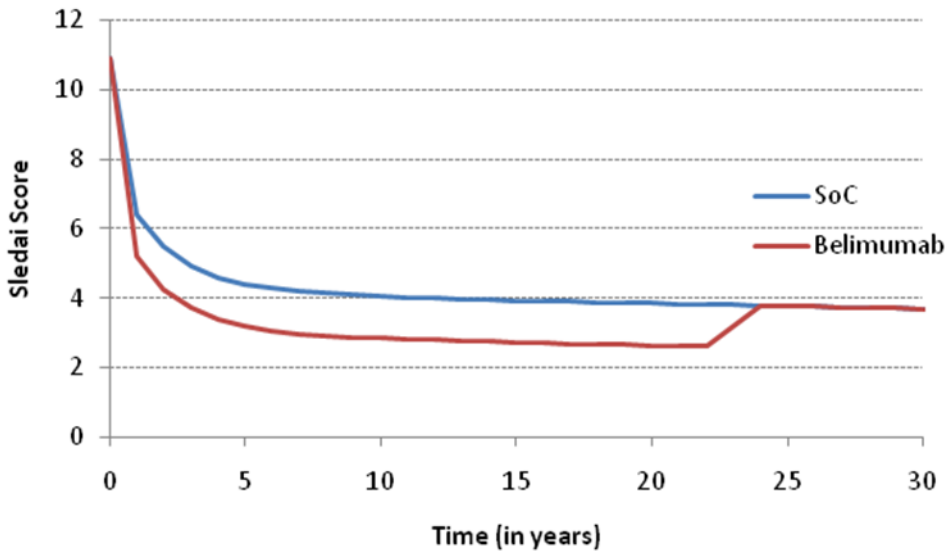


Figure 12. Example of SLEDAI score for a SOC patient and patient discontinuing belimumab treatment in year 23. It is assumed that discontinuation takes place in the middle of the year. The red curve does not go up immediately due to the fact that mean SS scores are only measured at integer time points (years).

Please also state how this differs from the original TA397 submission and whether the model choice now is in line with committee preferences at the time.

Response:

The model continues to process non responders on the belimumab arm in the same way as the original model provided as part of the TA397 submission.

Treatment effectiveness and extrapolation

B8. Priority question. In section B.3.3.3 of the CS, it is stated that the methodology used to determine a patient's change in SS score at week 52 is consistent with the previous submission. Compared to TA397, change in SS score for SoC is smaller [redacted] versus -0.390 in TA397, whereas the change in SS score is larger for SS_0 all belimumab ([redacted] versus -0.285).

a) **Please justify these differences. For example, why is SoC less effective than before? Can this difference be solely attributed to the change in population from HDA-1 to HDA-2?**

Response: The -0.390 value is the regression coefficient for standard therapy based on the total BLISS population whereas the -0.379 value is based on the HDA-2 population. For the HDA-1 population, the regression coefficient is -0.349. These differences are solely attributable to population differences.

b) **In TA397, an adjusted R2 was provided (0.699). Please provide the adjusted R2-values for the two models in Table 61 in the CS in order for the ERG to obtain a general idea regarding the fit to the data.**

Response: The adjusted R2 values for the corresponding regression for change in SS (as reported in the CS Table 61) are as follows:

Table 12. Adjusted R2 values for the IV and SC models in the HDA-2 population

	IV model - HDA-2	SC model – HDA-2
Adjusted R^2 values	[redacted]	[redacted]

c) The results in the HDA-1 population presented in appendix Q do not match the results presented in TA397 (Table 6.5). Please explain why these results differ and provide adjusted R2-values.

Response: Note the populations being compared here are different. Table 6.5 of the CS from TA397 presents the results of the regression for estimating the change in SS score at Week 52 for the pooled IV total population.

Table 2 in Appendix Q of the CS presents the regression for HDA-1 population (i.e. a sub-group to the total pooled population). The regression for IV in Table 2 match the values for the same regression in the same HDA-1 population in the previous submission TA397 Table 6.41 labelled '*Linear regression explaining change in SS score at week 52 – high disease activity group'*. Copied below for information:

Table 13. Linear regression explaining change in SELENA-SLEDAI (SS) score at week 52 – High disease activity (HDA-1) subgroup – Taken from CS TA397 - Table 6.41.

Parameter	Estimate	SE	p-value
SS0 SoC	-0.349	0.022	<0.001
SS0 all belimumab	-0.343	0.046	<0.001
SS0 belimumab responders	-0.280	0.052	<0.001

Table 14. Change in SELENA-SLEDAI score after 52 weeks compared to ST in HDA-1 - CS Appendix Q Table 2.

Parameter	IV model – HDA-1				SC model – HDA-1			
	Estimate	Std Error	t-value	p-value	Estimate	Std Error	t-value	p-value
SS_0 ST	-0.349	0.022	-15.919	0.000	-0.447	0.029	15.260	p<0.0001
SS_0 all belimumab	-0.343	0.046	-7.516	0.000	-0.262	0.046	-5.760	p<0.0001
SS_0 belimumab responders	-0.280	0.052	-5.410	0.000	-0.382	0.050	-7.680	p<0.0001

B9. Priority question. In section B.3.3.4.1 of the CS it is stated "Reasons for treatment discontinuation in the current submission remain consistent with the reasons for treatment discontinuation provided in the previous submission: natural discontinuation, and no longer deriving clinical benefit from treatment. (Please see Section 6.3.1 of the previous submission provided as part of TA397 for further details)". In TA397, the percentage of belimumab patients satisfying treatment continuation rule at 24 weeks was 52.4%. Please explain why this percentage is lower compared to the values mentioned in the CS (both for HDA-1 in appendix Q as well as HDA-2).

Response: The 52.4% value refers to the percentage of belimumab patients satisfying the treatment continuation rule at 24 weeks for the total pooled population. In the current submission, we present analyses for two high disease activity (HDA) sub-groups, HDA-1 (referred to as high disease activity sub-group in TA397) and HDA-2, which are a more severe population than the total pooled population. Our analysis shows that patients in these sub-groups have a higher response rate to belimumab and are more likely to satisfy the week 24 continuation rule as compared to the total pooled population.

B10. In section B.3.3.5 of the CS, it is stated that "Using the JH cohort data, a Weibull survival model was developed explaining the risk of death with AMS included and SS item involvement effects removed." Further details are missing and it is not explicitly stated that the same model was used as in TA397.

a) Could the company please confirm whether the long-term Adjusted Mean SLEDAI (AMS) score was modelled in exactly the same way as in TA397?

Response: We can confirm that the long-term Adjusted Mean SLEDAI (AMS) score was modelled in exactly the same way as in TA397.

b) Were other models considered besides a Weibull survival model? Please provide fit statistics of each considered model (e.g. AIC/BIC).

Response: An analysis considering the fit of other survival models was considered for TA397. The Weibull model was chosen as it had the best model fit for mortality. Please see the table below.

Table 15. Model fit for mortality used in the economic models

	Mortality
	AIC
Exponential	486.4941
Weibull	469.7584
Gompertz	484.1842
Loglogistic	475.0611

B11. It appears as though updates to the model in terms of new evidence were limited. For example, BLISS 76 US LTE was only used to assess long-term organ damage in patients with SLE treated with belimumab.

a) Please provide further detail on what additional evidence was incorporated in this updated model: e.g. from LTE BLISS-52 and LTE BLISS-76 for long-term clinical effectiveness.

Response: As described in section B.3.3.4.2 of Document B, data to calculate the natural discontinuation probability in years subsequent to year 1 were derived from an integrated P2 and P3 LTE studies analyses. No other additional LTE evidence was incorporated in this model update.

b) Please explain why none of the LTE studies were used to update long-term SLE mortality risks, or other long-term parameters, if any.

Response: The BLISS LTE studies were open label extension studies designed to evaluate the safety and tolerability of add-on belimumab and assess long-term organ damage accrual. There was no formal statistical hypothesis testing performed; all analyses were exploratory. Long-term organ damage accrual has been included in the modelling (via the PSM and subsequent model validation and calibration). Other endpoints were not collected sufficiently to establish the longitudinal effect. It should be noted that the mortality rate compared favourably with that reported in patients with SLE (Wallace et al., 2013 and van Vollenhoven et al., 2020).

c) Could the company elaborate on the use of observational data from OBSErve and BILAG-BR in the updated model, and provide justification for their choice?

Response: Patient weight determines the dosage of belimumab IV a patient receives. As the BILAG-BR collected patient weight data for patients in the UK, this

data was used to update patient weights to UK specific values in the IV model. This change increases model validity by better reflecting dosage used by patients in England, and therefore treatment costs associated for Benlysta IV.

Four of the six countries participating in the OBSErve registry - Germany, Spain, Canada, and Switzerland - had a data collection period of up to six months whereas data was collected over a period of 2 years for the US and Argentina. The OBSErve registry was not used in the economic model as they did not provide data on the HDA-1 and HDA-2 subgroups and were specific to local reimbursement criteria for those participating countries.

B12. Priority question. SLICC/ACR Damage Index (SDI) change from baseline difference between belimumab and standard therapy was estimated using PSM on the total patient population in the LTE, without restrictions in terms of SS score or complement levels. Please provide:

a) **the same PSM analysis in the HDA-2 subgroup.**

Response:

Of 567 (BLISS LTE n=195; TLC n=372) intention to treat (ITT) patients, 99 from each cohort were 1:1 PS matched. If a restriction is applied to only patients who meet the HDA-2 subgroup criteria (or any subgroup that restricts the numbers of patients as compared to the total BLISS LTE population), patient numbers would be small and therefore limit the power required for analyses to be conducted robustly.

b) the results of all sensitivity analyses regarding choice of statistical method and inclusion of patient characteristics.

Response:

Sensitivity Analyses – choice of statistical method

As a sensitivity analysis the primary end point (difference in change of SDI from baseline to 5 years) was also evaluated using inverse propensity score weighting (IPSW). This propensity score (PS) method uses the entire sample and the PS to weight the observations and was undertaken to confirm the robustness of the results obtained through the PSM.

Regression-augmented IPSW was also conducted as an additional sensitivity analysis to overcome any inadequate balance with the IPSW analysis, adding variables with bias >10% as covariates in the regression model. To assess the potential for nonlinearity in the magnitude of the 5-year change in SDI score, an ordered logistic regression model (SDI change equal to 0, 1 or 2+) was estimated using the PS-matched samples. Finally, changes from baseline in SDI organ damage system sub scores were compared using Fisher's exact tests.

Sensitivity Analyses: Results by choice of statistical method – primary end point

The results of the primary end point by statistical method is provided in the table below (and also in Urowitz et al., 2019). Using PSM, 99 patients from the BLISS LTE study and 99 patients from the TLC were 1:1 PS-matched from a larger pool of 567 patients (BLISS LTE n=195; TLC n=372). This sample was well balanced, with percentage bias <5% for 9 of 17 variables and <10% for all variables (mean bias=4.6%).

It should be noted that for the IPSW sensitivity analysis, based on the full patient sample, whilst the results showed the same trend for smaller increase in SDI score on belimumab, bias in the analysis was considered statistically inadequate. The regression-augmented IPSW analysis, adding variables with bias >10% as

covariates, produced similar results, with a smaller SDI score increase for patients treated with belimumab compared with ST.

Table 16. Change in SDI from baseline to 5 years using PSM, IPSW and regression augmented IPSW

Method/variable	ST	Belimumab	Difference
PSM sample			
n	99	99	
5-year SDI change, mean (SE)	0.717	0.283	-0.434 (0.119)
95% CI	0.500 to 0.934	0.166 to 0.400	-0.667 to -0.201
P value			P<0.001
IPSW sample			
n	372	195	
5-year SDI change, mean (SE)	0.777	0.336	-0.441 (0.116)
95% CI	0.607 to 0.947	0.184 to 0.488	-0.669 to -0.222
P value			P<0.001
Regression augmented IPSW sample			
n	372	195	
5-year SDI change, mean (SE)	0.782	0.333	-0.450 (0.116)
95% CI	0.630 to 0.935	0.167 to 0.498	-0.676 to -0.223
P value			P<0.001

Additional analyses (not sensitivity analyses pertaining to the choice of statistical methods)

An additional post-hoc analysis was conducted to re-estimate the regression model for the 5 year change in SDI, to adjust for baseline corticosteroid dose and decade of entry into the study. In this augmented model, the estimated coefficient of the belimumab treatment variable remained essentially unchanged (−0.448; 95% CI −0.739 to −0.157; p=0.003).

A post-hoc, regression-augmented model estimating the differences between groups in daily average cumulative corticosteroid usage through to Year 5, adjusted for decade of entry, indicated that cumulative corticosteroid usage was lower each day by 2.045 units (95% CI −3.625 to −0.465; p=0.011) for patients treated with belimumab compared with SoC. When immunosuppressive medication use was added as a covariate in the 5 year SDI score change model for the PS-matched samples, the estimated belimumab coefficient remained essentially unchanged

(−0.449; 95% CI −0.739 to −0.159); however, this was not statistically significant, and the variation in types of immunosuppressive medication used was not considered clinically meaningful.

As the 5 year SDI change measure has a significant floor effect (ie, zero change) and does not necessarily increase in a linear manner, the analysis was re-estimated using an ordered logistic regression model (for response levels 0, 1 and 2+), using the PS-matched sample. The results indicated that patients treated with belimumab plus ST were 60% less likely than patients from the TLC treated with ST to have a 5 year change in total SDI score. If patients treated with belimumab did experience a change, they were 60% less likely to have seen a change of more than 1 unit.

Sensitivity analyses – patient characteristics

In responding to this question, we have considered patient characteristics to be interchangeable with predictors of organ damage, i.e. matching variables. Sensitivity analyses were conducted on the selection of predictors, but not extended to the results.

Predictors of SLE organ damage were chosen for PSM matching variables – a range of patient characteristics were utilised. A systematic literature review[11] was conducted to identify factors influencing and predicting SLE organ. These were augmented by an internal GSK study which studied the impact of disease activity on mortality and organ damage progression. The predictors found in the literature (table below) were then reviewed by clinical experts and limited to those for which data was available in both BEL112233 and the Toronto Lupus Cohort. One variable was available – disease activity over time – but was not suitable as a PSM variable because it was not a baseline variable.

Table 17. Predictors of SLR organ damaged identified in the literature

Predictors
Age
Gender
Race/Ethnicity
Household income
Educational attainment
SLE duration
History - hypertension
History - dyslipidemia
History - proteinuria

History - lupus anticoagulant positivity
History - anticardiolipin positivity
History - anti-β2-glycoprotein I positivity
History – anti-Ro positivity
Current smoker
Number of ACR criteria satisfied at diagnosis7
Baseline SLEDAI score
Disease activity over time (i.e., time-weighted SLEDAI)
Corticosteroid use/dose
Hydroxychloroquine/other antimalarial drug use
Cyclophosphamide/other immunosuppressive use
Initial or prior SDI
SF-20 physical functioning

This process produced the list of 14 PSM variables (column one of **Figure *13*** (taken from the PSM CSR, Table 6). All 14 variables (17 operationalized variables) were used in the PSM for the primary and secondary analyses (checked in the second column of Figure 13). The exploratory analysis on the pooled (BEL112233, BEL112234) dataset had 13 PSM variables available (checked in the third column of Table 6). The PSM variable smoker was excluded from the exploratory analyses on the pooled dataset due to an inexplicably large difference in proportions between the pooled and TLC datasets; 2% versus 24%, respectively (reported in Urowitz et al., 2020). The PSM variables were operationalized as 17 variables in the BEL112233 dataset (checked in the fifth column of Figure 13) and as 16 variables in the pooled dataset (checked in the sixth column of Figure 13). Definitions of these operationalized variables from both the US LTE and the TLC cohorts are provided in Table 7 of the PSM CSR (reference 90 of reference pack).

Baseline SDI was operationalized as a categorical variable because there were so few patients with baseline SDI > 2. The references for the operationalized Race/Ethnicity and Baseline SDI variables were Caucasian and zero, respectively.

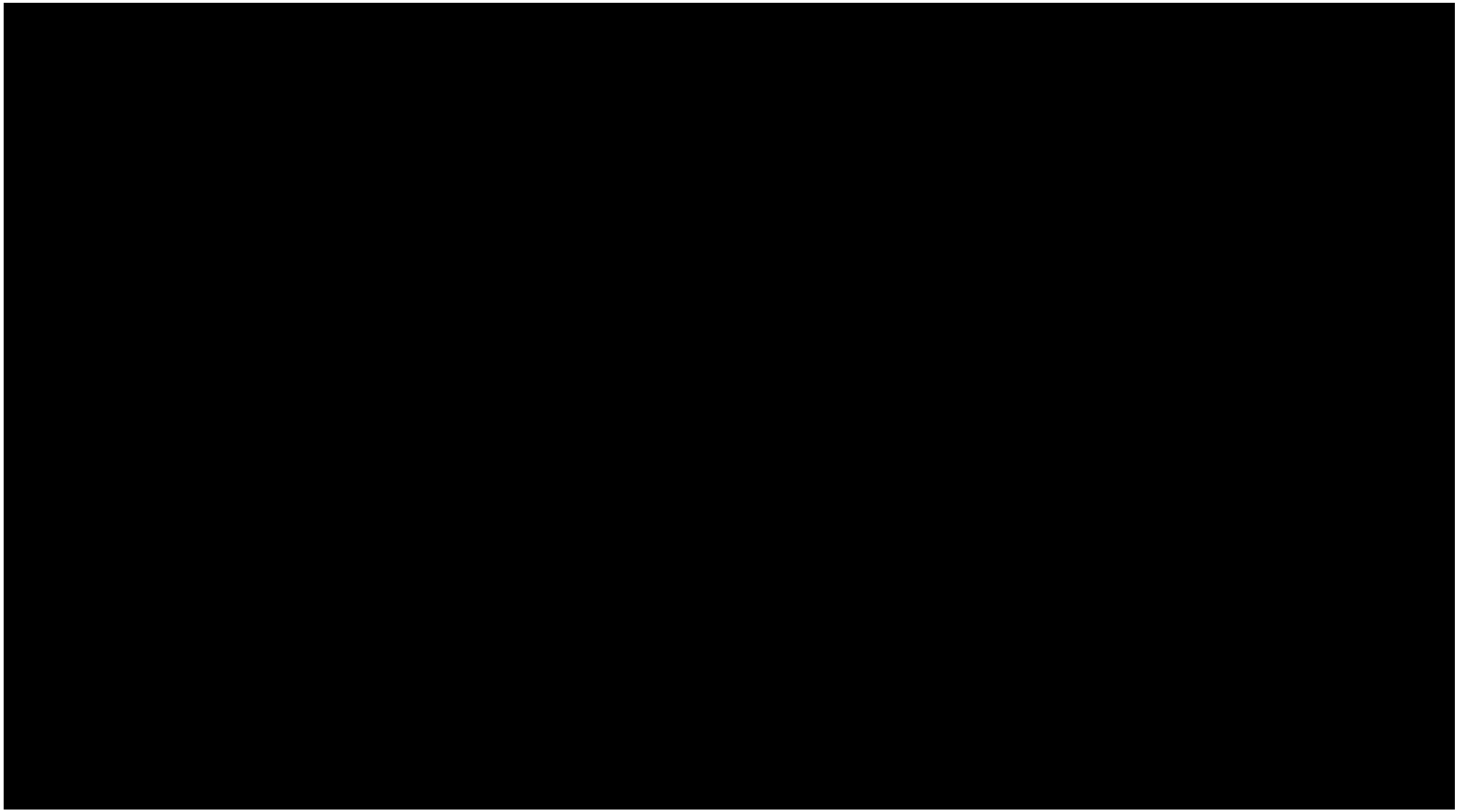


Figure 13. Propensity score baseline matching variables in the data and how they were operationalized.

Sensitivity analyses – patient characteristics

Propensity scores were calculated using logistic regression (Urowitz et al. 2019; reference 48 of the CS and the PSM CSR). The model specification included all potential predictor variables as independent variables. In a backward elimination step-wise fashion, the statistically least significant predictor was dropped from the propensity score model, until all included predictors had a p-value < 0.1. The specific predictors of organ damage included as covariates in the trimmed model was based on the model specification with the minimum Akaike information criterion (AIC) value. Attention was devoted to assessing the adequacy of the match for baseline SDI score (as likely the most important predictor of future organ damage), by comparing the frequency distribution of baseline SDI scores for the belimumab and SoC samples.

The PS value for matching was defined as the estimated log-odds from the logistic regression, rather than the predicted probability, to enhance the range of variation in the PS distribution for matching. Patients from the BLISS LTE study were matched 1:1 to patients from the TLC based on similar PS values (within a calliper value defined as 20% of the SD for the distribution of the PS variable in the full sample). Unmatched patients were excluded from the analysis of the PS-matched sample.

Note that four sets of matches were performed on the full and 'trimmed' model; 2x for BLISS US LTE/TLC cohort (primary and secondary analyses) and 2x pooled BLISS LTE/TLC cohort (exploratory):

- Analyses requiring 5 years of follow-up
- Time to event analyses requiring ≥1 year of follow-up

The full model was superior to the trimmed model (post-PSM co-variate balance) and was therefore used for all analyses.

Information sources for further detail

- Urowitz et al., 2019[12] - PSM comparative analysis for BLISS US LTE/TLC Cohort

- Urowitz et al., 2020[13] - PSM comparative analysis for Pooled BLISS LTE/TLC Cohort; exploratory post-hoc analysis that used a heterogeneous population of US and non-US patients receiving monthly intravenous belimumab from pooled BLISS LTE trials

- PSM CSR – Ref 90 of CS

c) further sensitivity analyses: includes all patients, including those that preceded 1990 and those with ≥ 15 years of follow-up

Response:

This is not available, data entry preceding 1990 was an exclusion criterion for the PSM analysis. The TLC has collected data on its patients for decades while the belimumab trials started in 2007. Therefore, an analysis could be confounded by change in treatment patterns over time. To minimize that possibility TLC patients with baseline dates before 1990 were excluded.

d) a detailed comparison of all potentially prognostic and treatment effect modifying patient characteristics between the LTE TLC studies

Response:

Please see response to b). The table below provide a summary comparison of the datasets from BLISS US LTE and TLC with 5 years follow-up (N=567) prior to propensity score matching.

Table 18. Bias prior to propensity score matching BLISS US LTE and TLC dataset with 5 years follow-up.

Variable	Mean			t-test	
	Belimumab	ST	% Bias	t	p>|t|
Age	42.769	37.303	45.5	5.01	<0.001
Age Squared	1947.4	1560.8	38.1	4.22	<0.001
Female	0.928	0.895	11.6	1.28	0.200
Black	0.231	0.153	19.7	2.29	0.022
Asian/Other Race	0.092	0.234	-39.0	-4.18	<0.001
SLE Duration	7.947	5.762	30.0	3.38	0.001
Smoker	0.036	0.237	-61.1	-6.27	<0.001
Hypertension	0.677	0.376	63.0	7.09	<0.001
Dyslipidemia	0.226	0.581	-77.5	-8.55	<0.001
Proteinuria	0.123	0.317	-48.1	-5.18	<0.001
ACR Criteria	5.923	5.651	19.8	2.22	0.027
Baseline SLEDAI	7.785	10.056	-48.4	-5.28	<0.001
Corticosteroid use	0.636	0.608	5.8	0.66	0.510

Antimalarial Use	0.738	0.519	46.6	5.17	<0.001
Immunosuppressive Use	0.538	0.315	46.4	5.31	<0.001
Baseline SDI = 1	0.272	0.148	30.7	3.60	<0.001
Baseline SDI = 2+	0.287	0.108	46.2	5.55	<0.001

Abbreviations: ACR, American College of Rheumatology; PS, propensity score; SDI, SLICC/ACR Damage Index; SE, standard error; SLE, Systemic lupus erythematosus; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index; TLC, Toronto Lupus Cohort

e) Sensitivity analyses for the cost-effectiveness analysis based on the results of any sensitivity analyses regarding the estimation of SDI change difference.

Response: No further validation of the IV economic model was undertaken based on the results of the sensitivity analyses that used alternative statistical methods to estimate the 5-year change in SDI.

The sensitivity analyses (based on IPSW and the augmented regression IPSW) were undertaken as a confirmatory step to confirm the robustness of the PSM approach. The IPSW uses the whole available population sample so the LTE and TLC populations were not considered balanced (based on bias analysis). Further, the augmented regression IPSW should be considered as a sensitivity analysis to the IPSW as it seeks to re-address the balance through the application of an additional 'matching step' through regression. With the high degree of matching that was achieved (although for a reduced cohort size) through the PSM approach, it was therefore not considered appropriate to use these sensitivity analyses to conduct alternative economic model SDI-based validations such as that conducted based on the PSM. Had they been the results across these methods are similar that limited difference would be expected following a validation to the economic model and further calibration.

f) A scenario without using the calibration exercise.

Response: Acknowledging the long-term evidence generation to evaluate organ damage accrual for those patients on belimumab and at the same time recognising the limitations on the methodology of applying this to enable the incorporation into the economic model, GSK has taken a conservative approach to balance these aspects. We re-iterate these important considerations here:

- Application of a calibration factor is applied following a model validation exercise to seek the difference in the simulated change in SDI between 6.5 and 1.5 years versus that seen in the PSM analysis.

- To ensure the value of lower accrual damage whilst on belimumab, the calibration factor is applied only to years 1.5 to 6.5 years of the model. As there is no evidence to suggest that the benefit on reducing organ damage accrual decreases over time while patients continue to receive belimumab, this is considered a conservative approach.

- We do not apply a calibration to 'worsen' the impact on damage accrual for those on ST despite the findings from the PSM.

We have provided the requested analysis where calibration factors have not been applied, despite evidence to the contrary. Please see the following analysis conducted for the IV and SC models for the HDA-2 population, where no calibration factor has been applied. To provide appropriate balance we have also provided an analysis where a calibration factor has been applied to both the belimumab and ST arms (for between 1.5 years and 6.5 years).

Table 19. IV model with the HDA-2 population - Calibration factors applied to belimumab only for 6 years (base case

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
ST	£160,470	16.90	9.81				
Belimumab IV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£30,001

Table 20. IV model with the HDA-2 population - Calibration factors applied to both belimumab and ST for 6 years

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
ST	£167,261	16.76	9.67				
Belimumab IV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£21,635

Table 21. IV model with the HDA-2 population – No calibration applied

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
ST	£160,470	16.90	9.81				
Belimumab IV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£47,872

Table 22. SC model with the HDA-2 population - Calibration factors applied to belimumab only for 6 years (base case)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
ST	£151,999	17.12	10.06				
Belimumab IV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£30,566

Table 23. SC model with the HDA-2 population - Calibration factors applied to both belimumab and ST for 6 years

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
ST	£158,791	17.00	9.92				
Belimumab IV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£23,353

Table 24. SC model with the HDA-2 population - No calibration applied

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
ST	£151,999	17.12	10.06				
Belimumab IV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£56,277

All model outcomes presented are discounted. Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years

B13. Priority question. On page 154 of the CS it states: "As the IV model captures the observed pooled analysis results from the pooled P3 studies, it was decided that the validation exercise of the deterministic model should be simulated as a 5-year increase in SDI score (further from the baseline duration of 1.5 years). The model starts at the beginning of the BLISS trial, hence the period from 1.5 to 6.5 years from the model was chosen to compare with the PSM analysis results. This simulated an SDI score increase of 0.568 in the

belimumab arm and 0.611 in the ST arm, respectively (Table 65, Figure 9).” Could the company please provide the following:

a) **Confirmation that ‘pooled P3’ refers to the Phase 3 BLISS-52 and BLISS-76 clinical studies.**

Response: Yes this is correct for the IV model.

b) **Confirmation that the 5-year values reported in Table 65 of 0.568 and 0.611 for ‘Cost-effectiveness model; matched LTE ITT population’ were extrapolated from 1.5 years of observed data from the Phase 3 BLISS-52 and BLISS-76 clinical studies.**

Response: Note that up to 1.5 years organ damage time to event models from the JHC are used. The value 0.568, in Table 65, is the SDI increase simulated in the model from 1.5 years, for a duration of 5 years.

c) **Detailed description of how the 5-year values reported in Table 65 of 0.568 and 0.611 for ‘Cost-effectiveness model; matched LTE ITT population’ were estimated. This would include the role of both the 1.5 years of observed data from the Phase 3 BLISS-52 and BLISS-76 clinical studies and data from the Johns Hopkins database.**

Response: A validation exercise simulated the 5-year SDI increase which is reported in Table 65 of the CS. To do this, the baseline characteristics of the model population (IV model) were adjusted to ensure comparability to the BLISS LTE population. These settings are provided in Table 64 of the CS.

Note that the PSM was conducted on a sub-set of the BLISS LTE (N=99) and this resulted in small differences in patient characteristics which were discussed and not assumed to impact on the relevance of the analysis.

The table below shows the baseline characteristics from the PSM analysis and the cost-effectiveness model with baseline characteristics adjusted for the validation exercise.

Table 25. Patient characteristics in the PSM analysis and cost-effectiveness analysis

Baseline characteristic	Propensity score-matched analysis		Cost-effectiveness model (total BLISS population trial data)
	Belimumab + SoC N=99	SoC N=99	Average of belimumab + SoC and SoC arms
Age, years (mean)	40.0	39.0	38.0
Female (%)	92.2	91.9	94.3
SLE duration, years (mean)	7.4	7.6	6.4
SLEDAI (mean)	8.5	8.5	9.74
SDI = 0 (%)	60.6	54.5	Average of 0.76
SDI = 1 (%)	24.2	27.3	
SDI = 2 (%)	15.2	18.2	

The IV model with adjusted baseline characteristics was then run deterministically. The 5-year SDI increase was then calculated from 1.5 years to 6.5 years for both the belimumab and ST. We took the SDI score at 6.5 years from the model results (average of 6 and 7 years, as the model runs in annual cycles) and subtracted the SDI score result at 1.5 years (average of 1 and 2 years) to get the SDI score increase between1.5 and 6.5 years.

The Johns Hopkins Cohort is used to derive and estimate time to event (TTE) models to describe the relationship between disease activity and other covariates on the risk of dying and on the risk of developing irreversible organ damage. The TTE models are then implemented in the model to simulate a patient's future disease course based on the severity of the population and the short-term outcomes observed in the BLISS trials.

The validation exercise sought to understand whether the IV model under- or- over-estimated the prevention of organ damage, which it does so (pre-calibration) indirectly through the reduction of disease activity, through the TTEs. The model validation shows that the indirect effect underestimates the observed effect of belimumab versus ST on long term organ damage accrual as seen in the PSM.

d) Given that the 1.5 years of observed data from the Phase 3 BLISS-52 and BLISS-76 clinical studies provides an unbiased estimate of the treatment effect (difference in SDI between belimumab and ST), justification for the use of data at high risk of selection bias (TLC data and the PSM analysis

of those data) for validating the extrapolation from the Phase 3 BLISS-52 and BLISS-76 clinical studies.

Response: It is important to consider the duration of these studies and then how they are implemented in the IV model. The pivotal Phase 3 studies BLISS-52 and BLISS-76 provide 52 weeks and 76 weeks of observed evidence respectively for ST alone and BEL added onto ST. One of the study end points was the Systemic Lupus International Collaborating Clinics (SLICC) Damage Index (SDI) change. SDI is a measure of organ damage progression. The damage items (recorded irrespective of their attribution to SLE) have to persist for a minimum of 6 months or be associated with immediate pathological scar indicative of damage.

Change in SRI beyond the 52 weeks and 76 weeks of the pivotal BLISS IV studies is captured in the BLISS LTE studies. As organ damage progresses slowly, a very low number of events would be expected during the RCTs follow-up (given it needs to persist for at least 6 months). Therefore, it is too short a time frame to draw meaningful conclusions about the difference in organ damage progression based on BLISS-52/76 RCTs.

Longer follow-up is clinically more relevant. The BLISS LTE studies were non-comparative and non-hypothesis testing studies designed to evaluate the safety and tolerability of add-on belimumab only. The PSM study was conducted to estimate the 5-year increase in SDI in patients on add-on belimumab and compare this to a matched cohort on ST (from TLC). Selection bias is, as best as possible, mitigated through appropriate matching of characteristics of belimumab patients with ST patients in the TLC.

The results from the PSM was used to validate the organ-damage progression simulated in the IV model for a belimumab patient and a ST patient where organ damage is indirectly estimated (beyond BLISS-52 and BLISS-76) through the reduction in disease activity (see response to c)).

e) **Justification for why the company consider that the cost-effectiveness model overestimated SDI progression in the belimumab arm and underestimated SDI progression in the ST arm, as opposed to the PSM**

analysis of the TLC data underestimating SDI progression in the belimumab arm and overestimating SDI progression in the ST arm.

Response:

There are important distinctions in the way in which SDI progression was measured versus how SDI progression is modelled. SDI progression as an outcome needs to be considered longitudinally as accrual of damage over time. At the time of the previous CS (2011), long-term evidence of SDI progression was limited (for belimumab, in the absence of the BLISS LTEs). So, for the IV model (in 2011), the relationship between the short-term outcomes captured in BLISS-52 and -76 and the long-term outcomes was estimated based on the Johns Hopkins Lupus Cohort. Therefore, without new evidence, the effect of belimumab on organ damage progression was only indirectly captured through the reduction in disease activity. For this reason, the appraisal committee previously recognised this as an area of underestimated benefit for belimumab (since the model over estimated SDI progression). For the PSM analysis, whilst based on a limited sample to ensure robust matching, the SDI progression for patients on belimumab was directly captured over the 5-year time frame and therefore it is less likely to be an overestimate of SDI progression.

With regards to SDI progression for patients on ST, there is relatively little difference between the IV model and that reported from the PSM analysis (approximately 15%). This may therefore suggest some underlying similarities between the Johns Hopkins Lupus cohort and the TLC cohort in a BLISS-LTE-like population. The 15% difference could be owing to the impact of 'matching' the TLC ST cohort to the BLISS-LTE cohort. Note, in our base case, we do not apply calibration factors to the ST arm we only seek the additional benefit for those on add-on belimumab. This is a conservative assumption and if we had applied it, the resultant ICER for belimumab would improve.

f) Discussion of appropriateness of, and potential bias induced by, applying the calibration factors in the SC model.

Response: No equivalent PSM analysis has been conducted for the belimumab SC formulation. At this current time, long-term follow-up with the SC formulation is

limited to approximately 6 months i.e. less than the duration of the Phase 3 study, BLISS-SC.

The application of the same calibration factors derived for the IV model and also utilised in the SC model is a reasonable approach and unlikely to contribute significant bias.

The baseline characteristics of the ITT population from the IV and SC LTEs are broadly similar (see Table 17 of the CS). The baseline characteristics (from the IV LTEs) were used to match to ST patients from the TLC for the PSM analysis. So in theory, similar baseline characteristics would have been matched if a similar exercise had been conducted for the SC formulation (if longer LTE data had been available for belimumab SC).

Further, as discussed in the CS, an ITC which used patient level data to compare the efficacy of the IV and SC formulations showed comparability between the formulations across the range of key end points (including SRI response and ≥4 point reduction in SS) (Ramachandran et al., 2018).

g) **A comparison between the SDI scores estimated using the observed data (up to 1.5 years) from the Phase 3 BLISS-52 and BLISS-76 clinical studies and the SDI scores estimated using the PSM. If there is a discrepancy between them then please comment on the validity of the PSM analysis for calibrating the model.**

Response: The question does not reflect the methodology undertaken. The explanation in the above questions and responses should seek to resolve this. We calibrate the model on the progression in 5 years i.e. difference between years 6.5 and 1.5 years.

h) **Scenarios where the treatment effect (difference between belimumab and ST) observed at 1.5 years is assumed to wane over time.**

Response: The question suggests that there may be a misunderstanding with the aim and method in conducting the calibration exercise. The explanation to the above questions should seek to resolve this.

At the time of the previous appraisal there was limited evidence to inform the maintenance of treatment effect. Clinical experts explained that in those people whose disease responded to rituximab and who then needed re-treatment with rituximab at a later stage had shown a good response to re-treatment. The 2016 Guidance was provided on the assumption that whilst a patient is maintained on belimumab, the treatment effect is maintained. Since TA397, clinical experience with belimumab IV, with a duration of follow-up beyond three years is limited; during the available follow-up, treatment effect observed aligns to the HDA-1 treatment effect in BLISS-52 and BLISS-76. The BLISS LTEs (open label, non-comparative, to evaluate the safety and tolerability of add-on belimumab), show that from baseline the organ damage accrual remained stable. The application of the PSM comparative analysis in the economic model is applied only to the belimumab cohort and for 5 years only. As there is no evidence to suggest a treatment effect waning over time even out to 13 years (Wallace et al., 2013), this has not been modelled.

B14. Priority question. In section B.3.3.6.4 in the CS, it is mentioned that, to derive the calibration factors, the model was simulated several times with varying calibration factors, until the model's results matched the observed results from the PSM up to 3 decimals. The results of this approach are not presented in the CS and it is not necessarily the case that the calibration factor which most closely resembles the 5 year estimates, is the factor that provides the best estimate over the whole time period. In Tables 34 to 38 of the CSR regarding the PSM (reference 90 in the CS), the change of total SDI score from baseline to end of years 1 through 5 is presented.

a) **Please provide a cross-validation between the results in these tables and the estimates derived from the model at each year by assuming various calibration factors, to help demonstrate whether the chosen calibration factor makes a good fit at all time points – and provide any other analyses or data to support this.**

Response: The total SDI difference of change from baseline at each year from the PSM analysis is plotted against the estimates derived from the model using different calibration factors in the figure below. The difference of change from baseline at year 1 through to year 4 were obtained from the regression equations in Section 6.2.5 in

the PSM CSR (tables 105-108). As explained in the CS, the calibration factor of 0.491 was determined by anchoring the model simulation estimates on the 5-year change from the PSM analysis, being the primary outcome of the study. The figure below shows that with this calibration factor, the estimated model results are reflected well for years 2, 4 and 5, and slightly below at years 1 and 3. A similar exercise was conducted for SOC (see figure below). This shows that the SOC calibration factor of 1.186 anchored on the 5-year PSM data structurally underestimates organ damage progression for SOC till year 5. This suggests that if the SOC calibration factor was based on all years, it would be higher than 1.186. The base case analysis in the CE model conservatively assumed no calibration for SOC (i.e. a calibration factor of 1), thereby potentially underestimating organ damage with SOC. This analysis shows that the degree of organ damage underestimation for SOC would be even higher if the calibration would have been based on the other timepoints as well, thereby further overestimating the ICER.

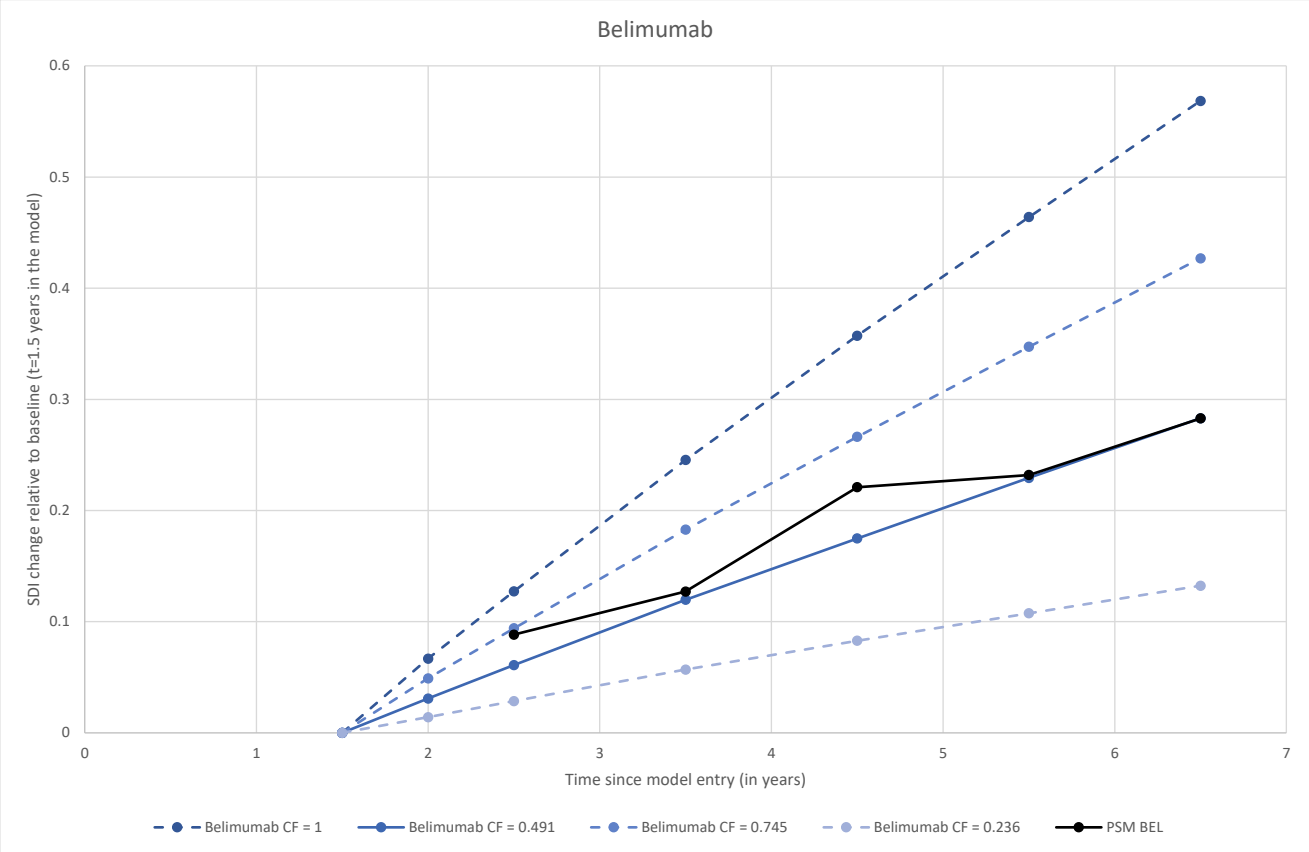


Figure 14. Time since model entry versus SDI change relative to baseline for calibration factors associated with belimumab

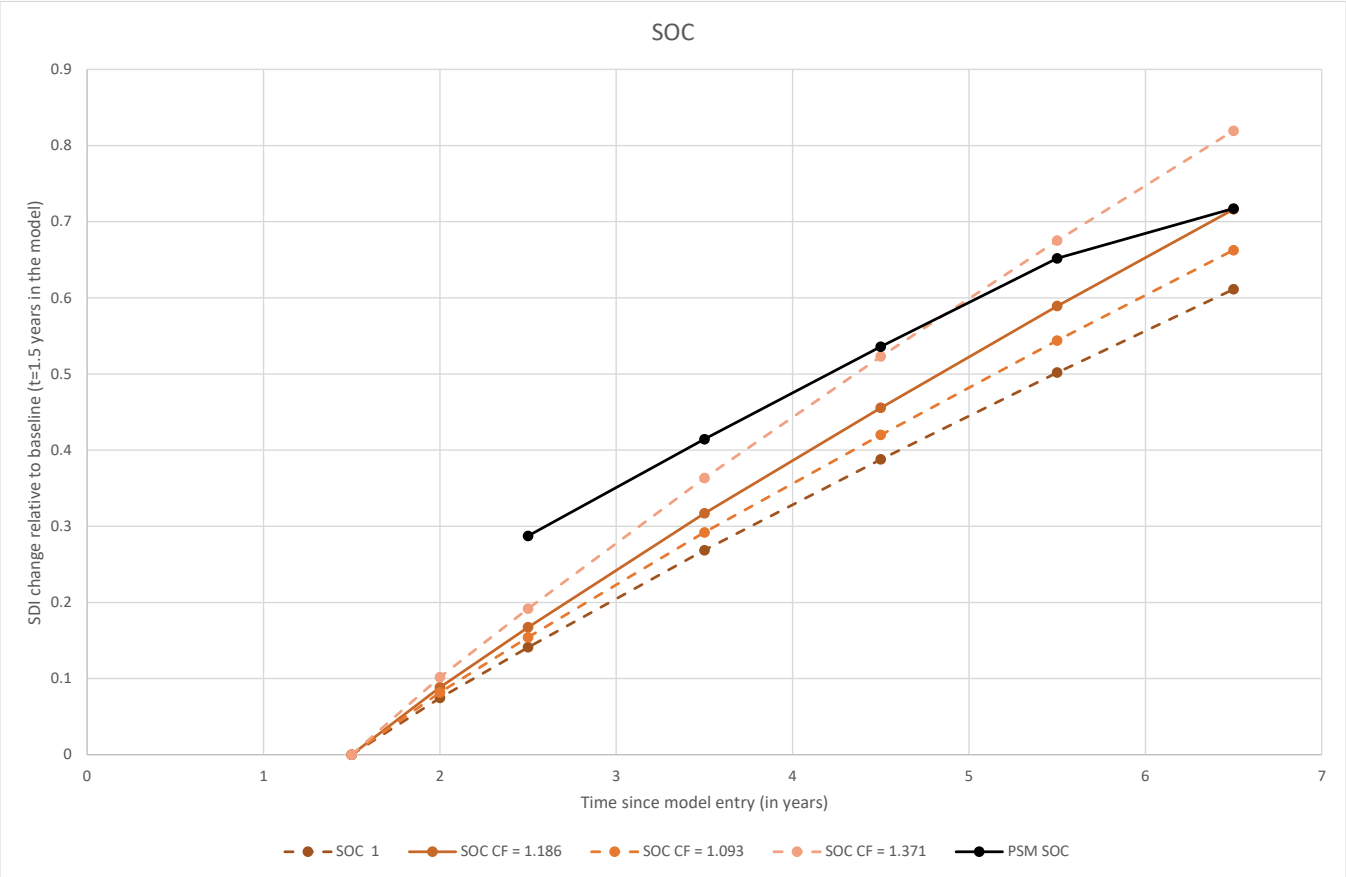


Figure 15. Figure showing time since model entry versus SDI change relative to baseline for calibration factors associated with standard of care

B15. Baseline patient characteristics:

a) Please confirm that no correlation has been taken into account in the simulation of patient baseline characteristics. Despite the potential correlation between baseline characteristics, they are sampled independently. Bootstrapping (i.e. sampling from the trial data) was considered but would underestimate the actual heterogeneity when simulating 50,000 patients. Due to the number of baseline characteristics and the different types of distributions, it was considered too complex to apply correlated sampling to the model. This is a limitation and can be accounted for in future model

changes, but it is expected that this will not greatly influence the average results.

Response: We can confirm that no correlation has been taken into account in the simulation of patient baseline characteristics.

b) Please reflect upon the possibility of generating implausible patient profiles (for instance short disease duration and high organ damage) and discuss the potential impact on model outcomes.

Response: It is possible to generate implausible patient profiles, but as described in part c) for the response for this question, this probability of this happening is extremely low. The existence of such profiles within the cost-effectiveness analysis is disadvantageous for belimumab and would adversely affect the ICER, as the opportunity to limit disease progression (in terms of long-term organ damage) for these patients is diminished.

c) Could you check generated patient profiles for their plausibility and provide an overview (e.g. proportion of simulated patients whose profiles are clinically implausible).

Response: For question B4, part c, 50,000 patient profiles were generated and included with this clarification response. The table below shows the proportion of patients simulated with SDI scores greater than 2 and disease duration less than 5 years. Although the chance of these combinations happening in practice is low, they are clinically not implausible. As argued above, the impact of these patient profiles on the model outcomes is considered negligible.

Table 26.Patient profiles (%) showing disease duration vs. SDI score

Disease duration (years)	SDI>2	SDI>3
1	0.8%	0.2%
2	0.7%	0.2%
3	0.6%	0.2%
4	0.5%	0.1%
5	0.4%	0.1%

There were no other combinations of implausible baseline characteristics. SLEDAI is an active disease score that can vary substantially, irrespective of organ damage (SDI) and disease duration. The same applies to baseline steroid dose, relative to SLEDAI, SDI and disease duration.

B16. For patients who discontinue belimumab due to not achieving response, it is unclear what levels of disease activity they revert to. Natural discontinuation appears to be only modelled for responders, as non-responders discontinue.

a) Can the company confirm that it assumed the average ST level of disease activity for the remainder of the model horizon?

Response: We can confirm that this is the case. Please see the answer provided as part of question B7 b).

b) Please confirm that all non-responders discontinue at 24 weeks and comment on the probabilities of natural discontinuation being independent of patient profiles (other than if treatment continuation criteria are not met) and whether this is appropriate and the potential bias this may introduce.

Response: We can confirm that all non-responders in the economic model discontinue at 24 weeks in line with the responder rule outlined by NICE in 2016 Final Appraisal Document. No variables in patient profiles have been noted as predictors for natural discontinuation, so these are currently assumed to be independent, and not to be a source of any bias.

B17. Adverse events were not included in the model. Please provide a scenario including adverse events and their impact on HRQoL and resource use and costs in the economic model.

Response: Adverse events (AEs) were not included in the health economic model. As discussed in Section 5.9.2 of the company submission provided as part of TA397, the Phase 2 and 3 studies did not find important differences in the incidence of all AEs and serious adverse events (SAEs) between the belimumab and placebo treatment groups. Importantly, the incidence of serious infections such as pneumonia, UTI, cellulitis, bronchitis, and pyelonephritis, which would require treatment in hospital and thus incur a significant cost to the NHS, was not

significantly higher in the belimumab treatment arms compared with placebo. Therefore, by not including AEs in the model, it is expected that this would not have an important impact on the cost-effectiveness results. The absence of adverse events in the economic model was not identified as an issue in the TA397 Final Appraisal Document.

Health-related quality of life

B18**.** Literature searches for utility values do not appear systematic. Please comment on the comprehensiveness and potential bias in selected utilities?

Response: The SLICC/ACR Damage Index captures 41 unique clinical problems in 12 organ systems. To obtain the potentially substantial quality of life impairments associated with organ damage in SLE a literature search was undertaken for the original model for each of these 41 damage items in the SLICC score. It should be noted that the scope of this search is almost unbounded due to the number and variety of different organ systems that are contained in the SLICC as well as the sometimes-broad definition of damage. For this reason, the utility values were not systematically identified.

The relative contribution to the ICER of most individual organ damage related costs and utilities were shown to be very limited in the original submission and remains the case for this submission. Therefore, before embarking on a full update we undertook a sensitivity analysis to understand which utilities for which organ systems contribute significantly to the ICER and therefore updates focused on the most influential organ systems; of which there were 7.

For these organ systems, a comprehensive approach has been taken by searching the NICE website for utility values used in recent submissions. In addition, a PubMed search was conducted if searching the NICE website did not provide any appropriate results. The search term used to identify utility literature in PubMed was MESH.EXACT.EXPLODE("Quality of Life"). As no additional search terms have been used, this can lead to bias if utility publications have not been labelled with the mesh term “Quality of Life”. Furthermore, the search term MESH.EXACT.EXPLODE("Cost-Benefit Analysis") was included which could capture inter alia cost-utility analysis publications as utility sources. However, as mentioned

before the search has not been conducted fully systematically in line with NICE requirements, and hence no SIGN search filters for HRQoL have been used. In principle, this could lead to bias, in the sense that the PubMed searches may not have captured all utility publications in the search results.

The company acknowledges that this approach could lead to bias but it is unknown in which direction that bias would be (i.e. it would increase or decrease the ICER). However, the relative contribution for these utility scores is low and any impact from bias would also be relatively minor.

B19. Please provide the appendix describing the process behind the weighting for the QoL scores for the organ damage multipliers (which is not available anymore) and any changes to the original modelling, if applicable.

Response:

Please see the file named ‘ID1591 B19. Clarification - QALY and Cost Weighting’. This file shows utilities and costs used in the current and previous economic model and enables a direct comparison.

B20. The SS score based and organ damage based utilities seem to overlap as for example arthritis and skin damage/rashes are mentioned in both. If this is the case, some items will be double counted and therefore receive unwarranted higher importance than others. Please elaborate on any potential double-counting and its impact on utility estimates and resource use and costs, and provide an updated analysis with any double-counting removed, if necessary.

Response: The SS score-based utility regression has been adjusted for organ damage items, thereby preventing double-counting. This was done by the addition of indicator variables for each organ damage system into the regression. Furthermore, if patients have multiple organ damage system involvement, only the domain featuring the lowest utility values is used. If all disutilities were used, this would overestimate the impact of multiple organ damage on a patient’s quality of life.

Resource use and costs

B21. Section B.3.5.5 explains how organ damage costs were updated. The company argues that "Due to restriction of the update searches to seven key organ systems, no searches for costs relating to diabetes, gastrointestinal, ocular organ systems were conducted." This is not clear to the ERG.

a) Please explain why update searches were restricted to seven key organ systems and why therefore these 3 were excluded.

Response: Please see the answer provided for question B18.

b) If appropriate, conduct searches for these 3 items and include them in the CE model.

Response: On reflection, it would be appropriate to conduct a search for the remaining organ systems not included in the search update; ocular, gastro-intestinal, skin, diabetes and gonadal failure.

When reflecting on the impact of not having conducted searches, Tables 7 and Table 8 of Appendix J, showing the disaggregated model costs for the HDA-2 population in the IV and SC models were checked to try and understand the potential impact of costs changing.

For the IV model, absolute costs difference between the ST and belimumab arms, were below £20 for each of 4 of the 5 domains. The absolute cost difference for the remaining domain, diabetes, the domain with the largest absolute cost difference over a patient's lifetime was £135, favouring ST. Even if the absolute cost difference for diabetes was doubled, the total impact would be an additional £135 for a belimumab IV patient to their total costs over their lifetime and thus the anticipated impact on the ICER would be minimal. For comparison, absolute cost difference for the pulmonary and renal domains between belimumab and ST in this model were £9,828 and £5,100 respectively.

For the IV model, absolute costs difference between the ST and belimumab arms, were below £9 for each of 4 of the 5 domains. The absolute cost difference for the remaining domain, diabetes, the domain with the largest absolute cost difference over a patient's lifetime was £192, favouring ST. Even if the absolute cost difference

for diabetes was doubled, the total impact would be an additional £192 for a belimumab SC patient to their total costs over their lifetime and thus the anticipated impact on the ICER is minimal. For comparison, absolute cost difference for the pulmonary and renal domains between belimumab and ST in this model were £9,828 and £5,100 respectively.

Due to time constraints, we are not able to provide the updated searches and model outputs in the timescales required for clarification questions. However, we feel confident that even if the models were updated using these searches, the anticipated impact on the ICER would be minimal.

B22. Section B.3.5.5 explains that costs for premature gonadal failure and skin organ systems were not updated because they were not searched previously. In the original company submission, no argument was found for why these systems were excluded in the first place.

a) Please explain why costs for skin and gonadal failure organ systems were excluded.

Response: Since belimumab has no effect on gonadal failure, costs and utilities were not considered for this condition. For the skin organ domain, no data were identified on the costs for treatment of scarring alopecia, scarring or skin ulceration.

b) If appropriate, include these costs in the CE model.

Response: Please see response to a).

B23. Section B.3.5.4 explains that SS related costs were inflated based on 2005/2006 NHS reference costs to 2018/2019 values using the consumer price index. There may be changes to the cost beyond the inflation of the NHS reference cost.

a) Please reflect on the appropriateness of not updating the resource use as clinical practice may have changed since 2005/2006.

Response: No evidence for significant change in the management of patients with SLE or resource use associated with belimumab IV (in terms of administration required, staff training, time taken to administer, NHS agenda for change staff used

to administer the IV formulation, and time required in specialised infusion suites) since 2005/2006 was noted.

b) Explain why updated reference costs were not used and provide an updated model with updated costs instead of inflating cost based on 2005/2006 NHS reference costs.

Response: On reflection this is an appropriate update request. In the time available and to provide some reassurance, we have evaluated the impact on the ICER and it is very limited. The base case ICER for the SC formulation for the HDA-2 population subgroup presented as part of the current submission was £30,566. In an additional sensitivity analysis where the SS related costs were doubled, the resulting ICER was £31,421. Another analysis where the relative values were halved resulted in an ICER of £30,139.

c) Please provide updated costs if appropriate.

Response: Please see response to b).

Sensitivity analyses

B24. It appears that not all parameters were included in the probabilistic sensitivity analysis (PSA). The naming of parameters explored in sensitivity analysis is not clear.

a) Please provide a table overview of all parameters used in the model including descriptions, highlighting those that were used in the PSA.

Response: Please see the table below.

Table 27. Parameters used in the model

			IV model with HDA-2 population			SC model with HDA-2 population		
Varied in PSA	**Rationale if excluded from PSA**	**Variable**	**Value (reference to appropriate table or figure in submission)**	**Measurement of uncertainty and distribution: CI (distribution)**	**Reference to section in submission**	**Value (reference to appropriate table or figure in submission)**	**Measurement of uncertainty and distribution: CI (distribution)**	**Reference to section in submission (Document B)**
		Patient characteristics at baseline						
No	Mean baseline patient characteristics not varied in PSA	Age (years)	[redacted]	[redacted]	Table 58	[redacted]	[redacted]	Table 58
No		Percentage females (%)	[redacted]	[redacted]		[redacted]	[redacted]	
No		Percentage Black Ethnicity (%)	[redacted]	[redacted]		[redacted]	[redacted]	
No		SLE disease duration (yrs)	[redacted]	[redacted]		[redacted]	[redacted]	
No		SLICC/ACR damage index score	[redacted]	[redacted]		[redacted]	[redacted]	

		Baseline disease activity parameters and steroid use simulated at baseline						
No	Mean baseline patient characteristics not varied in PSA	Baseline SLEDAI	[redacted]	[redacted]	Table 59	[redacted]	[redacted]	Table 59
No		Increased DNA binding	91.4%	0.914, Bernoulli		92.4%	0.924, Bernoulli	
No		Low Complement	83.1%	0.831, Bernoulli		66.6%	0.666, Bernoulli	
No		Vasculitis	11.8%	0.118, Bernoulli		10.5%	0.105, Bernoulli	
No		NP involvement	0.6%	0.006, Bernoulli		0.0%	0.000, Bernoulli	
No		Renal involvement	6.4%	0.064, Bernoulli		4.8%	0.048, Bernoulli	
No		Serositis involvement	1.1%	0.011, Bernoulli		5.7%	0.057, Bernoulli	
No		Haematological Involvement	6.4%	0.064, Bernoulli		1.8%	0.018, Bernoulli	
No		Skin Involvement	57.0%	0.57, Bernoulli		77.6%	0.776, Bernoulli	
No		[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	
		Linear regression explaining change in SELENA-SEDAI score after 52 weeks compared to ST						
Yes		SS^0 ST	[redacted]	[redacted]	Table 61	[redacted]	[redacted]	Table 61
Yes		SS^0 all belimumab						
Yes		SS^0 belimumab responders						
		Summary of percentage belimumab continuations and natural discontinuation						
Yes		% belimumab patients satisfying treatment continuation rule at 24 weeks	[redacted]	[redacted]	Table 62	[redacted]	[redacted]	Table 62
Yes		Natural discontinuation rate for patients satisfying treatment continuation criteria at 24 weeks in Year 1	[redacted]	[redacted]		[redacted]	[redacted]	
Yes		Natural discontinuation rate for patients in year 2 and subsequent years	[redacted]	[redacted]		[redacted]	[redacted]	
		Calibration factors						
Yes		Belimumab + ST	[redacted]	[redacted]	Table 66	[redacted]	[redacted]	Table 66
Yes		ST	[redacted]	[redacted]		[redacted]	[redacted]	Table 66
		Utility multipliers per organ system (SD for all organ systems assumed at 10%)						
		Organ System	**Year 1**	**Year 2**	**Reference**			
Yes		Cardiovascular	0.717	0.779	Table 68	These utility multipliers are used in both the IV and SC models		
Yes		Diabetes	0.910	0.910				
Yes		Gastrointestinal	0.786	0.786				
Yes		Malignancy	0.919	0.837				
Yes		Musculoskeletal	0.665	0.655				
Yes		Neuropsychiatric	0.679	0.713				

Yes		Ocular	0.974	0.974				
Yes		Peripheral vascular	0.856	0.863				
Yes		Premature gonadal failure	1	1				
Yes		Pulmonary	0.693	0.713				
Yes		Renal	0.972	0.972				
Yes		Skin	0.943	0.943				
		Model cost inputs						
		Type	**Cost**	**Varied in PSA**	**Reference**	**Used in IV model**	**Used in SC model**	
N/A		Belimumab 120mg vial	[redacted]	No	Table 69	Yes	No	
N/A		Belimumab 400mg vial	[redacted]	No		Yes	No	
N/A		Belimumab 200mg subcutaneous prefilled pre-filled pen	[redacted]	No		No	Yes	
No		Admin cost per IV infusion	£154	No		Yes	No	
No		Cost of specialist hospital-based nurse per hour to deliver SC training	£113	No		No	Yes	
		Other variables used in the model						
		Item	**Value**	**Varied in PSA**	**Reference**	**Used in IV model**	**Used in SC model**	
No	Not uncertain. Fixed dosing scheme	Number of IV infusions in year 1	14	No	Table 69	Yes	No	
No	Not uncertain. Fixed dosing scheme	Number of IV infusions in year 2 onwards	13	No	Table 69	Yes	No	
N/A		Exposure to drug	100%	No	Table 73	Yes	Yes	
N/A		Vial sharing	Off	No	Table 73	Yes	No	
No	Baseline patient characteristic	Average weight	70.4kg	No	B.3.2.3.1.1	Yes	No	
N/A		Discount rate for costs	3.5%	No	B.3.2	Yes	Yes	
N/A		Discount rate for effects	3.5%	No	B.3.2	Yes	Yes	
		Disease activity related costs per year 2018/2019						
		SELENA-SLEDAI Score	**Yearly cost**	**Reference**				
Yes		0	£1294.53	Table 70				

Yes		1	£1444.42					
Yes		2	£1594.30					
Yes		3	£1700.50					
Yes		4	£1762.95					
Yes		5	£1825.40					
Yes		6	£1887.86					
Yes		7	£1950.31					
Yes		8	£2012.76					
Yes		9	£2085.65					
Yes		10	£2168.92					
Yes		11	£2252.19					
Yes		12	£2335.46					
Yes		13 - 20	£2418.76					
		Johns Hopkins cohort characteristics						
		Item	**Value**	**Reference**				
No	Average background baseline characteristics not appropriate for varying in PSA	Number of patients	1282	Table 6.7 of the previous submission				
No		Females	1,190 (92.8%)					
No		Black ethnicity	492 (38.4%)					
No		Caucasian	672 (52.4%)					
No		Age at diagnosis (mean (SD)	33.1 (13.0)					
No		Age at cohort entry (mean (SD)	38.2 (12.8)					
No		Disease duration at cohort entry (mean (SD)	5.15 (6.5)					
No		SLEDAI score at first visit (mean (SD)	3.32 (3.7)					
No		Steroid dose at first visit (mean (SD)	9.95 (15.3)					
No		Past smoker (%)	38.9%					
No		Hypertension (yearly risk)	15.8%					
No		Anti-cardiolipin antibodies positive (%)	3.0%					
No		Lupus anticoagulant positive (%)	9.6%					

		Coefficient results for the linear regression model predicting change in mean SLEDAI – Johns Hopkins Cohort						
		Item	**Coefficient**	**95% Confidence Intervals**		**Reference**		
Yes		Mean SLEDAI score in previous period	-0.4163	-0.4396	-0.3929	Table 6.9 of the previous submission		
Yes		Male gender	-0.0991	-0.2544	0.0562			
Yes		Black ethnicity	0.3524	0.2566	0.4482			
Yes		Log of age	-0.3586	-0.5072	-0.2100			
Yes		Constant	2.0577	1.4855	2.6299			
N/A		Sigma ui	0.4093					
N/A		Within R^2	0.3624					
N/A		Overall R^2	0.1668					
		Linear regression model explaining average steroid dose per year (mg/day) based on SLEDAI score (model input) - Johns Hopkins cohort						
		Regression parameter	**Coefficient (95% CIs)**	**P-value**	**Reference**			
Yes		Average SLEDAI score during current year	0.7199 (0.617, 0.823)	<0.001	Table 6.11 of the previous submission			
Yes		Constant	3.410 (3.073,3.747)	<0.001				
		Weibull survival model explaining risk of death with AMS included and item involvement effects removed - JH cohort						
		Covariates	**Model coefficient**	**Reference**				
Yes		Constant	-10.366	Table 6.12 of the previous submission				
Yes		Black ethnicity	0.7814					
Yes		Age at diagnosis	0.0321					
Yes		Cholesterol	0.0044					
Yes		AMS over lifetime	0.2135					
Yes		Cumulative Average Prednisone Dose (mg/month)	0.0012					
Yes		Renal damage	0.652					
Yes		Musculoskeletal damage at previous visit	0.415					
Yes		Peripheral vascular damage at previous visit	0.9783					

Yes		Gastrointestinal damage at previous visit	0.4684					
Yes		Diabetes at previous visit	0.6764					
Yes		Malignancy at previous visit	1.1489					
Yes		Any infection at time of death at current visit	0.7409					
N/A		Parametric distribution parameter for Weibull	1.6799					
		Average SLICC scores per organ - based on all recordings in Johns Hopkins cohort						
		Organ	**Score**	**Reference**				
No	Clinical sub-outcome only, does not affect costs/utilities	Cardiovascular	1.42	Table 6.16 of the previous submission				
No	Clinical sub-outcome only, does not affect costs/utilities	Diabetes	1.00					
No	Clinical sub-outcome only, does not affect costs/utilities	Gastrointestinal	1.09					
No	Clinical sub-outcome only, does not affect costs/utilities	Malignancy	1.00					
No	Clinical sub-outcome only, does not affect costs/utilities	Musculoskeletal	1.41					
No	Clinical sub-outcome only, does not	Neuropsychiatric	1.37					

	affect costs/utilities							
No	Clinical sub-outcome only, does not affect costs/utilities	Ocular	1.23					
No	Clinical sub-outcome only, does not affect costs/utilities	Peripheral vascular	1.21					
No	Clinical sub-outcome only, does not affect costs/utilities	Premature gonadal failure	1.00					
No	Clinical sub-outcome only, does not affect costs/utilities	Pulmonary	1.31					
No	Clinical sub-outcome only, does not affect costs/utilities	Renal	1.83					
No	Clinical sub-outcome only, does not affect costs/utilities	Skin	1.14					
		Abbreviations: CI, confidence interval; SD, standard deviation; SE, standard error; ST, Standard Therapy						

b) Please update the PSA to incorporate all parameters used in the model, if necessary.

Response: This is not necessary as all parameters subject to uncertainty were varied in the PSA.

Section C: Textual clarification and additional points

C1. Thank you for providing the full CSRs for the BLISS studies. Could you provide CSRs for the trials listed on page 36 of the CS? This will supplement the details provided in the CS appendices.

Response: These have now been provided.

C2. Please could you provide PDFs for all the references in the appendices including those cited as ‘data on file’.

Response: These have now been provided.

C3: We are unable to open the PDF file appendix C (SmPCs and EPARs) as an error message occurs that the file is damaged. Please can you provide another version of this file that can be opened.

Response: An undamaged version of the file has been shared.

C4: We are receiving error messages when opening appendices D, L, M and N (including redacted version) which states that there is unreadable content in these documents. Please could you provide versions of these documents that can be opened without this error message and upload any embedded documents in the appendices separately.

Response: Embedded documents have been removed and these files will be shared.

C5: We are unable to open the extracted reference files from the reference pack included in the submission. Please can you try to upload these files in a different format so that we are able to open them.

Response: The reference pack will be reshared without being added to an archive folder to avoid errors with access.

References

1. Furie R, Rovin BH, Houssiau F, Malvar A, Teng YKO, Contreras G, et al. Two-Year, Randomized, Controlled Trial of Belimumab in Lupus Nephritis. New England Journal of Medicine. 2020;383(12):1117-28.
2. NHS England. Clinical Commissioning Policy Rituximab for refractory Systemic Lupus Erythematosus (SLE) in adults and post-pubescent children [200402P] 2020 [Available from: https://www.england.nhs.uk/wp-content/uploads/2020/07/Rituximab-for-refractory-Systemic-Lupus-Erythematosus-in-adults-and-post-pubescent-children-2.pdf.
3. Fanouriakis A, Kostopoulou M, Alunno A, Aringer M, Bajema I, Boletis JN, et al. 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. Annals of the Rheumatic Diseases. 2019;78(6):736-45.
4. van Vollenhoven R, Voskuyl A, Bertsias G, Aranow C, Aringer M, Arnaud L, et al. A framework for remission in SLE: consensus findings from a large international task force on definitions of remission in SLE (DORIS). Annals of the Rheumatic Diseases. 2017;76(3):554-61.
5. Franklyn K, Lau CS, Navarra SV, Louthrenoo W, Lateef A, Hamijoyo L, et al. Definition and initial validation of a Lupus Low Disease Activity State (LLDAS). Annals of the Rheumatic Diseases. 2016;75(9):1615-21.
6. Polachek A, Gladman DD, Su J, Urowitz MB. Defining Low Disease Activity in Systemic Lupus Erythematosus. Arthritis Care Res (Hoboken). 2017;69(7):997-1003.
7. Petri M, Kim MY, Kalunian KC, Grossman J, Hahn BH, Sammaritano LR, et al. Combined Oral Contraceptives in Women with Systemic Lupus Erythematosus. New England Journal of Medicine. 2005;353(24):2550-8.
8. Buyon JP, Petri MA, Kim MY, Kalunian KC, Grossman J, Hahn BH, et al. The effect of combined estrogen and progesterone hormone replacement therapy on disease activity in systemic lupus erythematosus: a randomized trial. Ann Intern Med. 2005;142(12 Pt 1):953-62.
9. Merrill JT, Neuwelt CM, Wallace DJ, Shanahan JC, Latinis KM, Oates JC, et al. Efficacy and safety of rituximab in moderately-to-severely active systemic lupus erythematosus: the randomized, double-blind, phase II/III systemic lupus erythematosus evaluation of rituximab trial. Arthritis Rheum. 2010;62(1):222-33.
10. Sutton EJ, Davidson JE, Bruce IN, editors. The systemic lupus international collaborating clinics (SLICC) damage index: a systematic literature review. Seminars in arthritis and rheumatism; 2013: Elsevier.
11. Urowitz MB, Ohsfeldt RL, Wielage RC, Dever JJ, Zakerifar M, Asukai Y, et al. Comparative analysis of long-term organ damage in patients with systemic lupus erythematosus using belimumab versus standard therapy: a post hoc longitudinal study. Lupus Science & Medicine. 2020;7(1):e000412.

NICE National Institute for Health and Care Excellence

Patient organisation submission

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (Review of TA397) [ID1591]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type. [Please note that declarations of interests relevant to this topic are compulsory].

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	[redacted]

2. Name of organisation	LUPUS UK
3. Job title or position	[redacted]
4a. Brief description of the organisation (including who funds it). How many members does it have?	LUPUS UK is the only national registered charity supporting people affected by lupus. The charity produces high-quality information for patients, carers, employers and clinicians. Through volunteer-led regional groups the charity provides support group meetings and raises awareness of the disease within local communities. LUPUS UK also funds medical research and Specialist Lupus Nurses in UK hospitals. The charity has approximately 3,500 subscribed members, however, we are here for all people affected by lupus and therefore engage with many more people with the disease in the UK.
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.] If so, please state the name of manufacturer, amount, and purpose of funding.	LUPUS UK was awarded a grant of £20,000 of restricted funding from GlaxoSmithKline in August 2020. This funding was to allow the charity to develop and provide a series of interactive virtual patient education seminars. This is part of LUPUS UK's digital outreach initiative in response to the COVID-19 pandemic to ensure that lupus patients are still able to access important patient education and support throughout social distancing restrictions. The virtual patient education seminars are being developed and produced independently from GSK and they have no editorial oversight of the contents. The virtual seminars are unrelated to belimumab. LUPUS UK was also awarded a grant of £5,000 from GlaxoSmithKline in February 2020 to assist with the distribution costs of our new book, "Lupus – Diagnosis & Treatment 2020 Edition" – an important education and information resource for clinicians.

4c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
5. How did you gather information about the experiences of patients and carers to include in your submission?	LUPUS UK conducted an anonymous online survey which was completed by 67 respondents between 30/09/2020 – 30/10/2020. The survey was open to people living with SLE and their carers in the UK. The survey asked a range of questions about the experiences of living with lupus and treatment. Views were collected regarding current standard therapy and belimumab (Benlysta). Results from a previous LUPUS UK membership survey were also used in this submission. The survey was completed by 2,527 patients who were members of the charity in 2014 and the results were subsequently analysed by The Arthritis Research UK Centre for Epidemiology, University of Manchester and accepted for publication in the journal 'Lupus' on 16 November 2017 - https://journals.sagepub.com/doi/10.1177/0961203317749746 Evidence was also taken from the Rare Autoimmune Rheumatic Diseases Alliance (RAIRDA) report, "Reduce, Improve, Empower" published in February 2018 and available at https://rairdaorg.files.wordpress.com/2020/06/rairda-survey-report-2018.pdf . The report followed a survey completed by 2,101 RAIRD patients, of which 1,098 reported having a diagnosis of lupus. The final draft of the submission was circulated to LUPUS UK Trustees to review and provide additional comments. Our Board of Trustees is entirely formed of people with personal lived experiences of lupus and their close family.
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?	SLE is a disease which varies significantly in presentation, is often unpredictable and difficult to successfully manage with medication. In our survey of LUPUS UK members, respondents reported fatigue (81%) and joint pain/swelling (60%) as the most difficult symptoms to live with. In our online survey, respondents reported that the most challenging aspects of living with lupus are the symptoms (particularly fatigue and joint/muscle pain), the impact on their ability to work and their mental wellbeing. *"I've had to develop a very strong mindset as I'm living in constant pain, plus having treatments that don't work and people telling you that you don't look sick and you should be working full time can all make you feel depressed."* *"The impact of lupus on me is such that I have never been able to work and have had to claim disability benefit. It has had a huge impact on my life, which can feel like a never-ending struggle against symptoms and fatigue."*

	The symptoms of lupus can limit a person's mobility and independence. In our recent online survey over 58% of respondents indicated that they require assistance with household care, over 43% require assistance with mobility and 1-in-3 said they need assistance with their personal care. *"I struggle with day to day activity, such as preparing food, and personal care. Household tasks are only possible on a good day."* *"There are things I can't do that my husband has to do but mostly I carry on with pain."* *"I need help in and out of the bath and to wash my hair. My husband has taken over all the household tasks. I would not be able to go out and about shopping without someone to lean on and guide me in the right direction."* One important area that is often impacted by lupus is a person's ability to maintain employment. LUPUS UK's member survey revealed that almost 1-in-4 respondents had retired on medical grounds and just over half were receiving welfare benefits. In our recent online survey approximately 58% of respondents indicated that they found maintaining employment 'difficult' or 'very difficult'. *"Had to retire from work 10 years early. Major financial and mental hardship as a result. Have to rely on reduced pension and demean myself for PIP."* *"I had to take ill health retirement at the age of 42 after struggling at work/home for over 20 years. I was doubted, misdiagnosed, and treated very poorly. It took a further 11 years to be 'diagnosed' by which time permanent damage has been done and I am regarded as being 'very complex' with multiple autoimmune conditions. My whole life has been devastated."* *"I've had to give up my part-time work as a trained NHS nurse due to overwhelming fatigue and joint pain, my husband has to support me".* RAIRDA's 2018 report showed that 25% of respondents to their survey indicated that either they or their partner/carer had reduced working hours as a result of their condition. A further 20% reported that either they or a partner/carer had been forced to give up working due to their condition. The social and psychological impact of having lupus is also reported as being very significant, with mental health problems such as depression & anxiety and loss of confidence/self-esteem being ranked as some of the most challenging aspects of living with the disease. In many cases, lupus presents with few visible symptoms (if any) making it difficult for family, friends, colleagues and medical professionals to appreciate the extent to which fatigue, pain and other symptoms have an impact. RAIRDA's 2018 report found that lupus patients were likely to feel isolated, with 24% feeling that way every day and 57% at least once a week. *"Living with lupus and other related autoimmune conditions is extremely hard. It is very complex (as in my case) and very individual. I feel like I am just a huge list of symptoms, problems and complications - not an actual person struggling and suffering to some greater or lesser level all the time. There is no break from it - physically or mentally. Very few people (friends, family and even many medical professionals) have an idea how complex and life changing it is - I am always being told I am too complicated. My rheumatologist is not a specialist in*

	lupus and there is no lupus nurse or specialist unit within about 100 miles of my area. It is very isolating. I have lost so much and do fear for the future." Whilst we received very few responses to our surveys from carers, we know anecdotally that the impact of caring for someone with lupus can be significant. This may be especially true for those with severe lupus that hasn't responded well to standard therapy – those who could potentially benefit most from belimumab (Benlysta). Fatigue, pain and weakness can be limiting factors in the mobility and capacity for activities for someone living with lupus. This often means a partner or carer will need to provide additional assistance with transport for medical appointments and essential personal and household care. As indicated in the 2018 RAIRDA report, some partners/carers either reduce their working hours or give up working to support the person with lupus. RAIRDA's 2018 report indicated that 50% of lupus patients feel that their condition has a negative effect on their family.
Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	Standard therapy is not always effective in controlling the symptoms of lupus and may not be tolerated well by all patients. Most currently available immunosuppressives are found to be significantly more effective than placebo in a little over 50% of individual lupus patients. *"When you've tried almost all the available treatment options there is the fear that you're running out of options and your lupus is starting to become untreatable. That fear is even bigger when you have lupus nephritis as your kidneys are suffering and time is running."* *"I've been treated with hydroxychloroquine, methylprednisolone, azathioprine, methotrexate and mycophenolate. I have had different reactions to each one; some I can tolerate, some not. I've had recurrent infections of all kind when taking them."* The side-effects from currently available treatments often have a significant impact on the lives of lupus patients. Steroids are renowned for their many side-effects with weight gain and changes to sleeping patterns being reported as the most difficult side-effects to tolerate. Other medication side-effects reported as being most difficult to tolerate by people with lupus include fatigue, nausea, hair loss, and changes in mood. Many people with lupus will have been prescribed several different medications to try and manage their condition. It is often the case that a treatment does not sufficiently control symptoms or causes adverse effects that cannot be tolerated. Many lupus treatments can take months before the full benefit may be experienced, meaning a significant period with a lower quality of life. *"My current treatment is not giving me a very good quality of life; I struggle with significant symptoms but owing to coronavirus disruptions I have not been able to start on anything else."* Most immunosuppressive and biologic treatments are not safe during pregnancy and breast-feeding; this is especially relevant

	given the proportion of cases diagnosed in young women who may be planning or considering a family. In our recent online survey, the areas that respondents most reported their treatment as having a negative impact were "managing current treatment (collecting prescriptions, checking for interactions etc.)", "social activities" and "changes to diet/lifestyle". *"Contraindications have caused friction and difficulties between different specialist hospital departments - I am stuck in the middle and often left with sorting things out. There is little co-ordination/communication between them - This has made everything more complicated and stressful than it should or needs to be. Again, my physical and mental health have been adversely affected. I regularly have to literally spend days phoning around regarding prescriptions, explaining complex issues, locating medications etc."* The provision of care for lupus patients in England is inconsistent. Many patients living closer to larger cities and able to access a specialist centre with multidisciplinary clinics report a much higher level of satisfaction with the care they receive. Patients without access to specialist lupus services may experience additional consultations with multiple specialties, poor communication between clinicians, a lack of a coordinated care plan and barriers to accessing some treatments, such as biologic therapies. RAIRDA's 2018 report indicated important findings related to treatment of people with RAIRDs (including lupus): • Only 34% of patients received all their routine care at the same hospital in the past year. • Two-thirds of patients routinely visit two or more hospitals for their care, with 1-in-20 visiting five or more hospitals in the past year for their care. • 8% of patients reported regular journeys of two or more hours for their treatment. • 93% of patients see clinicians from multiple specialisms as part of their routine treatment, yet among those people, less than 1-in-5 were able to see multiple specialists at a joint clinic. • 46% of lupus patients stated that they do not feel the different professionals involved in their care have a plan for their treatment. In our recent online survey, we asked, "How would you rate your overall treatment and care from the NHS for your lupus?". On a scale from one (very poor) to ten (very good) the average score was six. Approximately 25% of respondents in our recent online survey stated that their current treatment was a "large" or "very large" interruption to work/study.
8. Is there an unmet need for patients with this condition?	Earlier diagnosis of lupus is needed to allow for faster intervention with treatment, the prevention of damage accumulation and improved outcomes and quality of life for patients. LUPUS UK's member survey indicated that the average length of time to obtain a diagnosis after the initial onset of symptoms was 6.4 years. In addition to delays in diagnosis, people with lupus often experience delays in seeing a specialist. RAIRDA's 2018 report found

	that just over half (54%) of patients were seen by a specialist in under three months, while almost a quarter (22%) reported that they had waited longer than six months for their specialist appointment. These findings suggest that waiting time targets continue to be missed for people with rare autoimmune rheumatic diseases (including lupus). Additionally, there is real concern that these targets themselves do not adequately reflect the need for prompt diagnosis of rare diseases to reduce the risk of irreversible organ damage occurring prior to treatment. Fatigue is commonly reported as the symptom which is most difficult to manage for people living with lupus. Standard therapy is generally ineffective at alleviating this sometimes-debilitating symptom. *"Fatigue is ongoing, and no medication has helped. I have three/four days a week; the rest I am resting. Losing my job due to lupus and UV sensitivity was extremely traumatic."* Cardiovascular disease (CVD) is the leading cause of mortality in lupus patients, with the condition representing a significant risk factor. Effective treatments to control lupus inflammation and reduce the development of CVD are essential in the length and quality of life for those with SLE. Many existing treatments used in lupus generally suppress the immune system and leave patients more at risk from infection. Treatments that effectively control the disease whilst not making patients vulnerable to infection are needed. Stratified medicine is needed for lupus because of its heterogenous nature and unpredictable response to treatment. Stratified medicine could aid newly diagnosed patients in accessing the treatment most likely to be effective for them earlier, saving months or years of trial and error with side-effects, poorly managed disease, and disruption to their life. *"They don't 'cure' lupus, and you still have symptoms even on random days and the accrued damage is still happening. Your lupus damage keeps progressing. Remission is rare. You will need to keep monitoring your lupus, which means that you can have another flare anytime. Thus, this can affect your work and family life again when you thought that everything was stable."*
Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	One of the respondents to our online survey who has personal experience of belimumab (Benlysta) had this to say; *"The belimumab infusions are much less time consuming than having other infusions. It is the only treatment that has improved my symptoms in 3 years. I am 32 years old, live on my own and work full time. Lupus has taken over my life and belimumab is the only thing that has given me some of my life back to allow me to do things other than work and sleep. The thought of this drug being taken away is a huge source of anxiety for me and what my life would become again without it. The 4 weekly treatment is completely manageable, and I receive excellent care when I have it."* It has been demonstrated in trials that belimumab, when used alongside standard therapy, can produce a significant reduction in disease activity (https://lupus.bmj.com/content/3/1/e000118). It has also been shown in randomised controlled trials that this effect may be more pronounced than standard therapy alone (https://www.nejm.org/doi/full/10.1056/NEJMoa2001180). This

	treatment addresses an unmet need for patients with severe SLE which does not respond to current treatments and can result in significantly reduced quality of life and premature mortality. *"I haven't felt well since being diagnosed seven years ago. My other drugs have helped but I have lost a lot of bone mass from steroids. In order to work I have to limit what I do in my spare time. Going out is difficult; by the evening I am too tired and during the day I can only manage a couple of hours. I feel ill after exercise. A new drug which could control fatigue and malaise would be life changing. I know I am running out of drug options, which is difficult when you have a potentially life-threatening disease. Benlysta could be life-saving if my condition deteriorates."* For those patients who do not respond sufficiently well to standard therapy alone, belimumab could assist them in reducing their steroid dose over time, helping to reduce the risk of side-effects and future comorbidities. *"Belimumab has the potential to act as a steroid sparing agent and allow withdrawal of steroids. My current treatments help but do not fully control my disease and I have not been able to withdraw steroids. I have concerns about my long-term bone health due to cumulative steroid use."* *"I currently have belimumab infusion once a month. I'm on my 5th dose and touch wood it the only drug which has worked for me. It's brought my DNA levels down from 450+ to 17. I have had rituximab, cyclophosphamide and neither of them worked for me."* For those patients who experience a significant improvement in the management of their condition because of belimumab, it could have a massive impact on their quality of life. It could potentially mean they are better able to continue in employment and experience further benefits to their social and psychological wellbeing. With their lupus better controlled, it could reduce the number of hospital visits and admissions they may otherwise have needed which would be a positive change for them and their family/carers. *"I hope it would relieve the burden on my family as they have to do a lot to support me in my everyday living."* *"I am not yet 40. Lupus is severely impacting my career, due to fatigue, pain, and cognitive function. I have already had to reduce my hours, and it is likely I will have to take a 70%+ pay cut or give up work entirely in the next few years. Any treatment that can reduce my symptoms sufficiently to keep me financially independent has to be a good investment. Not to mention the impact on my mental health, relationships, and life outside of work."* *"Excellent drug - has had a significant positive impact on my daughter's health."* It is important to remember that many of the patients who would be considered for treatment with belimumab have highly active SLE which has not responded adequately to current standard treatment. It therefore provides an additional treatment option and hope.

Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	Lupus presents differently in each patient and their response to treatments can vary similarly. A treatment that works well for one patient will not necessarily work for another. Adverse reactions to medications are seen in many people with SLE, resulting in a need to switch to another treatment option. It is therefore likely that some patients will not be able to tolerate belimumab or will not respond as hoped. A few patients made comments in our recent online survey expressing concern about possible side-effects from belimumab. We expect that this trepidation would be applied to any similar new technology. *"Lupus has had a devastating impact on my life. It has crippled me. Any new treatments welcome but side effects need to be taken into account and discussed fully with the patient and a joint decision should be made between consultant and patient."* *"Reducing the B cell count would possibly make the lupus patient at an increased risk of a serious infection."* *"Might increase vulnerability to infection?"*
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.	At present there are strict criteria for SLE patients to access belimumab. They must: • Be autoantibody-positive • Have positive anti-double-stranded DNA and low complement • Have a disease activity score (SELENA-SLEDAI) greater than or equal to 10 despite standard therapy Therefore, it is a small sub-set of people living with SLE in the UK who are currently eligible for this treatment. *"Treatment options for those with lupus appear to be fairly limited with few new treatments appearing. Many of us develop adverse drug reactions and this further reduces treatment options - some options are not available unless disease is extremely active - leaving some with no medication options (other than symptom control) and thus exposed to potential progression to organ damage. Is it appropriate that some have to wait to suffer actual damage in order to access medications? Shocking!"* *"Belimumab should not be limited to patients with dsDNA antibodies and low complement. Patients with SLE can be very sick without having both these criteria and there are multiple autoantibodies associated with SLE. Clinical trials have to have strict entry criteria but NICE takes these criteria and applies them to funding decisions when in real life patients don't fit neatly into categories. NICE should be more pragmatic about funding of treatments for SLE and not strictly define subtypes, as doing so rations treatments and patients lose out on therapies."* *"I have terrible symptoms of lupus and it impacts on my life every day. However, I don't qualify for a drug that has been specifically created*

	for use in lupus. Instead I take a concoction of medications including very long-term use of high doses of prednisolone and my symptoms are still not under control. Surely the cost of all these medications (7 different items in total, including rituximab) would be better spent on trying something that has been approved for use in lupus. It has even been suggested to me by a Dr to come off my medication, make myself really poorly and then I might (might!) qualify. The impact of living like this for 20 years has lost me the best days of my life."
Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	Currently IV belimumab is only administered at specialist centres. A Rare Disease UK study (https://www.raredisease.org.uk/media/1601/centres-of-excellence.pdf) has previously shown that only 27% of patients with rare diseases are cared for in specialist centres. This presents a barrier to access for some patients who may live a considerable distance from a specialist centre or have difficulty travelling due to their ill-health and/or disability. As such, those living in more remote parts of the country or with mobility issues may be less likely to benefit from this treatment if it continues to be administered only at specialist centres by IV infusion. Increasing the availability of subcutaneous belimumab could help to resolve this. The administration of IV belimumab at specialist centres may also present a barrier to access for some patients of working age. The 4-weekly, hour long infusions (with additional travelling time) can incur a considerable amount of time away from the workplace. Many lupus patients are of working age and may choose not to access this treatment if they fear their employment may be put at risk by regular absences. Increased availability of subcutaneous belimumab would involve fewer hospital attendances and potentially improve access for people who could benefit from this technology. SLE affects people of all ethnic groups but is more prevalent in people of African, Caribbean and Asian heritage. People from these ethnic groups are also more likely to experience severe disease and higher rates of premature mortality. There are some important considerations that must be made for these groups of patients: • Double-stranded-DNA antibodies are less common in patients of African descent, so it could be perceived as discriminatory to stipulate dsDNA antibody positivity as a criterion and not consider other lupus-related antibodies. • People from these ethnic groups are already at a high risk of developing diabetes and hypertension. It should be considered whether steroid-sparing treatments such as belimumab could have additional advantages over standard treatments by reducing some adverse effects and risks of comorbidities. SLE disproportionately affects women and commonly presents in those of childbearing age. Cyclophosphamide is still used to treat severe lupus with major organ involvement despite presenting a risk of infertility. The role of belimumab in the treatment of young women should be carefully considered.

Other issues	
13. Are there any other issues that you would like the committee to consider?	The current criteria for belimumab prevent many people with SLE from accessing the treatment and should be reconsidered. Requiring both positive anti-dsDNA antibodies and low complement as well as clinical disease activity could be unnecessarily restricting the treatment from some patients who could respond well and experience significant improvements in their quality of life. The patients currently being excluded because they don't meet all of the criteria will end up having more corticosteroids and being re-tried on standard therapies that have already failed (and therefore have a low probability of response) with significant increased risks of side-effects and future comorbidities. We are aware that uptake of belimumab under the Managed Access Agreement was much lower than anticipated. It is important to understand the reasons for this low uptake. • Was this due to administration being restricted to specialist centres? • Were the eligibility criteria too restrictive? • Is there a lack of clinician education and awareness of this treatment? • Were patients in local hospitals receiving appropriate referrals to specialist centres? In late October 2020 we were informed by a patient that, due to a change in the price of sub-cutaneous belimumab injections, their hospital had to seek an increase in budget and could not issue further injections until it was approved. This could result in some patients returning to monthly outpatient infusion clinics during the COVID-19 pandemic.

Key messages

14. In up to 5 bullet points, please summarise the key messages of your submission:

- SLE often has a significant impact on the lives of people with the disease and their close family.
- Standard therapy is not effective at controlling symptoms for all patients.
- Most current standard treatments are ineffective at treating fatigue.
- Belimumab offers an additional treatment option, representing hope for those with active disease who do not respond to standard therapy.
- The criteria for accessing belimumab need to be carefully reviewed to ensure health equality and improved outcomes for more people living with SLE in England.

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.

.....

NICE National Institute for Health and Care Excellence

Professional organisation submission

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (Review of TA397) [ID1591]

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 for Rheumatology**

3. Job title or position	[redacted]
4. Are you (please tick all that apply):	☒ an employee or representative of a healthcare professional organisation that represents clinicians? ☒ a specialist in the treatment of people with this condition? ☒ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5a. Brief description of the organisation (including who funds it).	**The British for Rheumatology is the leading UK specialist medical society for rheumatology and musculoskeletal care professionals.**
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.] If so, please state the name of manufacturer, amount, and	**In the past 12 months, BSR has received funding from the manufacturers of several comparator products. We received funding from Pfizer ([redacted]), Roche Pharmaceuticals ([redacted]), Celltrion ([redacted]) and Sandoz GmbH ([redacted]) relating to our Rheumatoid Arthritis register.** **We also received costs/fees from Pfizer, Roche and Celltrion relating to the cancellation of our annual conference.**

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.)	The aim of Belimumab treatment is to reduce disease activity in patients with Systemic Lupus Erythematosus (SLE), thus improving symptoms and quality of life, reducing the risk of permanent organ damage and limiting exposure to alternative therapies (particularly corticosteroids) that are associated with long-term toxicities.
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	There is no 'gold standard' method of defining a clinically meaningful response in lupus, and the situation is increased in complexity but the use of two alternative validated disease activity scoring systems which come with their own advantages and disadvantages. In recognition of these problems, the group of currently published belimumab trials in non-renal lupus (bearing in mind this is a disease that can present in many ways) used a rather technical composite definition of 'significant' response, called the SLE responder index 4 (SRI4), which requires 1) an improvement in the SELENA-SLEDAI disease activity score by 4 or more points, 2) no new BILAG A scores and no more than one new BILAG B score and 3) no deterioration in a physician global assessment by ≥ 0.3 points. Roughly speaking, this would correspond to the improvement of one major clinical manifestation of lupus, for example arthritis, without any major new clinical manifestations developing.

	Our view is that SRI4 is too complex to be used in routine clinical practise. SELENA-SLEDAI heavily weights on severe-rare manifestations, contains elements that are not routinely available in UK NHS practice (urine sediment analysis for example) and tends to gloss over many of the subtleties of presentation and response that are seen in the real world patients and better encapsulated by the BILAG index. On balance for lupus in general we would probably adopt a slightly simpler definition of clinically meaningful response which would also fall in line with the current Rituximab policy which is either 1) Fall in SLEDAI score by ≥ 4 points or 2) Fall in one or more baseline BILAG A score or two or more BILAG B scores (providing this is not requiring unacceptably high levels of corticosteroid to sustain). It would be reasonable to assess this at 3 and 6 months, with treatment discontinued if improvement not met at 6 months. Lupus nephritis as a specific manifestation requires a different approach, because assessment is focussed on blood and urine abnormalities. In the recently published study looking at lupus nephritis, a significant response was considered one that improved renal parameters to a level at which previous data suggest the risk of progressive chronic kidney disease is significantly reduced (namely a urine protein:creatinine ratio <70 mg/mmol + a serum creatinine no more than 20% below baseline). Assessment was made after 1 year of treatment reflecting the fact urine parameters can take time to respond and that the evidence around risk reduction for chronic kidney disease is at this one year endpoint. If NICE agree to usage in patients with newly presenting renal disease as per the clinical trial, then the above endpoint would be applicable. Further challenges arise when considering using belimumab for relapses in the context of more advanced chronic kidney disease where urine protein may be chronically elevated due to renal scarring. An alternative pragmatic figure for urine protein improvement may have to be adopted, for example ‘a fall in urine protein to within 25% of the pre-relapse baseline’
8. In your view, is there an unmet need for patients and healthcare professionals in this	Yes, there remains a very significant un-met health need in lupus. Current medication will induce a long-lasting, low-disease activity state in less than half of patients with lupus – the majority will either have ongoing disease activity or period of remission punctuated by frequent relapse. We know that disease activity and relapse is associated with the development of progressive organ damage and reduced quality of life. Still around 10-20% of patients with lupus nephritis will progress to end-stage kidney disease and the

condition?	massive quality of life and healthcare costs associated with dialysis and/or transplantation. While disease activity can in many cases be controlled by high dose steroids, we are aware that these also come with increased risk of irreversible organ damage and major healthcare problems such as type 2 diabetes mellitus and osteoporosis. We would comment that the current guidelines around belimumab usage are very strict and exclude a large cohort of patients with moderate to severe disease activity despite immunsuppressants (a review of UK registry data suggested that only 13% of patients with disease activity severe enough to consider biologic treatment were actually eligible for belimumab). Many of these will be requiring inappropriately high doses of corticosteroid to maintain disease stability. that may benefit but are currently excluded
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	Currently, patients with lupus are managed long-term in secondary healthcare settings, with more severe disease such as nephritis focussed in a smaller number of specialist centres. There are national guidelines (British Society of Rheumatology/NICE accredited) which provide a good framework for decision making in lupus. Lupus is managed with a combination of corticosteroids, antimalarials (hydroxychloroquine) and conventional immunosuppressants (e.g. methotrexate, azathioprine and mycophenolate). More refractory cases are treated with belimumab, cyclophosphamide and Rituximab (within the parameters defined by NICE, NHSE and BSR pathways). For lupus in general, the evidence base to support these treatment strategies is weak – there are few randomised controlled trials and trial design and interpretation remains complex. Many of these treatments have been around for many years and are used on the basis of expert opinion, case-control and cohort data as reviewed in BSR guidelines for management of SLE. Lupus nephritis as a specific manifestation is slightly better evidenced, particularly the roles of mycophenolate, cyclophosphamide and azathioprine.
• Are any clinical guidelines used in the treatment of the	There are a number of guidelines available. For lupus in general, there are 2018 British Society of Rheumatology Guidelines (NICE accredited) for the management of systemic lupus erythematosus – rheumatology.org.uk/practice-quality/guidelines/

condition, and if so, which?	For lupus nephritis, there are three sets of guidelines which are highly concordant: EULAR/ERA-EDTA 2019 guidelines – ard.bmj.com/content/79/6/713 American College of Rheumatology 2012 guidelines – Arthritis Care Res (Hoboken) 2012;64:797-808 KDIGO (Kidney Diseases Improving Global Outcomes) 2012 with updated evidence summary in 2018. For usage of Rituximab and Belimumab in England we follow the prescribing rules set in place by NHS England and NICE respectively.
• 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.)	Broadly speaking, there is consensus about overarching treatment principals, although there may be clinician-specific differences, in particular in aspects such as corticosteroid dosing. It seems likely that within specialist centres care is more concordant, and there is a well establish network of lupus-interested clinicians (the BILAG group) that would perpetuate that tendency. A recent large multicentre audit (rheumatology.org.uk/practice-quality/audits/lupus) indicated that most variation from guideline-based practice occurred in smaller regional centres.
• What impact would the technology have on the current pathway of care?	This really depends on whether current NICE policy is changed as a result of this technology approval. We would argue, as outlined in other sections or this document, for continued usage, with some changes to the policy to expand usage to renal disease and to patients with active disease who may not meet the current stipulation of both dsDNA antibody positivity and low C3/C4 complement levels.
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	As answered above. Belimumab is currently used within the guidance issues by NHSE/NICE. We would argue, as outlined in other sections or this document, for continued usage, with some changes to the policy to expand usage to renal disease and to patients with active disease who may not meet the current stipulation of both dsDNA antibody positivity and low C3/C4 complement levels.
• How does healthcare	Belimumab as used currently (i.v. formulation) is the only treatment for lupus that requires regular and

resource use differ between the technology and current care?	indefinite intravenous administration, with the associated inconvenience for patients and associated administrative and financial costs for the providing hospital. The current pathway, which mandates provision within a specialist centre, does place barriers to acceptance. Many patients live some distance from specialist centres and refuse treatment on the basis of the time and expense of travel. Providing these treatments when resources are stretched (this has only been amplified by the current COVID situation) can be a burden. We would strongly support the adoption of subcutaneous Belimumab for lupus in general, since this also now comes with randomised control trial evidence. Emergency measures put in place at the beginning of the COVID outbreak allowed us to switch patients to s.c. belimumab to help reduce patient footfall. Anecdotal evidence of specialist colleagues has been that this has been well received by patients and reduced hospital attendances.
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	It seems reasonable for the current arrangements for authorising belimumab usage to remain in place (discussion with a regional MDT via network arrangements). Clearly, this drug should be supervised within secondary care, but I do not see the rationale for insisting the drug is actually administered in specialist centres. All secondary care rheumatology units are familiar with the safe administration of biologic therapies and for stable patients infusions of other agents is often successfully provided in patients homes. I see absolutely no reason for making belimumab a 'special case' as far as the practicalities of administration are concerned. Regional networks should be allowed to innovate and find the most appropriate solutions for their areas.
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	The only investment required would be if belimumab usage increased significantly to the point that ongoing insistence on delivery at a specialist centre began to overwhelm the specialist centres infusion capacity. Bearing in mind the intravenous version of this drug is given by indefinite monthly infusion, this does present a risk, but this could be mitigated by authorising usage of the subcutaneous formulation (as outlined in answer to previous question).
11. Do you expect the technology to provide clinically meaningful benefits compared	The available clinical trial evidence suggests that clinically meaningful benefits are certainly seen over and above standard care for selected groups of patients. There are also extensive reports of patients obtaining benefit in other countries with less stringent criteria for administration (according to original FDA and EMA licenses). Evidence also exists to point to a reduction in accrual of long-term damage in patients receiving belimumab compared with standard of care (Organ damage in patients treated with belimumab versus

with current care?	standard of care: a propensity score-matched comparative analysis. Urowitz et al. Annals of Rhuematic Diseases 2019)
• Do you expect the technology to increase length of life more than current care?	There would certainly be no direct evidence to support the suggestion this is a life-prolonging treatment. There is indirect evidence to suggest it does slow the accumulation of organ damage (referenced above). There is also evidence it supports reduction in corticosteroid usage (itself asscieted with excess mortality and morbidity). The belimumab lupus nephritis trial primary outcome was selected to match parameters that have been shown in other studies to be associated with a slower rate of progression to chronic kidney disease. It is also relevant to say there that there is no evidence of side-effects likely to limit life with this drug.
• Do you expect the technology to increase health-related quality of life more than current care?	HR-QOL outcomes have been included in all the current clinical trials for belimumab and there appears to be statistically and meaningfully improvement.
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Evidence in general suggests more benefit in patients with more active lupus. Patients with high disease activity (SLEDAI score >8) who are antibody positive (ANA or dsDNA historically) with either positive dsDNA antibodies or low C3/C4 (at the time a decision is made to use belimumab). Patients from a wide-range of ancestral backgrounds have been included in the available lupus clinical trials, but more data would be ideal as evidence for other therapeutics suggest there may be ancestry-specific differences in response Belimumab is not appropriate for women who are pregnant or actively trying to fall pregnant as the effects on foetus are unknown and drug will cross placenta from around 16 weeks gestation.
The use of the technology	
13. Will the technology be easier or more difficult to use	Please refer to more detailed answer in section 10. Belimumab in its i.v. formulation requires indefinite monthly intravenous administration as a day case patient in hospital and is therefore more difficult to

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.)	support than current standard of care therapies.
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	The rheumatology/renal community in general have shown willingness to use belimumab within the terms of its NHSE approval (appropriate level of disease activity at baseline, prior treatment attempts with other agents, ongoing assessment of response and registry enrolment). We are sure the clinical community would continue to respect rules, but would look to a change in the current set of rules based on new evidence and on-going experience. In particular: In response to recent randomised-control trial evidence as an add on-therapy in early lupus nephritis, we would look for approval to use for lupus nephritis (currently not within NHSE rules). In the light of randomised controlled trial evidence for subcutaneous belimumab and as we look to increase acceptance of therapy and to limit the use of secondary care healthcare resources (both in relation to COVID and because of resource implications), we would like to see approval to use subcutaneous belimumab more widely long-term and not just during Covid-19 pandemic. Thirdly, although we agree it is appropriate to restrict the use of belimumab to patients with more active disease, we would like to see the requirement for both dsDNA positivity *and* low complement levels to be revised. We appreciate the evidence from post hoc analysis that patients with SELENA-SLEDAI ≥ 10, high dsDNA antibodies and low complement levels may derive most benefit from belimumab, but in the 'real world' there are certainly patients with exceptionally

	active lupus who miss out on treatment because of this restriction and the treatment should be available if there is either high dsDNA antibodies *or* low complement with high disease activity by SELENA-SLEDAI, as each of these was associated with response in the phase 3 trials (in univariate and multivariate analyses). Each of these was a pre-determined end-point but the combination as currently used was an ad-hoc analysis devised after the end of trials and treatment should not be restricted to only this combination, although the principal of serological activity with clinical activity helps to ensure that the patient has disease that is likely to respond to the mechanism of action of belimumab. This was not an inclusion criteria in any of the clinical trials and is certainly not of relevance if we look to usage in renal disease (where disease activity should be determined by renal biopsy).
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?	We do not have the expertise in health economics to answer this question. However, we would point out that lupus is very heterogeneous in presentation, but also in terms of long-term consequences. A small number of patients do still die. A reasonable proportion of patients develop end-stage kidney disease and require dialysis or a transplant. We are aware lupus is associated with an acceleration of cardiovascular events, perhaps by as much as 20 years in comparison with general population. We know that some other items of 'damage' due to lupus and its treatment can have profound consequences – for example I have patients who have required 4 or more joint replacement due to the development of avascular necrosis (severe joint damage due to lupus and steroid treatment).
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 improve the way that current	Lupus treatment has evolved by incremental steps. Outcomes are clearly better that 20 or more years ago. Belimumab provides a further incremental benefit. It is one of only three treatments we have that actually holds a licence for use in lupus. Clinical trial evidence suggests that significant benefits are seen for a proportion of patients. With any chronic and complex disease there clearly remain uncertainties about the overall impact on long-term outcomes, but this evidence is impossible to provide unless we can bring the treatment into more widespread usage. Existing evidence suggests there is a reasonable correlation between short-term outcomes in lupus (disease activity and steroid dosage) and long-term damage (the really health affecting and costly consequences such as renal failure), so it is reasonable to assume this long-term benefits are likely is short-term responses are observed.

need is met?	
• Is the technology a 'step-change' in the management of the condition?	Belimumab is a step change in so far as it has emerged as one of only three licenced therapeutics for lupus and has a now growing record of successful trial data and post-licensing observational data. Clinically, it is probably best regarded as an important incremental treatment strategy.
• Does the use of the technology address any particular unmet need of the patient population?	Refer back to other answers – lupus is a long-term illness with no definitive cure. It addresses the needs of the substantial proportion of patients with lupus who have non-responding or relapsing disease with current standard of care treatment and that are dependent on corticosteroids to treat and prevent active disease which cause many well-known complications.
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 side effect profile that has emerged from the multiple trials is good – certainly no more than any other available standard of care option and probably substantially less than commonly used higher dose steroids.
Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	The clinical trials certainly represent a substantial subset of patients seen in UK clinical practice, but of course not all. The trials for non-renal lupus tended to look at a population of patients who had had lupus for 5-10 years and had on-going moderate to severe disease activity despite steroids and/or immunosuppressants. (NHSE guidelines have added additional restrictions to current usage – in particular the requirement for positive dsDNA antibodies *and* low complement levels that were not required in the trials). The recent trial in lupus nephritis looked at adding belimumab in to early treatment for inducing remission alongside

	standard of care (this is recent data and there is currently no UK authorisation to use for lupus nephritis). There are of course other scenarios not directly addressed in the trials but encountered in the clinics – for example patients refractory to multiple treatment courses or patient with non-renal lupus very early on in their disease. There is an important unmet need for managing patients with severe active lupus including high SELENA-SLEDAI ≥ 10 with either high dsDNA antibodies or low complement who are not eligible for belimumab but would be likely to benefit from it. These patients were included in the trials,
• If not, how could the results be extrapolated to the UK setting?	It seems reasonable to approve belimumab as studied in the post-hoc analysis of the original clinical trials and the recent lupus nephritis trials and to continue the current approach which is to use the highly developed UK BILAG biologics registry to gather dateaon belimumab efficacy in a wider variety of 'real world' situations.
• What, in your view, are the most important outcomes, and were they measured in the trials?	It is really hard to encapsulate all the myriad manifestations of lupus within a single clinical trial end-point, but within the limits of what can be realistically achieved, the trials were comprehensive in their collection of disease activity over a reasonably good time frame (1-1.5 years blinded but then with prolonged open label extensions). What doesn't get directly captured is the development of 'damage' – which are the items of irreversible organ dysfunction/failure that occur as a long-term consequence of disease. These take many years to develop and it is simply impossible to maintain a blinded trial over the timeframe required.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	There is reasonable evidence to suggest shorter term control of disease activity and lower usage of steroids reduce the risk of organ damage, so measuring these outcomes is as good a surrogate as we have got.
• Are there any adverse effects that were not apparent in clinical trials but have come to light	There has been some data to indicate an associated with development of depression and suicidal thoughts. It is recommended patients are screened for risk as part of the overall discussion of risk and benefit before starting treatment.

subsequently?	
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	I know that NHSE are working with the UK BILAG biologics registry to audit real-world efficacy of belimumab in a UK setting, but don't know the timescale for these data to be assessed.
20. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance [TA397]?	Yes, since then there has been a randomised controlled trial in lupus nephritis (Two-year, randomized, controlled trial of belimumab in lupus nephritis. Furie et al., New England Journal of Medicine 2020), there has been a trial of subcutaneous belimumab (Efficacy and safety of subcutaneous belimumab in systemic lupus erythematosus: a fifty-two-week randomiized, double-blind, placebo-controlled study. Stohl et al. Arthritis rheumatology 2020), there has been a trial of i.v. belimumab in far-eastern populations (A pivotal phase III, randomised, placebo-controlled study of belimumab in patients with systemic lupus erythematosus located in China, Japan and South Korea. Zhang et al, Ann Rheum Dis 2018).
21. How do data on real-world experience compare with the trial data?	As mentioned in question 19, real-world data on UK usage is being widely collected, although not reported on to date. We are mindful that due to a combination of pretty stringent rules around usage and the natural tendency of cautious clinicians to use novel therapies after all else fails that a group of very refractory lupus patients may have been started on belimumab in the UK. We are reluctant for this initial cohort to decide the issue around belimumab authorisation once and for all. Actually, some relaxation of rules and some more time for experience to grow may allow the full potential of belimumab to be revealed in the real world. The published literature is full of case-reports and case-series/cohorts of belimumab being used in a variety of lupus scenarios with success, although we recognise issues related to publication bias here.
Equality	
22a. Are there any potential equality issues that should be	Not specifically – just the observation we need to be mindful that there is limited data on ancestral differences in response to therapy and previous comments that the insistence of intravenous usage at a specialist centres unfairly discriminates against people living in rural communities who must travel further

taken into account when considering this treatment?	and incur additional transport costs to receive this treatment.
22b. Consider whether these issues are different from issues with current care and why.	Comment about ancestral responses may be relevant to current care as well. Comment about discrimination based on location does not apply to other treatments as other treatments are not indefinite monthly infusions.

Key messages

23. In up to 5 bullet points, please summarise the key messages of your submission.

- There is growing clinical trial evidence to support usage of belimumab and on this basis we would seek to expand the range of scenarios in which belimumab was approved for usage in the UK. In particular:
  - We would see authorisation for use as an adjunctive treatment in active lupus nephritis
  - We would seek authorisation for long-term
  - We would seek relaxation of the requirement for dual dsDNA positivity/low complement in patients with active lupus
  - We would seek relaxation of the requirement for delivery at specialist centres only to make delivery of treatment more practical and equitable.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

……

Your privacy

NICE National Institute for Health and Care Excellence

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

.....

NICE National Institute for Health and Care Excellence

Professional organisation submission

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (Review of TA397) [ID1591]

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	**On behalf of the Renal Association**

3. Job title or position	[redacted]
4. Are you (please tick all that apply):	**x** an employee or representative of a healthcare professional organisation that represents clinicians? **x** a specialist in the treatment of people with this condition? **x** a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5a. Brief description of the organisation (including who funds it).	The leading professional body for the UK renal community, dedicated to improving lives by supporting professionals in the delivery of kidney care and research. Funded by membership fees and NHS capitation fees.
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.] If so, please state the name of manufacturer, amount, and	No

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 reduce disease activity – ideally to achieve remission but at least the lowest disease activity (minimal disease activity (MDA)) level possible in order to: a) Improve quality of life b) Reduce the risk of progression to more severe disease e.g reduce risk of a renal flare c) Prevent treatment related toxicity d) Prevent permanent damage – from disease or as a complication of treatment In light of BLISS-LN data, likely new indication upcoming for treatment of lupus nephritis.
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	Reduction in: a) Patient reported symptoms b) Disease activity scores e.g. SLEDAI reduction of >4 c) Steroid dose and ideally steroid need altogether d) If licensed for lupus nephritis – induction of remission and preservation of kidney function
8. In your view, is there an	• Absolutely – symptoms can often be controlled by steroids but at huge longterm cost to the patient in

unmet need for patients and healthcare professionals in this condition?	terms of adverse side effects that can lead to lasting damage. • Large proportion of patients have ongoing symptoms despite reasonable doses of steroids and other agents. • Already recommended for refractory moderate disease
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	Standard of care for non renal lupus would be hydroxychloroquine (first line) then steroids then a steroid sparing agent. Steroid sparing agents include azathioprine and methotrexate – both can be liver toxic, bone marrow toxic and require careful blood monitoring. Often not tolerated.
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	https://academic.oup.com/rheumatology/article/57/1/e1/4318863 British Soc Rheumatology guideline for the management of SLE in adults – 2017 EULAR guideline Fanouriakis A, et al. Ann Rheum Dis 2019;78:736–745. https://pubmed.ncbi.nlm.nih.gov/30926722/
• 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 are recommendations on best treatment but most are based on relatively poor evidence. The two key recent guidelines divide patients into having mild, moderate or severe disease and first line in mild is hydroxychloroquine which should be maintained long term. Then steroids and then second line steroid sparing / disease modifying agents. Belimumab reserved for refractory mild/ moderate disease requiring high dose steroids. Belimumab has been trialled more rigorously and in larger trials than any other medication for non renal non cerebral lupus. Steroids are used routinely at too high a dose and for too long for symptoms that are severely troubling to patients but not organ threatening. Second line agents that steroid spare e.g. azathioprine or methotrexate, are used but often late in the day and require daily tablets, regular monitoring of blood counts and liver function and are variably tolerated. All of these are very non-specific immunosuppressants. Practice , but in general far too high doses of steroid remain the most problematic

	aspect of treating the same indications in lupus that belimumab can target.
• What impact would the technology have on the current pathway of care?	Undoubtedly, if able to use early, would reduce steroid exposure, reduce need for monitoring and would improve QoL. Particularly if could continue access to sub cutaneous admin (sc) which has started in the pandemic – allows self medication, avoids many oral daily tablets
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	It's currently in use largely as IV monthly injections; would be preferable to have the choice to move to s.c. administration
• How does healthcare resource use differ between the technology and current care?	Current care – if not using belimumab at all – regular review in outpatients, frequent flares and need for monitoring increasing and decreasing doses of steroids, monitoring of steroid sparing agents such as azathioprine and methotrexate – regular blood counts and LFTs. Patients are more stable on belimumab which does not require as frequent blood monitoring, and if on IV belimumab, then all bloods could be done at same time as infusion – reduces burden on patient and clinic. If move to sub cutaneous administration – reduces cost to patient (in terms of hospital visits, IV access, time taken, risk of exposure to Covid or disruption of services by Covid) and to NHS (cheaper than IV infusions, less nurse time, can manage patients virtually for longer). Could arrange to have bloods done at local unit and liaise with central unit.
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Definitely secondary care and currently limited to centres of excellence. This is in general good for lupus patients as improves access to expert care which undoubtedly leads to reduction in steroid use. However, might be advisable to set up shared care pathways – belimumab initiated (1st month of treatment?) and sanctioned in centre of excellence for managing lupus and then set up shared care pathway with local unit for monitoring. Would encourage better management in general, likely to lead to lower use of steroids and now with Teams meetings for MDTs so well established, could set up joint MDTs to discuss all patients in a monthly MDT or two monthly MDT. Particularly should have patient back at centre or on MDT to discuss response and continuation at the end of first 6 months and thereafter at least annual review. Would raise

	standard of care for lupus patients overall and facilitate more rapid referral to centre for e.g. a kidney flare.
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	None in centres of excellence. Local units would need training – if giving IV but should not be giving first months (3 doses) doses; local units would need training in use of SLEDAI scoring.
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes. Already does – a significant group of patients who require high dose steroids to control disease and who are refractory or intolerant of aza or MTX, respond very well to belimumab. Indeed, for some it has been transformative.
• Do you expect the technology to increase length of life more than current care?	Yes. For several reasons: a) Minimises steroid use which is a major cause of damage and complications of therapy which in themselves can shorten life – eg diabetes and hypertension b) Reduces flare rate so reduces overall burden of additional immunosuppression c) May well reduce risk of a renal flare – and renal disease definitely associated with shortened life expectancy
• Do you expect the technology to increase health-related quality of life more than current care?	Yes. For reasons outlined earlier (and especially if given s.c by patient) : a. Reduces frequency of visits b. When given IV, all bloods can be done at the same time c. Improved symptom control as more effective drug whilst having a lower pill burden and almost certainly lower doses of steroids d. Self management if giving sc allowing more autonomy e. Data from open label extension studies shows reduction in damage scores when compared to cohorts of patients not receiving belimumab

12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Lupus predominantly affects women of child-bearing age and most commonly women from BAME communities – both are disadvantaged groups. They often have child care issues, and challenging home situations. They also can have major mental health issues not only dealing with the burden of a chronic incurable disease but also changing body image especially with drugs such as steroids and bad lupus skin disease. Belimumab (esp if sc but even IV) can reduce the burden of all of these – fewer visits to hospital as could combine outpatients with infusion, and even fewer if on sc; no impact on body image and is steroid sparing so could improve that; increased autonomy and lower daily pill burden.
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.)	Some units not geared up for IV infusions on a regular basis but all rheumatological / renal centres of excellence will be so this should not be a limitation. All units will be used to training patients to give sc injections as common route of administration for rheumatological DMARDs. Patients will welcome lower daily pill burden and readily adapt to regular infusions / sc administration of drugs. Most centres giving belimumab managed to continue this throughout lockdown number 1 and indeed patients were willing to attend for their infusions whilst refusing to attend ordinary outpatient appointments – suggests a strong perceived benefit from infusions.
14. Will any rules (informal or	Current rules require improvement of SLEDAI-2K by 4 points at 6 months to allow continuation. This does

formal) be used to start or stop treatment with the technology? Do these include any additional testing?	not seem unreasonable to continue. Would mandate annual review of continued medication.
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?	Not expert in what is included in QALY calculation but would expect marked improvement in life participation due to improvement in fatigue scores (shown in trials) and SF36 scores (shown in trials) as well as general well being. Also reduction in use of steroids and reduction in flare rate should impact on this as both of those associated with worse damage and long term outcomes. Appears to reduce renal flare rate and renal flares associated with increased morbidity, NHS costs and mortality
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 improve the way that current need is met?	Yes – already showing benefits to patients: Reduced steroid use Reduced pill burden Improved quality of life SC administration will be beneficial to adherent patients requiring more autonomy. As described above – shared care pathways will improve overall management of patients with lupus – build

	links, facilitate rapid referral to centre when needed, encourage less use of steroids etc.
• Is the technology a 'step-change' in the management of the condition?	Yes – first newly licenced drug for the treatment of lupus in 50 years. Targeted immunosuppression Steroid sparing.
• Does the use of the technology address any particular unmet need of the patient population?	As stated earlier yes – steroid sparing
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?	Good safety profile now established and built upon with BLISS-LN. Low burden of side effects much more likely to lead to adherence (huge issue for lupus patients) vs drugs with much more adverse safety profiles e.g. steroids, azathioprine.
Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	NO. UK NHSE rules mean the patients have to have both a high dsDNA and low complement in order to start treatment – this excludes a lot of patients who otherwise might benefit. The trials required positive ANA or antidsDNA antibody only – not low complement. Analysis showed greatest benefit in those with high disease activity ie low complement and high dsDNA, but the NHS rules are currently too restrictive to

	as insist on high disease activity ie high dsDNA and low complement. Whilst those who are allowed to receive belimumab are those most likely to benefit, the rigid rules exclude many who might benefit.
• If not, how could the results be extrapolated to the UK setting?	Simply require either positive ANA / anti dsDNA or low complement as eligibility for the drug as well as a SLEDAI of at least 10
• What, in your view, are the most important outcomes, and were they measured in the trials?	The most important outcomes are always poorly captured in trials – namely patient reported quality of life. Fatigue was measured and improved and SF36 was measured. There are measures of skin and joint involvement but in lupus par excellence – symptoms can be vary variable, arthralgias can be crippling without leading to a SLEDAI score of note as no formal arthritis, skin is notoriously difficult to score – and what really needs measuring is patient participation in their lives – can they do what they want to do or does their illness limit them? This is not readily assessed, is certainly not captured in a SLEDAI and yet we routinely see patients who might not have had very high SLEDAI scores at the beginning, may only have modest drops over time but feel transformed by belimumab – walk better, feel better, reduce their steroids. I would suggest video recordings of patients at baseline – affect, walking, mobility – and at 6 months – would identify a greater proportion who have done well. We all forget how we felt just as we forget pain. Because a lot of symptoms not captured by SLEDAI need additional measures of response.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	Work towards disease remission or minimal disease activity: Reduction in steroid use would definitely be a good surrogate

	Reduction in flare rate requiring dose escalation of steroids or new additional drugs Reduction in renal involvement / flares – hugely important
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No and indeed concerns about suicidal ideation do not appear to be borne out.
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	Open label extension of average of 6.5 years f/up compared with the Toronto cohort with 16.5 years follow up - propensity matching showed reduction in damage scores in those on belimumab vs those who had not had belimumab, even in those high risk patients with pre existing organ damage suggesting a favourable effect on future damage development – Bruce IN et al Lupus 2016 25:699-709 https://pubmed.ncbi.nlm.nih.gov/26936891/ Just published – BLISS-LN data – two year RCT of belimumab in LN – Furie R et al NEJM 2020; 383:1117-1128 https://pubmed.ncbi.nlm.nih.gov/32937045/ "**Conclusions:** In this trial involving patients with active lupus nephritis, more patients who received belimumab plus standard therapy had a primary efficacy renal response than those who received standard therapy alone. (Funded by GlaxoSmithKline; BLISS-LN ClinicalTrials.gov number, NCT01639339.)." Not only likely to lead to extension of label to include treatment of lupus nephritis but extends safety data to addition to background MMF and steroids or Cyclophosphamide and steroids and no adverse signals accrued in this large trial over a long period.
20. Are you aware of any new evidence for the comparator treatment(s) since the	BLISS-LN trial just published – use in lupus nephritis; very favourable results with no safety signals despite addition to MMF and steroids. See comments in answer to question 20 and the guidelines referenced in answer to question 9. There are many more – below is just an initial literature search back for last year or so. NICE has more resource to do

publication of NICE technology appraisal guidance [TA397]?	these searches! • Comparative analysis of long-term organ damage in patients with **systemic lupus erythematosus** using **belimumab** versus standard therapy: a post hoc longitudinal study. Urowitz MB, Ohsfeldt RL, Wielage RC, Dever JJ, Zakerifar M, Asukai Y, Ramachandran S, Joshi AV.Lupus Sci Med. 2020 Oct;7(1):e000412. doi: 10.1136/lupus-2020-000412. • CALIBRATE: A Phase 2 Randomized Trial of Rituximab Plus Cyclophosphamide Followed by **Belimumab** for the Treatment of **Lupus** Nephritis. Atisha-Fregoso Y, Malkiel S, Harris KM, Byron M, Ding L, Kanaparthi S, Barry WT, Gao W, Ryker K, Tosta P, Askanase AD, Boackle SA, Chatham WW, Kamen DL, Karp DR, Kirou KA, Lim SS, Marder B, McMahon M, Parikh SV, Pendergraft WF 3rd, Podoll AS, Saxena A, Wofsy D, Diamond B, Smilek DE, Aranow C, Dall'Era M.Arthritis Rheumatol. 2020 Aug 4. doi: 10.1002/art.41466. Online ahead of print. • Safety and efficacy of intravenous **belimumab** in children with **systemic lupus erythematosus**: results from a randomised, placebo-controlled trial. Brunner HI, Abud-Mendoza C, Viola DO, Calvo Penades I, Levy D, Anton J, Calderon JE, Chasnyk VG, Ferrandiz MA, Keltsev V, Paz Gastanaga ME, Shishov M, Boteanu AL, Henrickson M, Bass D, Clark K, Hammer A, Ji BN, Nino A, Roth DA, Struemper H, Wang ML, Martini A, Lovell D, Ruperto N; Paediatric Rheumatology International Trials Organisation (PRINTO) and the Pediatric Rheumatology Collaborative Study Group (PRCSG).Ann Rheum Dis. 2020 Oct;79(10):1340-1348. doi: 10.1136/annrheumdis-2020-217101. Epub 2020 Jul 22. • Safety and efficacy of **belimumab** after B cell depletion therapy in **systemic LUPUS erythematosus** (BEAT-**LUPUS**) trial: statistical analysis plan.Muller P, Chowdhury K, Gordon C, Ehrenstein MR, Doré CJ.Trials. 2020 Jul 16;21(1):652. doi: 10.1186/s13063-020-04391-2. (study due to report in March 2021) • Long-term effects of combined B-cell immunomodulation with rituximab and belimumab in severe refractory SLE: 2 year results Kraaij T, Arends EJ, van Dam LS, Kamerling SWA, van Daele PLA, Bredewold OW, Ray A, Bakker JA, Scherer HU, Huizinga TJW, Rabelink TJ, van Kooten C, Teng YKO.Nephrol Dial Transplant. 2020 Jun 27:gfaa117. doi: 10.1093/ndt/gfaa117. Online ahead of print.PMID: 32591783 • Patient Satisfaction And Disease Control In Patients With **Systemic Lupus Erythematosus** Is Not Affected By Switching From Intravenous **Belimumab** To Subcutaneous Injections. Mucke J, Brinks

	R, Fischer-Betz R, Richter JG, Sander O, Schneider M, Chehab G.Patient Prefer Adherence. 2019 Nov 4;13:1889-1894. doi: 10.2147/PPA.S227208. eCollection 2019.PMID: 31806937 • The safety and efficacy of biologic agents in treatment of **systemic lupus erythematosus**: A network meta-analysis. Tao MJ, Cheng P, Jin LR, Zhou J, Shi W, Peng H, Xu L, Li Z, Yuan H.Pak J Med Sci. 2019 Nov-Dec;35(6):1680-1686. doi: 10.12669/pjms.35.6.771.PMID: 31777515 • Effect of **belimumab** treatment on antiphospholipid antibody levels: post-hoc analysis based on two randomised placebo-controlled trials in **systemic lupus erythematosus**. Chatzidionysiou K, Samoli E, Sfikakis PP, Tektonidou MG.Ann Rheum Dis. 2020 Feb;79(2):304-307. doi: 10.1136/annrheumdis-2019-216367. Epub 2019 Nov 11.PMID: 31712248 • Changes in quality of life in relation to disease activity in **systemic lupus erythematosus**: post-hoc analysis of the BLISS-52 Trial.Jolly M, Annapureddy N, Arnaud L, Devilliers H.Lupus. 2019 Dec;28(14):1628-1639. doi: 10.1177/0961203319886065. Epub 2019 Nov 1.PMID: 31674267 • Long-term safety and limited organ damage in patients with **systemic lupus erythematosus** treated with **belimumab**: a Phase III study extension.van Vollenhoven RF, Navarra SV, Levy RA, Thomas M, Heath A, Lustine T, Adamkovic A, Fettiplace J, Wang ML, Ji B, Roth D.Rheumatology (Oxford). 2020 Feb 1;59(2):281-291. doi: 10.1093/rheumatology/kez279.PMID: 31302695 • Achieving **lupus** low-disease activity and remission states under **belimumab** in refractory **systemic lupus erythematosus**: time and organ involvement matter.Sbeih N, Mathian A, Pineton de Chambrun M, Lhote R, Zahr N, Pha M, Cohen-Aubart F, Haroche J, Hié M, Jouffroy S, Benameur N, Devilliers H, Amoura Z.Ann Rheum Dis. 2020 Nov;79(11):e148. doi: 10.1136/annrheumdis-2019-215732. Epub 2019 Jul 8.PMID: 31285257 • Organ system improvements in Japanese patients with **systemic lupus erythematosus** treated with **belimumab**: A subgroup analysis from a phase 3 randomized placebo-controlled trial.Tanaka Y, Bass D, Chu M, Egginton S, Ji B, Roth D.Mod Rheumatol. 2020 Mar;30(2):313-320. doi: 10.1080/14397595.2019.1630897. Epub 2019 Jul 4.PMID: 31199180 • Predictors of low disease activity and clinical remission following **belimumab** treatment in **systemic lupus erythematosus**.Parodis I, Johansson P, Gomez A, Soukka S, Emamikia S, Chatzidionysiou K.Rheumatology (Oxford). 2019 Dec 1;58(12):2170-2176. doi: 10.1093/rheumatology/kez191.PMID: 31157891 • Phase III, multicentre, randomised, double-blind, placebo-controlled, 104-week study of subcutaneous **belimumab** administered in combination with rituximab in adults

	with **systemic lupus erythematosus** (SLE): BLISS-BELIEVE study protocol.Teng YKO, Bruce IN, Diamond B, Furie RA, van Vollenhoven RF, Gordon D, Groark J, Henderson RB, Oldham M, Tak PP.BMJ Open. 2019 Mar 20;9(3):e025687. doi: 10.1136/bmjopen-2018-025687.PMID: 30898822 (ongoing study) • Safety and Efficacy of **Belimumab** Plus Standard Therapy for Up to Thirteen Years in Patients With **Systemic Lupus Erythematosus**.Wallace DJ, Ginzler EM, Merrill JT, Furie RA, Stohl W, Chatham WW, Weinstein A, McKay JD, McCune WJ, Petri M, Fettiplace J, Roth DA, Ji B, Heath A.Arthritis Rheumatol. 2019 Jul;71(7):1125-1134. doi: 10.1002/art.40861. Epub 2019 Jun 5.PMID: 30771238
21. How do data on real-world experience compare with the trial data?	Well tolerated in long term – numerous publications on this in recent years. Need more data on pregnancy as ideally would continue through pregnancy – no suggestion of teratogenicity
Equality	
22a. Are there any potential equality issues that should be taken into account when considering this treatment?	See my answer to question 12. This is a disease predominantly affecting women and those from BAME backgrounds. They are severely disadvantaged by this chronic disease and need effective, safe long term treatments that do not alter body image or increase damage in themselves. Belimumab offers this.
22b. Consider whether these issues are different from issues	

with current care and why.	

Key messages

23. In up to 5 bullet points, please summarise the key messages of your submission.

- It is the first licenced drug and is targeted therapy for SLE in 50 years; grade A evidence for efficacy and now for lupus nephritis as well
- The current NHSE criteria for access to belimumab are too rigid and should be extended to patients who are simply seropositive and have a SLEDAI of 10
- Maintain initiation of treatment has to be by centres of excellence and experience but develop shared care pathways to improve access, and raise the standard of care for patients with lupus more generally
- Growing comparator and real world data on safety, tolerability and benefit of belimumab – need to widen access
- Importance of approving and increasing use of sc vs IV preparation after initiation of treatment.

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

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (review of TA397) [ID1591]

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
Sabine Grimm, Health Economist, Maastricht UMC, the Netherlands
Ben Wijnen, Health Economist, Maastricht UMC
Annette Chalker, Systematic Reviewer, KSR Ltd
Sean Harrison, Systematic Reviewer, KSR Ltd
Nigel Armstrong, Health Economist, KSR Ltd
Thomas Otten, Health Economist, Maastricht UMC
Mohammed Islam, Health Economist, KSR Ltd
Gill Worthy, Statistician, KSR Ltd
Lisa Stirk, 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, UK
YO19 6FD

Date completed 19/01/2021

Source of funding: This report was commissioned by the NIHR Evidence Synthesis Programme as project number STA 13/15/80.

Declared competing interests of the authors

None.

Acknowledgements

We gratefully acknowledge the expert clinical advice input from Professor C. Gordon (Emeritus Professor of Rheumatology, University of Birmingham).

Commercial in confidence (CiC) data are highlighted in blue throughout the report.

Academic in confidence (AiC) data are highlighted in yellow 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, Grimm S, Wijnen B, Chalker A, Harrison S, Armstrong N, Otten T, Islam M, Worthy G, Stirk L, Joore MA, Kleijnen J. Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (review of TA397) [ID1591]: 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. Sabine Grimm acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Ben Wijnen, 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. Annette Chalker and Sean Harrison acted as systematic reviewers, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Gill Worthy and Sean Harrison acted as statisticians, critiqued the analyses in the company's submission and contributed to the writing of the report. Lisa Stirk 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

ACEi	Angiotensin-converting enzyme inhibitor
ACR	American College of Rheumatology
AE	Adverse events
AESI	Adverse events of special interest
AIC	Akaike information criterion
ANCOVA	Analysis of covariance
ARB	Angiotensin II receptor blocker
AWR	Access with Research
BI	Budget impact
BIC	Bayesian information criterion
BILAG	British Isles Lupus Assessment Group
BLyS	B lymphocyte stimulator
CD	Cluster of differentiation
CE	Cost effectiveness
CEA	Cost effectiveness analysis
CEAC	Cost effectiveness acceptability curve
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CNS	Central nervous system
COVID	Coronavirus disease
CRD	Centre for Reviews and Dissemination
CS	Company's submission
CSR	Clinical study report
C-SSRS	Colombia Suicide Severity Rating Scale
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded DNA
DSU	Decision Support Unit
EMA	European Medicines Agency
EPAR	European Public Assessment Report
EQ-5D	European Quality of Life-5 Dimensions
ERG	Evidence Review Group
EUR	Erasmus University Rotterdam
FAD	Final appraisal document
FACIT	Functional Assessment of Chronic Illness Therapy
FDA	Food and Drug Administration
GSK	GlaxoSmithKline
HDA	High disease activity
HR	Hazard ratio
HRQoL	Health-related quality of life
HTA	Health technology assessment
IC	Indirect comparison
ICER	Incremental cost effectiveness ratio
IFN	Interferon
ITC	Indirect treatment comparison
ITT	Intention-to-treat
IV	Intravenous
IVIG	Intravenous immunoglobulin
KSR	Kleijnen Systematic Reviews
LTE	Long-term extension
LYs	Life years
LYG	Life years gained
MAIC	Match-adjusted indirect comparison
MCID	Minimal clinically important difference

MCS	Mental component score
MeSH	Medical subject headings
MHRA	Medicines and Healthcare Products Regulatory Agency
MTA	Multiple technology appraisal
MTC	Mixed treatment comparison
NA	Not applicable
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NMA	Network meta-analysis
NMSC	Non-melanoma skin cancer
NR	Not reported
NSAID	Non-steroidal anti-inflammatory drug
OR	Odds ratio
OS	Overall survival
PAS	Patient access scheme
PCS	Physical component score
PD	Pharmacodynamics
PFS	Progression-free survival
PGA	Physician's global assessment
PK	Pharmacokinetics
PRESS	Peer Review of Electronic Search Strategies
PRISMA	Preferred reporting items for systematic reviews and meta-analyses
PRO	Patient reported outcome
PS	Propensity score
PSA	Probabilistic sensitivity analysis
PSM	Propensity score-matching
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
QALY	Quality-adjusted life year
QoL	Quality of life
RCT	Randomised controlled trial
RR	Risk ratio
RWE	Real-world evidence
SAE	Serious adverse events
SC	Subcutaneous
ScHARR	School of Health and Related Research
SD	Standard deviation
SDI	SLICC/ACR Damage Index
SE	Standard error
SELENA	Safety of Estrogen in Lupus Erythematosus National Assessment
SF-36	Short Form-36
SFI	SLE Flare Index
SLE	Systemic lupus erythematosus
SLEDAI	Systemic lupus erythematosus disease activity index
SLICC	Systemic lupus international collaborating clinics
SLR	Systematic literature review
SMC	Scottish Medicines Consortium
SmPC	Summary of product characteristics
SRI-4	SLE responder index-4
ST	Standard therapy
STA	Single technology appraisal
TA	Technology assessment
TEAE	Treatment emergent adverse events
TLC	Toronto lupus cohort

TNF	Tumour necrosis factor
TTO	Time trade-off
UK	United Kingdom
UMC	University Medical Centre
USA	United States of America
WHO	World Health Organization
WTP	Willingness-to-pay

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 discussed the decision problem, Section 1.3 issues related to the clinical effectiveness, and Section 1.4 issues related to the cost effectiveness. Background information on the condition, technology and evidence and information on non-key issues are in the main ERG report, see Sections 2 (background), 3 (decision problem), 4 (clinical effectiveness) and 5 (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

ID1457	Summary of issue	Report sections
1.	Evidence is missing for specific populations, such as children and patients with severe active central nervous system (CNS) lupus.	Section 3.1
2.	Some comparators listed in the NICE scope were not included	Section 3.3
3.	Short follow-up in the main comparative trials (BLISS-SC, BLISS-52 and BLISS-76)	Section 4.2.1
4.	Using the propensity score-matching (PSM) analysis in calibrating the cost-effectiveness model can severely bias the results in favour of belimumab	Section 4.3 and 4.4
5.	Data from the BILAG Biologic Register (BILAG BR) are not suitable for a comparison of belimumab with rituximab	Section 4.3 and 4.4
6.	Rituximab + standard therapy was not included as comparator in the model	Section 5.2.4
7.	IV and SC formulations are not compared with each other, as two separate model files are provided.	Section 5.2.4
8.	Use of calibration factor is likely biasing results	Section 5.2.6
9.	Implementation of 24-week response and treatment continuation in the model is inconsistent	Section 5.2.6
10.	Error in calculation of belimumab non-responder disease activity at 52 weeks	Section 5.2.6
11.	Violation in utility estimation	Section 5.2.8
12.	Uncertainty about organ damage utility multipliers	Section 5.2.8
13.	Sampling of organ damage and death occurs after allocation to treatment	Section 6.4

The key difference between the company's preferred assumptions and the ERG's preferred assumptions is the use of the calibration factor for adjusting long-term effects of belimumab on organ damage.

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 length of survival
- improving health-related quality of life associated with systemic lupus erythematosus (SLE) disease activity and through the delay or prevention of organ damage

Overall, the technology is modelled to affect costs by:

- its higher unit price than current treatments
- a decrease in disease activity related costs and costs related to organ damage

The modelling assumption that has the greatest effect on the ICER is:

- the use of the calibration factor for adjusting long-term effects of belimumab on organ damage.

1.3 The decision problem: summary of the ERG's key issues

The decision problem addressed in the company submission (CS) is broadly in line with the final scope issued by NICE. However, there is a lack of evidence for specific populations, such as children and patients with severe active CNS lupus (Table 1.2) and some comparators listed in the NICE scope were not included (Table 1.3).

Table 1.2: Key issue 1: Evidence is missing for specific populations

Report section	Section 3.1
Description of issue and why the ERG has identified it as important	Evidence is missing for specific populations, such as children and patients with severe active CNS lupus. - Although belimumab is indicated as add-on therapy in patients aged 5 years and older with active, autoantibody-positive systemic lupus erythematosus (SLE) with a high degree of disease activity despite standard therapy, patients with severe active CNS lupus were excluded from the BLISS trials and no evidence to support the use of belimumab is available in this population. In addition, no searches were performed in people over the age of 5 as the CS focuses on an adult population with SLE as does the economic modelling.
What alternative approach has the ERG suggested?	The ERG is unable to suggest an alternative approach, because no evidence is presented for these populations.
What is the expected effect on the cost effectiveness estimates?	The change to the ICER is unclear.
What additional evidence or analyses might help to resolve this key issue?	No evidence in these populations is currently available according to the company.

Table 1.3: Key issue 2: Some comparators listed in the NICE scope were not included

Report section	Section 3.3
Description of issue and why the ERG has identified it as important	The company only included one comparator: Standard therapy alone. Rituximab plus standard therapy or cyclophosphamide plus standard therapy were not considered relevant comparators by the company.
What alternative approach has the ERG suggested?	The ERG is unable to suggest an alternative approach, because no reliable evidence is presented for these comparators.
What is the expected effect on the cost effectiveness estimates?	The change to the ICER is unclear.
What additional evidence or analyses might help to resolve this key issue?	Regarding cyclophosphamide, the ERG believes the company has a point because it is mainly used for different populations than belimumab and the adverse event profile of cyclophosphamide means that it is avoided if possible. The comparison with rituximab will be difficult because the evidence for rituximab is weaker as the phase 3 trials were negative due to very stringent end-points (and different to those used for belimumab) and is mostly from registries. BILAG BR data cannot be used to compare them easily due to the different criteria for the use of rituximab and belimumab (See also Section 3.3 of this report).

1.4 The clinical effectiveness evidence: summary of the ERG's key issues

The clinical effectiveness evidence presented in the CS is mostly based on the same studies as in the original submission (TA397). The issues with these studies have been critiqued in the ERG report for TA397 and will not be repeated here (see also Section 4.6 of this report). The current appraisal is different from the original appraisal (TA397) in three ways:

1. The definition of 'high disease activity' (i.e. HDA-1 versus HDA-2, see Section 3.1).
2. Age: in TA397, belimumab was approved for adults only. This appraisal includes people aged five years or more.
3. Formulation: The original appraisal included an intravenous (IV) formulation only. The current appraisal also includes a new subcutaneous (SC) formulation in the form of a pre-filled pen.

The ERG identified three major concerns with the evidence presented on the clinical effectiveness in the current submission, namely short follow-up in the main comparative trials (BLISS-SC, BLISS-52 and BLISS-76) (see Table 1.4), using the propensity score-matching (PSM) analysis in calibrating the cost effectiveness model can severely bias the results in favour of belimumab (see Table 1.5) and BILAG BR data are not suitable for a comparison of belimumab with rituximab (see Table 1.6).

Table 1.4: Key issue 3: Lack of reliable long-term comparative follow-up data.

Report section	Section 4.2.1
Description of issue and why the ERG has identified it as important	All three trials used the same primary efficacy endpoint, which was the SRI-4 response rate at Week 52. BLISS-SC and BLISS-52 had a maximum follow-up of 52 weeks, while BLISS-76 had a maximum follow-up of 76 weeks. All three trials did have long-term extension (LTE) phases. However, all patients received belimumab during the LTE.

	Therefore, these extension studies do not provide comparative evidence.
What alternative approach has the ERG suggested?	There is no reliable comparative evidence available beyond 76 weeks.
What is the expected effect on the cost effectiveness estimates?	The expected change to the ICER is unclear.
What additional evidence or analyses might help to resolve this key issue?	Long term comparative studies of Belimumab versus the main comparators would help to resolve this key issue.

Table 1.5: Key issue 4: Using the PSM analysis in calibrating the cost effectiveness model can severely bias the results in favour of belimumab.

Report section	Sections 4.3 and 4.4
Description of issue and why the ERG has identified it as important	The CS uses a propensity score matching (PSM) analysis to calibrate the Johns Hopkins (JH) model for organ damage over time for patients on belimumab. The main issue with this analysis is that the calibration factor is derived after 5 years of belimumab treatment, using 99 patients from the BLISS US-LTE matched with 99 patients from the Toronto Lupus Cohort (TLC) as the data source. This means that the calibration factor is biased towards belimumab preventing organ damage, as most patients withdrew from the BLISS US LTE before 5 years, regardless of discontinuation of belimumab over that period, and those remaining on treatment at 5 years are likely to either have responded unusually well to belimumab or had slower progressing SLE than those that withdrew. It is the application of the calibration factor derived at 5 years to the whole 5 years that is the key issue. Additionally, the patients are not necessarily representative of either BLISS US-LTE patients or TLC patients, but representative only of patients matching between these cohorts: this is unlikely to be representative of patients in the UK. Most importantly, the calibration factor derived from the PSM analysis of 0.491 effectively doubles the effectiveness of belimumab for preventing organ damage, compared with the JH model.
What alternative approach has the ERG suggested?	In the absence of better evidence, the ERG has removed the calibration factor.
What is the expected effect on the cost effectiveness estimates?	The calibration factor derived from fitting the cost effectiveness model to the PSM analysis likely biases the model to make belimumab seem more cost effective than it is. However, it is unknown how large this increase might be, as analysing the results weighting for a UK cohort will also change the cost effectiveness, possibly in either direction.
What additional evidence or analyses might help to resolve this key issue?	Ideally, data from an RCT would inform the effectiveness of belimumab versus standard therapy for at least five years follow-up, rather than relying on different observational data to inform the belimumab and standard therapy arms to inform the CEA. One way of producing a less biased estimate of long-term SDI at least for those on belimumab treatment would be as follows: instead of matching BLISS LTE patients with TLC, create

	propensity scores using UK SLE cohort data (e.g. BILAG BR data), which could be tailored so only HDA-1 and HDA-2 subgroups are included, then weight the BLISS LTE data to make it generalisable to a UK cohort using the propensity scores. This does not require follow-up data for the UK cohort, just enough information to weight the BLISS LTE data so it is generalisable to a UK cohort. There remains an issue of patients who drop out of the belimumab arm, as they are more likely than those on standard therapy to have more quickly progressing SLE, and therefore are unlikely to have SDI increases over time comparable to those on standard therapy. However, to our knowledge there is no dataset that will allow good estimation of outcomes for these patients. One possible sensitivity analysis would be to give those who drop out of belimumab the expected increase in SDI as the most quickly progressing percentile, e.g. 50%, of standard therapy patients, to account for the greater likelihood of progression for these patients. It should be noted this does not remove the potential for confounding between the belimumab and standard therapy arms (although this was not removed by using a calibration factor solely for the belimumab arm, and the PSM analysis itself is probably biased towards belimumab being effective so could not resolve this either). It also does not completely resolve the bias in the outcome data for patients who drop out of the belimumab arm, though will likely be less biased than at present.

Table 1.6: Key issue 5: BILAG BR data are not suitable for a comparison of belimumab with rituximab

Report section	Section 4.4
Description of issue and why the ERG has identified it as important	Data from the BILAG Biologic Register (BILAG BR) cannot be used to make a reliable comparison of the effectiveness of belimumab versus rituximab due to the different criteria for the use of rituximab and belimumab (See BSR guidelines for SLE).
What alternative approach has the ERG suggested?	There is no reliable evidence for a comparison of belimumab versus rituximab.
What is the expected effect on the cost effectiveness estimates?	The expected change to the ICER is unclear.
What additional evidence or analyses might help to resolve this key issue?	Given the different criteria for the use of rituximab and belimumab it is unlikely a head-to-head trial of belimumab versus rituximab is feasible.

1.5 The cost effectiveness evidence: summary of the ERG's key issues

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 key issues in the cost effectiveness evidence are discussed in Tables 1.7 to 1.16.

Table 1.7: Key issue 6: Rituximab excluded as comparator

Report section	Section 5.2.4
Description of issue and why the ERG has identified it as important	Rituximab was excluded as a comparator but may be a relevant comparator and was mentioned in the scope.
What alternative approach has the ERG suggested?	Include rituximab as comparator
What is the expected effect on the cost effectiveness estimates?	Unknown
What additional evidence or analyses might help to resolve this key issue?	The ERG understands that this comparison will be difficult.

Table 1.8: Key issue 7: IV and SC formulations are not compared

Report section	Section 4.2.6 and 5.2.6
Description of issue and why the ERG has identified it as important	IV and SC formulations are not compared with each other. Two separate model files are provided.
What alternative approach has the ERG suggested?	Include both formulations in one model file and enable comparison of IV and SC formulations (full incremental analysis).
What is the expected effect on the cost effectiveness estimates?	Cost effectiveness will not be affected but this would enable a comparison of cost effectiveness of IV and SC.
What additional evidence or analyses might help to resolve this key issue?	

Table 1.9: Key issue 8: Use of the calibration factor

Report section	Section 4.2.6 and 5.2.6
Description of issue and why the ERG has identified it as important	The calibration factor lacks validity and should not be used. It does not resolve uncertainty about long-term organ damage in patients treated with belimumab versus standard therapy due to issues with methodology, the BLISS-LTE evidence and the PSM study. The use of this calibration factor is a major driver of cost-effectiveness outcomes. For example, the company provided a scenario without using the calibration exercise, resulting in an ICER of £47,872 per QALY gained for the IV model with the HDA-2 population and an ICER of £56,277 per QALY gained for the SC model with the HDA-2 population.
What alternative approach has the ERG suggested?	The ERG considers it most appropriate to remove the calibration factor in its base-case.
What is the expected effect on the cost effectiveness estimates?	The use of a calibration factor likely results in an overestimation of treatment effect of belimumab and hence underestimates the ICER.
What additional evidence or analyses might help to resolve this key issue?	Empirical evidence is lacking to validate the calibration factor. The issues with the long-term belimumab data are unlikely to be resolved.

Table 1.10: Key issue 9: Implementation of 24-week response and treatment continuation in the model

Report section	Section 5.2.6
Description of issue and why the ERG has identified it as important	The probability of being a responder is based on the baseline SS score, which is linked to the responder criteria applied to patients in the BLISS trials (i.e. only for patients with a reduction of ≥ 4 points in SS score). Hence, it is estimated at baseline in the model and not directly linked to the actual improvement in SS score in the model. In turn, actual SS scores are estimated based on a regression model where response is an independent variable, given that a 24-week time point does not exist in the model. As a result, a large proportion of patients is classed as non-responder but experiences >4 points reduction in SS at 52 weeks.
What alternative approach has the ERG suggested?	An adjustment of the model to align reduction in SS at 52-weeks with the defined criteria for being responder/non-responder (i.e. >4 points reduction in SS).
What is the expected effect on the cost effectiveness estimates?	This could lead to under-estimation of belimumab costs in the model compared to clinical practice as patients with no response do not continue belimumab.
What additional evidence or analyses might help to resolve this key issue?	Not applicable

Table 1.11: Key issue 10: Error in calculation of belimumab non-responder disease activity at 52 weeks

Report section	Section 5.2.6
Description of issue and why the ERG has identified it as important	Due to a modelling error, belimumab non-responders have the same reduction in disease activity at 52 weeks as patients in the standard therapy arm.
What alternative approach has the ERG suggested?	The ERG proposes a correction of the model to align reduction in SS at 52-weeks for non-responders with the regression function mentioned in the CS.
What is the expected effect on the cost effectiveness estimates?	This difference leads to an overestimation of treatment effectiveness in the belimumab arm at 52 weeks (in particular the belimumab non-responders).
What additional evidence or analyses might help to resolve this key issue?	Correction of the company's modelling

Table 1.12 Key issue 11: Violation in utility estimation

Report section	Section 5.2.8
Description of issue and why the ERG has identified it as important	The SLE-related utility estimate excludes key organ damage covariates without adjusting the remaining coefficients.
What alternative approach has the ERG suggested?	The ERG suggests re-estimating the utility model coefficients after excluding the organ damage covariates.
What is the expected effect on the cost effectiveness estimates?	Probably minor, direction unknown

Report section	Section 5.2.8
What additional evidence or analyses might help to resolve this key issue?	The ERG suggests re-estimating the utility model coefficients after excluding the organ damage covariates and providing an explanation regarding the discrepancy between model coefficients presented in the CS and the ones used in the model.

Table 1.13 Key issue 12: Uncertainty about organ damage utility multipliers

Report section	Section 5.2.8
Description of issue and why the ERG has identified it as important	Uncertainty about organ damage utility multipliers – this may over-estimate the impact of organ damage on HRQoL as the utility estimation function may capture this to a certain extent.
What alternative approach has the ERG suggested?	To explore this uncertainty, use a scenario in which the organ damage utility multipliers are disabled
What is the expected effect on the cost effectiveness estimates?	When the company's calibration factor is used, the ICER increases upon removal of organ damage utility multipliers. When the calibration factor is removed, the ICER decreases upon removal of organ damage utility multipliers.
What additional evidence or analyses might help to resolve this key issue?	The company could investigate whether the weighting of organ damage items corresponds with the latest evidence and consult expert opinion on the magnitude of the organ damage utility multipliers.

Table 1.14 Key issue 13: Sampling of organ damage and death occurs after allocation to treatment

Report section	Section 6.3
Description of issue and why the ERG has identified it as important	In the VBA, first, a simulated patient is allocated to a treatment and organ damage and death are only sampled within the treatment arm. This leads to the same simulated patient (same age, gender, SS score) experiencing differential organ damage and times of death only dependent on allocation to treatment arm but not caused by this allocation (so just because of sampling). This induces noise in the model and makes validation difficult.
What alternative approach has the ERG suggested?	A structural model adjustment in which organ damage items involved and death are sampled from before treatment allocation
What is the expected effect on the cost effectiveness estimates?	Unknown. Maybe no effect if all this did was induce noise.
What additional evidence or analyses might help to resolve this key issue?	A structural model adjustment in which organ damage items involved and death are sampled from before treatment allocation

1.6 Other key issues: summary of the ERG's view

No other key issues were identified by the ERG.

1.7 Summary of the ERG's view

The following tables summarise the ERG's changes to the company's base-case to arrive at an ERG base-case (Tables 1.15-1.16). In addition, Tables 1.17-1.18 present the ERG scenarios.

Table 1.15: Deterministic ERG base-case for the IV formulation (HDA-2 subgroup, PAS price)

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (/QALY)
CS base-case					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£30.000
Standard therapy	£160.470	9.809			
Fixing errors 1: 1st year: SS reduction for belimumab non-responders					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£31,695
Standard therapy	£160,470	9.809			
Matter of judgement 2: Calibration factor removed conditional on FE1 = ERG base-case					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£52,891
Standard therapy	£160,470	9.809			

Table 1.16: Deterministic ERG base-case for the SC formulation (HDA-2 subgroup, PAS price)

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (/QALY)
CS base-case					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£30,566
Standard therapy	£151,999	10.056			
Fixing errors 1: 1st year: SS reduction for belimumab non-responders					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£32,617
Standard therapy	£151,999	10.056			
Matter of judgement 2: Calibration factor removed conditional on FE1 = ERG base-case					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£61,057
Standard therapy	£151,999	10.056			

Table 1.17: ERG scenarios for the IV formulation (HDA-2 subgroup, PAS price)

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (/QALY)
ERG base-case					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£52,891
Standard therapy	£160,470	9.809			
Scenario 1: Use unadjusted JH model					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£63,951
Standard therapy	£161,467	10.798			
Scenario 2: Use calibration factor					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£31,695
Standard therapy	£160,470	9.809			
Scenario 3: Use calibration factors on both arms					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£24,847
Standard therapy	£167,261	9.669			
Scenario 4: Remove impact of organ damage					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£48,347
Standard therapy	£160,470	11.941			
Scenario 5: Patient weight based on trial					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£50,451
Standard therapy	£160,470	9.809			
Scenario 6: HDA-1 subgroup					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£48,849
Standard therapy	£166,658	10.216			
Scenario 7: HDA-1 subgroup conditional on company's base-case with FE1					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£28,265
Standard therapy	£166,658	10.216			

Table 1.18: ERG scenarios for the SC formulation (HDA-2 subgroup, PAS price)

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (/QALY)
ERG base-case					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£61,057
Standard therapy	£151,999	10.056			
Scenario 1: Use unadjusted JH model					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£68,909
Standard therapy	£151,873	11.036			
Scenario 2: Use calibration factors					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£32,617
Standard therapy	£151,999	10.056			
Scenario 3: Use calibration factors on both arms					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£25,418
Standard therapy	£158,791	9.916			
Scenario 4: Remove impact of organ damage					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£56,901
Standard therapy	£151,999	12.082			
Scenario 6: HDA-1 subgroup					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£60,241
Standard therapy	£156,692	10.476			
Scenario 7: HDA-1 subgroup conditional on company's base-case with FE1					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£31,706
Standard therapy	£156,692	10.476			

2. BACKGROUND

2.1 Introduction

In this report, the ERG provides a review of the evidence submitted by GSK in support of belimumab, trade name Benlysta®, as an add-on to standard therapy for patients with autoantibody-positive systemic lupus erythematosus (SLE), with a high degree of disease activity despite standard therapy. The Company Submission (CS) extends a previously approved intravenous (IV) formulation of belimumab (TA397 in 2016) to a) a younger population (five years or more versus 18 years or more previously), b) a new patient population (patients with a SELENA-SLEDAI score ≥10 AND at least one of the following serological features: low complement OR positive anti-dsDNA [HDA-2] versus patients with a SELENA SLEDAI score ≥10 AND low complement AND positive anti-dsDNA [HDA-1] previously), and c) a new subcutaneous (SC) formulation of belimumab. In this section, the ERG summarises and critiques the company's description of the underlying health problem and the company's overview of the current service provision. The information for this critique is taken from Document B of the CS.[1]

2.2 Critique of company's description of the underlying health problem

The underlying health problem of this appraisal is systemic lupus erythematosus (SLE), an autoimmune, multi-system disease with varying manifestations characterised by an unpredictable clinical course, autoantibody production, abnormal B lymphocyte function and chronic inflammation leading to high morbidity and mortality rate.[2] The exact cause of SLE is unknown but there is evidence for multiple genetic and environmental factors contributing to development of the disease. Gender, race, socioeconomic status, family history and environmental exposures appear to be important disease determinants.[3]

The CS highlighted that SLE is challenging to diagnose due to the complexity and heterogeneity of the condition, with no definitive tests for diagnosis, considerable variation in presentation, extent and severity of clinical signs and symptoms that can occur in any organ system. The CS described a UK survey that demonstrated a mean time to diagnosis from the first symptom of SLE of 6.4 years (SD=9.5), with 47% of patients diagnosed with a different condition prior to SLE.[4]

SLE affected nearly 0.1% of the population of the UK in 2012, most typically women between the ages of 20 and 60 years, with a female to male ratio of 9:1, and a peak incidence between 40 and 49 years for women and 60 and 69 years for men.[5] SLE is also more common in the UK in people of African-Caribbean and South Asian descent.[5-7] The CS stated that although the mortality rate has improved over time, the mortality rate of SLE was still high at 10% over 20 years, with a mean age of death of 53.7 years.[8] The CS noted that around one in three SLE patients in the UK develop lupus nephritis, which can lead to end-stage renal failure,[2] and that a patient diagnosed [in the US] with lupus at 20 years of age has a one in six chance of dying by 35 years of age, most often from the complications of lupus or infection.[9]

In the original American College of Rheumatology criteria, four of 11 clinical and laboratory criteria must be met to diagnose SLE,[10] though other classification criteria exist, such as the 2012 Systemic Lupus International Collaborating Clinics (SLICC) criteria[11] and the ACR/EULAR joint criteria.[12] SLE is a relapsing and remitting condition, characterised by periods of exacerbation (flares) and relative quiescence, and can affect multiple organs, giving rise to a wide range of clinical manifestations and serological features.[13, 14] Typical SLE presentations include fatigue and symptoms involving the skin, such as facial scarring and hair loss, and joints, such as pain and impaired physical function.[15] The CS notes that a large proportion of SLE patients are unable to remain in paid employment,[16, 17] and that

SLE *"inevitably forces a patient to relinquish control of their lives, nullifying their ability to maintain normalcy or predictability."*[1] The CS also notes that SLE impacts all aspects of health-related quality of life (HRQoL), including "*physical and mental health, vitality, pain, social and emotional functioning and activities of daily living"*.[1]

The majority of patients with long-term active SLE develop permanent organ damage, leading to organ dysfunction, which progresses steadily over time.[18-23] In addition, long-term high-dose glucocorticoid treatment (the CS reported that 72% of SLE patients in Europe were receiving ongoing treatment with corticosteroids 10 years after diagnosis[24]) can also contribute to myopathy, osteoporosis, hypertension, diabetes, atherosclerotic vascular disease, infections, and death.[22] As the CS notes, "*patients with SLE also appear to be at greater risk of developing other diseases and therapy related morbidity, including infections, especially of the respiratory and urinary systems,[25, 26] atherosclerosis, vascular disease and coronary artery disease,[27-29] and haematological and solid tumours,[30-32] as well as increased risk for mortality.[33, 34] SLE is also associated with significant maternal and foetal morbidity, including spontaneous abortion, pre-eclampsia, intrauterine growth restriction, foetal death and pre-term delivery.[35,,1]*

ERG comment: Overall, the overview of SLE was reasonably well balanced and referenced, and most of the evidence was accessible and appears to be applicable to the UK population, though the references were somewhat out of date with many quoted publications pre 2010, and there were some specific issues that are detailed below. One large issue is that children are not mentioned in the description of SLE: as only adults were considered in TA397, we would have expected details about the presentation, incidence and prevalence of SLE in children aged five to 18 years.

- The CS stated that SLE affected "*1 in 1000 of the [UK] population"*: the cited reference[5] used the UK Clinical Practice Research Datalink (CPRD) to estimate the SLE prevalence with three different definitions, of which one is roughly accurate for 2012, although it should be noted the prevalence increased between 1999 and 2012 (the study period), so the prevalence in 2020 is likely to be higher. However, the female to male ratio was 6.75:1 in 2012, not 9:1 as stated in the CS (this was the female to male ratio for prevalent cases during the *entire* study period, not for 2012 as for the overall prevalence). The ratio of females to males increased during the study period, as the prevalence of SLE increased faster in females than in males.
- The CS only mentioned that SLE was more common in "*people of African-Caribbean and South Asian descent"*: a cited reference[5] gives the prevalence of SLE in 2012 in the UK as five in 1,000 people of Black Caribbean descent, and between 1.8-1.9 per 1000 people of Chinese/other Asian descent.
- The mean age of death in the CS was not from their cited reference[36] but from a different study based in Birmingham rather than the UK as a whole:[8] the median age of death in the cited study (and relevant to the UK as a whole) was 72.8 years (IQR = 61.4 to 80.3 years). Our clinical expert disputes this and suggests very few patients in UK clinical practice are that old.[37]
- It is unclear from where in the cited reference the following statement in the CS was taken: *"mortality remains high with a 10% mortality over 20 years"*.[1] Only the standardised mortality rate over 10 years is given in the cited reference.
- The CS stated that *"A patient in whom lupus is diagnosed at 20 years of age still has a 1 in 6 chance of dying by 35 years of age"*,[1] but whilst the cited reference[9] includes this statement, the reference cites the statement from a textbook published in 2007, which KSR could not check for reliability.[38] However, this statement is unlikely to be true for the UK in 2020 given both the differences in healthcare between Canada and the UK and the advances made in treating SLE that have reduced the mortality rate over time.

- The CS noted that *"In order for a diagnosis SLE to be established, four of 11 clinical and laboratory criteria must be met"*.[1] This statement is incorrect, these are classification criteria not diagnostic criteria. In addition, it should be noted the citation adds that meeting four out of 11 criteria means "lupus can be classified with 95% specificity and 85% sensitivity",[10] meaning some people with SLE may be missed, and some people incorrectly classified as SLE, with these threshold criteria. The newer SLICC and ACR/EULAR classification criteria had similar sensitivities/specificities.
- The CS states that "*Patients typically present with symptoms involving the skin and joints, of which pain and fatigue are amongst the most debilitating symptoms interfering with daily life, domestic and professional activities, and social and sexual lives*",[1] but this is from a qualitative study with only 15 participants (from a convenience sample of SLE patients in an outpatient rheumatology clinic in Portugal), and therefore unlikely to be representative of the UK SLE population. However, our clinical expert did confirm that skin and joints are the most common features of lupus and fatigue is the commonest complaint (but usually multifactorial).[39]
- The CS states that "*a large proportion of those with SLE unable to remain in paid employment*",[1] citing two studies.[16, 17] The first study was a review where the most recent study followed participants to 2006 (2001 for studies looking at UK participants), and the second study was conducted with only SLE patients from the USA up to 2004. As treatment for SLE and working patterns (particularly for women, who are much more likely to develop SLE) have evolved over time, and as the USA has a very different relationship between chronic illness and employment compared with the UK, these results may not generalise to the contemporary UK population.[40-42] It should also be noted that as the peak incidence in both men and women are close to the historical retirement age in the UK, it would be useful to know how many years of work people with SLE in the UK lose due to their condition. For reference, the two studies in the review[16] cited by the CS that looked at UK populations found that 54% and 53% of patients were in employment, in 1995-1997 and 1995-2002 respectively.
- The CS stated that "*In a European observational study in patients with SLE, after 10 years of disease duration 72% of patients were receiving ongoing treatment with corticosteroids*",[1] but it should be noted that the study[24] recruited patients in 1994 diagnosed at least 10 years previously, thus the study is over 25 years out of date. The patients were also recruited from all over Europe, with only 29% from Western Europe, and thus the generalisability of the results to present-day UK patients is very limited. This is particularly true, as since 1994 alternative treatments have become available that have reduced the mortality rate for SLE, and therefore likely reducing the need to move on to high-dose corticosteroids (with the highest doses reserved for the most severe disease) by 10 years after diagnosis.[37, 43] Although, the BSR audit shows that steroids are still widely used: "Overall, 497 (48.7%) clinic visits documented prednisolone (including 28.5% of visits when disease was assessed to be inactive".[37] Confirming a current need for additional therapies to control SLE.

2.3 Critique of company's overview of current service provision

There are no published NICE clinical guidelines for the treatment of SLE. There is one published NICE technology appraisal (which this CS reviews), TA397 "Belimumab for treating active autoantibody-positive systemic lupus erythematosus", published in 2016, where IV belimumab was approved for use in adults with HDA-1 SLE.[44] There is also a published NICE evidence summary for oral mycophenolate (ESUOM36).[45] However, both the British Society for Rheumatology (BSR)[39] (NICE accredited) and EULAR[46] have recently published updated guidelines on the management of SLE.

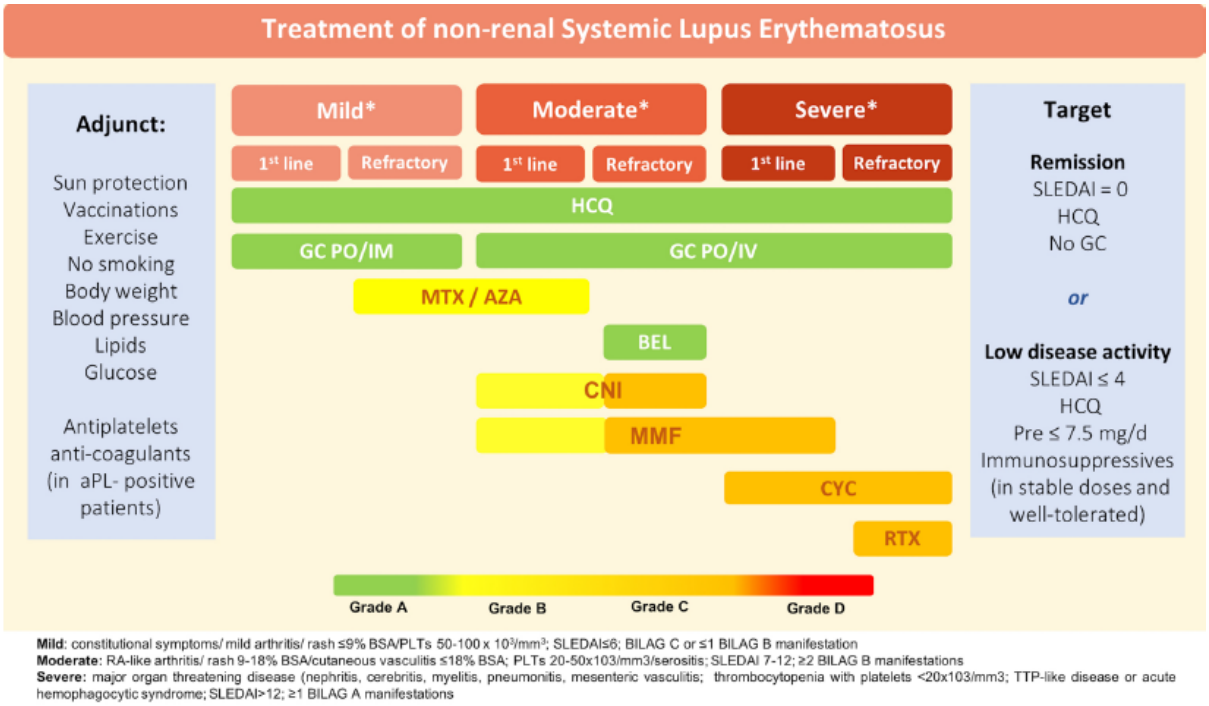
Standard therapy for SLE includes (alone or in combination) antimalarials (e.g. hydroxychloroquine), non-steroidal anti-inflammatory drugs (NSAIDs), corticosteroids and immunosuppressants.[39, 46] When standard therapy fails to control a patient's SLE, this can lead to an increase in corticosteroid use, which can lead to use of belimumab or unlicensed treatments (e.g. rituximab, clinical trials). As the CS states, "*To improve long-term patient outcomes, the overarching aim of treatment should be the remission of disease symptoms and signs, the prevention of flares, the prevention of organ damage accrual, the minimisation of drug side effects, and improvement in patients' quality of life.[46]More specifically preventing flares and maintaining symptoms with the lowest possible dose of glucocorticoids."*[1]

The EULAR guidelines propose that belimumab should be considered in patients with non-renal SLE of moderate severity, defined as: "*rheumatoid arthritis-like rash / rash 9-18% body surface area / cutaneous vasculitis ≤ 18% body surface area / platelets 20-50x10^3/mm^3 / serositis; Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) 7-12; ≥2 British Isles Lupus Assessment Group (BILAG) B manifestations*",[46] that is inadequately controlled on first-line treatments (e.g. hydroxychloroquine, corticosteroids, immunosuppressants), with an inability to taper daily corticosteroids doses to 7.5 mg/day or less. The BSR guidelines[39] for belimumab use are similar to the EULAR guidelines, though suggests considering clinical trials before starting on belimumab. The current EULAR guidelines for the treatment of non-renal SLE are presented in Figure 2.1,[46] and the proposed positioning of belimumab within the clinical pathway of care for SLE from the CS is presented in Figure 2.2.[1]

Figure 2.1 shows the recommended drugs with respective grading of recommendation for systemic lupus erythematosus (SLE) from the 2019 EULAR recommendations. In these recommendations, belimumab is indicated for refractory (not controlled by first-line treatment) moderate SLE.[46]

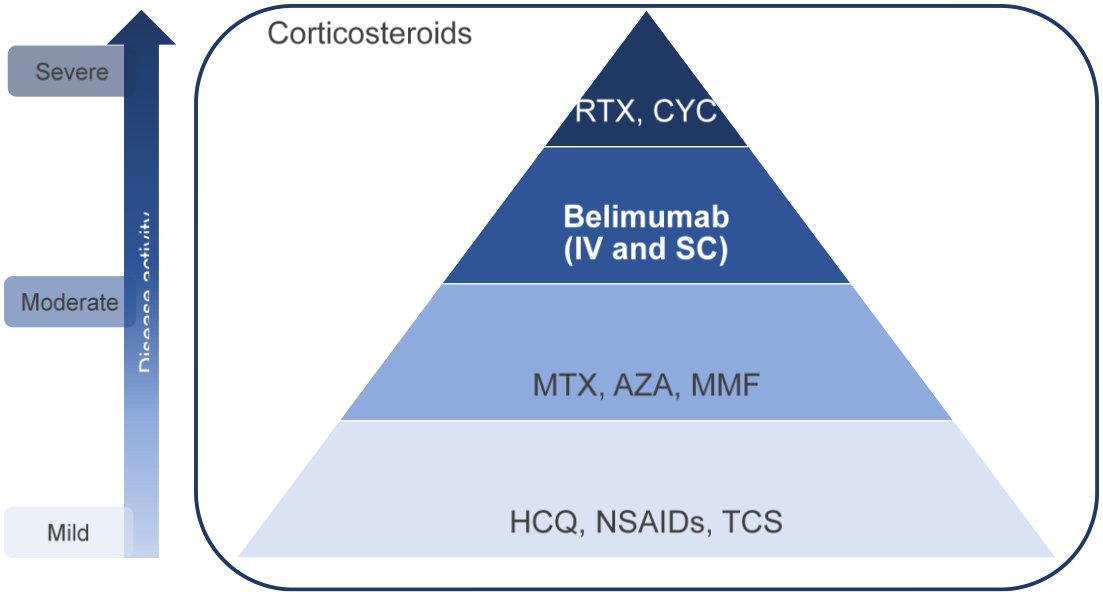
Figure 2.2 shows the proposed positioning of belimumab within the clinical pathway of care for systemic lupus erythematosus (SLE). In the proposed positioning, the company submission (CS) specified belimumab as treatment for moderate-severe SLE, when earlier treatments no longer fully control the disease.

Figure 2.1: Treatment of non-renal SLE—recommended drugs with respective grading of recommendation, from the 2019 EULAR recommendations



Source: 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus[46]
aPL = antiphospholipid antibodies; AZA = azathioprine; BEL = belimumab; BILAG = British Isles Lupus Assessment Group disease activity index; CNIs = calcineurin inhibitors; CYC = cyclophosphamide; GC = glucocorticoids; HCQ = hydroxychloroquine; IM = intramuscular; MMF = mycophenolate mofetil; MTX = methotrexate; Pre = prednisone; PO = per os; RTX = rituximab; PLTs = Platelets; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index.

Figure 2.2: Proposed positioning of belimumab within the clinical pathway of care for SLE



Source: Section B.1.3.5 of the CS[1]
AZA = azathioprine; CYC = cyclophosphamide; HCQ = hydroxychloroquine; IV = intravenous; MMF = mycophenolate mofetil; NSAIDs = non-steroidal anti-inflammatory drugs; MTX = methotrexate; RTX = rituximab; SC = subcutaneous; TCS = topical corticosteroids

In TA397, the SLE patient population under consideration were adults with a Safety of Estrogens in Lupus National Assessment (SELENA) SLEDAI score ≥10, low complement and positive anti-dsDNA, termed the high disease activity 1 (HDA-1) group. The CS proposes that the HDA-1 classification is too restrictive and wishes to extend the patient population to patients five years of age and older with a SELENA-SLEDAI score ≥10 and either low complement or positive anti-dsDNA, termed the HDA-2 group. The CS gives the following reasons for the proposed change: *"patients will often experience levels of high disease activity but only have one of the two defined serological biomarkers. Furthermore, patients who have both biomarkers at the time of diagnosis and are managed with current standard therapies, may subsequently experience normalisation of one of the two serological biomarkers but continue to have high disease activity clinically due to a suboptimal treatment response. Additionally, some patients with high disease activity may have an underlying complement deficiency and therefore access to belimumab would be unattainable with the current criteria."*[1]

ERG comment: The CS overview of current service provision seems appropriate, given the dearth of approved treatments for moderate/severe SLE. Belimumab is already approved for use and the positioning of the drug is not changing; the CS instead wishes to expand the patient population eligible to receive belimumab, both in disease severity and age, which may meet unmet needs in the SLE population given the current restrictions on belimumab use.

In the clarification letter the ERG asked the company whether belimumab will mainly be offered as a third-line treatment (Clarification letter, Question A15).[47] The company responded that it anticipates that belimumab will be offered primarily as third-line treatment after failure of antimalarials and immunosuppressants (both of which may be supplemented with corticosteroids). However, the company also stated that a minority of patients may receive belimumab at second line, in line with the EULAR guidelines which state that "Belimumab should be considered in extrarenal disease with inadequate control (ongoing disease activity or frequent flares) to first-line treatments (typically including combination of an antimalarial and prednisone with or without immunosuppressive agents), and inability to taper glucocorticoids daily dose to acceptable levels (i.e., maximum 7.5 mg/day)". The company noted that pivotal trials of belimumab did not require patients to have received a certain number of prior therapies, or prior therapy with immunosuppressants.

3. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 3.1: 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	People aged 5 years or more with active, autoantibody-positive SLE with a high degree of disease activity despite standard therapy.	Phase 3 Trial Population: Patients with active autoantibody-positive SLE as enrolled in belimumab pivotal trials. High Disease Activity Subgroup-1 (HDA-1): Patients with a SELENA SLEDAI score ≥10 AND low complement AND positive anti-dsDNA (current NICE guidance population; TA397) High Disease Activity Subgroup-2 (HDA-2): patients with a SELENA-SLEDAI score ≥10 AND at least one of the following serological features: low complement OR positive anti-dsDNA – *The Base case*	Mindful of NHS resources, the proposed population of interest to this decision problem is a subgroup of the phase 3 trial population which applies the additional criteria of evidence for high serological (low complement AND positive anti-dsDNA) and clinical (SELENA-SLEDAI score of ≥10) disease activity. This subgroup experienced an additional treatment benefit of belimumab, resulting in the HDA-1 population becoming the recommended population within TA397. Following TA937, data collected as part of the managed access agreement since 2016 through the British Isles Lupus Assessment Group (BILAG) Biologics Registry (BR) has revealed that the number of patients receiving belimumab in England was substantially smaller than anticipated. This suggests that the HDA-1 population was too restrictive when applied in clinical practice and, to better address the unmet need in SLE and more accurately reflect patients with high disease activity, we propose belimumab be considered in the HDA-	The population is 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
			2 population defined as 'patients with a SELENA-SLEDAI score ≥10 AND at least one of the following serological features: low complement OR positive anti-dsDNA'. To support the adoption of the HDA-2 subgroup, it is our proposed base case for the economic modelling. GSK presents the results of PLUTO, the paediatric trial of IV belimumab compared with placebo within an appendix of the submission. The paediatric population recruited in PLUTO is limited (due to the rarity of childhood SLE) and the study was not statistically powered to show a difference between treatment groups. The economic evaluation will not specifically address a paediatric population; all inputted data pertains to an adult population. We assume that the resultant NICE guidance would apply to a paediatric population under the NHS England Commissioning policy for adolescents and paediatrics.	
Intervention	Belimumab as an add-on to standard therapy.	As per the NICE scope. Please note that this submission refers to the previously appraised IV formulation and introduces a new subcutaneous (SC) formulation in the form of a pre-filled pen	SC formulation has been developed as an additional formulation to the currently available IV formulation, to offer physicians and patients a choice of treatment modalities based on the individual's needs, supporting increased access to treatment and	The intervention is 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
			adherence. It also reduces the burden on NHS resources as regular clinic time is not required for administration.	
Comparator(s)	• Standard therapy alone. For people in whom it is considered appropriate: • Rituximab plus standard therapy • Cyclophosphamide plus standard therapy.	Evidence from clinical trials is available versus standard therapy alone; this is presented in this submission. Rituximab Although GSK acknowledges that rituximab would be used in patients eligible for belimumab if belimumab were not made available in the future, we have not conducted a formal indirect comparison versus rituximab. Cyclophosphamide is not included as a comparator.	Rituximab: With the lack of positive RCT data, and limited robust published observational data for rituximab, particularly in terms of long-term follow-up data, no attempt has been made to conduct a formal indirect comparison between rituximab and belimumab. The data provided for rituximab (Appendix P) from the BILAG-BR demonstrates the difficulty in assessment - how patients are managed on rituximab. Although a comparison of the two medicines is provided in Appendix P, these results should be interpreted with caution due to the observational nature of the study. Other statistical techniques, such as a matching adjusted indirect comparison, were not possible, due to the small sample size, particularly for belimumab. Considering rituximab as a comparator is not straightforward. Although rituximab could be used in patients eligible for belimumab if belimumab were not made available in the future, the recently published NHS England commissioning policy for rituximab in	The comparators are not 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
			the treatment of SLE states that belimumab should be considered prior to rituximab in the treatment pathway. Data for rituximab collected from the BILAG-BR are presented in Appendix P to this submission for completeness. Cyclophosphamide: Used to treat patients with severe lupus. It is largely reserved for the treatment of lupus nephritis or CNS lupus, both of which are outside of the current marketing authorisation for belimumab. Therefore, cyclophosphamide plus standard therapy is not a relevant comparator for this appraisal. In addition and as stated by clinical experts in Section 4.3 of TA397adverse effects associated with long-term exposure to cyclophosphamide (bladder cancer, bone marrow suppression, haematologic malignancies, infections, myelodysplasia, and infertility[48])severely limits the use of cyclophosphamide in patients with SLE, who are more often women of childbearing age.	
Outcomes	The outcome measures to be considered include: • disease activity • rate and duration of response	As per the scope, except for the rate and duration of remission.	The rate and duration of remission were therefore not considered to be suitable outcomes in this submission.	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
	• rate and duration of remission • incidence and severity of flares • impact on disease manifestations • incidence of long-term complications and/or organ damage • corticosteroid use • rate and duration of corticosteroid-free remission • mortality • health-related quality of life • 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 QALY. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective.	As per the NICE reference case.	No deviation from NICE scope; however, only the adult SLE population was modelled as described above. The economic analysis used a lifetime horizon and captured relevant direct health effects and costs.	The company's health economic model is in line with the reference case and 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
	The availability of any commercial arrangements for the intervention, comparator and subsequent treatment technologies will be taken into account. The availability of any managed access arrangement for the intervention will be taken into account.			

Source: CS, Table 1, pages 12-15.[1]

BILAG-BR = British Isles Lupus Assessment Group - Biologics Registry; CNS = Central nervous system; GSK = GlaxoSmithKline; HDA = High disease activity; IV = Intravenous; NHS = National Health Service; NICE = National Institute of Health and Care Excellence; QALY = quality-adjusted life year; SC = Subcutaneous; SELENA = Safety of Estrogen in Lupus Erythematosus National Assessment; SLE = Systemic Lupus Erythematosus; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index.

3.1 Population

The population defined in the scope is: "People aged 5 years or more with active, autoantibody-positive systemic lupus erythematosus with a high degree of disease activity despite standard therapy".[49] The scope does not provide a definition for 'a high degree of disease activity'. In the CS, the company provides two definitions:

- High Disease Activity Subgroup-1 (HDA-1): Patients with a SELENA SLEDAI score ≥10 AND low complement AND positive anti-dsDNA (current NICE guidance population; TA397)
- High Disease Activity Subgroup-2 (HDA-2): patients with a SELENA-SLEDAI score ≥10 AND at least one of the following serological features: low complement OR positive anti-dsDNA – The Base case

The current appraisal is different from the original appraisal (TA397[44]) in three ways:

1. The definition of 'high disease activity' (i.e. HDA-1 versus HDA-2, see above)
2. Age: in TA397, belimumab was approved for adults only. This appraisal includes people aged five years or more.
3. Formulation: The original appraisal included an intravenous (IV) formulation only. The current appraisal also includes a new subcutaneous (SC) formulation in the form of a pre-filled pen.

According to the marketing authorisation, belimumab is indicated as add-on therapy in patients aged five years and older with active, autoantibody-positive systemic lupus erythematosus (SLE) with a high degree of disease activity (e.g., positive anti-dsDNA and low complement) despite standard therapy.[50]

ERG comment: In the response to clarification (Clarification letter, Question A11), the company acknowledged that patients with severe active CNS lupus were excluded from the BLISS trials and no evidence to support the use of belimumab is available in this population. Patients with lupus nephritis (LN) were also excluded from the BLISS trials; however, a clinical trial in this population, BLISS-LN, has recently been published,[51] and a Type II variation for an indication extension has been submitted to the EMA on 24th June 2020 and an outcome is anticipated in H1 2021. Therefore, SLE patients with lupus nephritis may be eligible for treatment with belimumab in the future.[47]

In addition, the company confirmed that 'no searches were performed in people over the age of five as the CS focuses on an adult population with SLE as does the economic modelling. The majority of clinical effectiveness data available on belimumab is for adult patients aged 18 years and older with SLE.' (Clarification letter, Question A17).[47]

3.2 Intervention

The intervention (belimumab) is in line with the scope.

IV formulation: The recommended dose regimen is 10 mg/kg on Days 0, 14 and 28, and at four-week intervals thereafter. Premedication including an antihistamine, with or without an antipyretic, may be administered before the infusion of belimumab. The infusions should be administered by a qualified healthcare professional trained to give infusion therapy.[50]

SC formulation: The recommended dose is 200 mg once weekly, administered subcutaneously. Dosing is not based on weight. The recommended injection sites are the abdomen or thigh. When injecting in the same region, patients should be advised to use a different injection site each week.[50]

According to the company, no additional tests or investigations are needed for selection of patients eligible for belimumab treatment other than those currently used routinely in clinical practice (CS, page 17).[1] The patient's condition should be evaluated continuously and discontinuation of treatment with belimumab should be considered if there is no improvement in disease control after six months of treatment.[50]

ERG comment: In the clarification letter, the ERG asked the company what proportion of patients will be using the SC formulation and what proportion will use the IV formulation (Clarification letter, Question A13).[47] The company responded that based on the uptake of belimumab SC in other EU markets, they expect that there will be gradual uptake, resulting in approximately 60% of patients on SC by the end of year one (12 months from date of launch) and 70% of patients on SC by the end of year two. The company does not expect the ratio to exceed 70:30 (SC:IV).[47] However, our clinical expert suggested that the majority of UK patients have switched to SC during the pandemic and so there might be a higher proportion of patients wanting SC in future if the disease appears to be as well controlled on SC as IV.

3.3 Comparators

The description of the comparators in the NICE scope is as follows: "Standard therapy alone. For people in whom it is considered appropriate: rituximab plus standard therapy or cyclophosphamide plus standard therapy".[49]

The company only included one comparator: Standard therapy alone. The company provides several reasons why rituximab was not included as a comparator: lack of data to perform an (indirect) comparison, and the recently published NHS England commissioning policy for rituximab in the treatment of SLE, which states that belimumab should be considered prior to rituximab in the treatment pathway.[1] Cyclophosphamide was not included as a comparator because "it is largely reserved for the treatment of lupus nephritis or CNS lupus, both of which are outside of the current marketing authorisation for belimumab".[1]

In the original appraisal (TA397), the Committee considered that, "because rituximab is provided through individual funding requests and its use in the NHS is likely to be limited, it should not be considered to be the main comparator in routine practice (although it had been specified in the scope for the appraisal). The Committee therefore concluded that standard care should be the main comparator for belimumab, as included in the final scope and in the manufacturer's decision problem" (FAD committee papers, page 29).[52] In addition, "The Committee was aware that cyclophosphamide was also included as a comparator in the scope for the appraisal, but noted the manufacturer's justification that it was largely used for lupus nephritis, which was a different population to the one included in the trials of belimumab and covered by the marketing authorisation for belimumab. Furthermore, it heard from clinical specialists that cyclophosphamide is used infrequently because of side effects" (FAD committee papers, page 29).[52]

ERG comment: We asked our clinical expert whether she thought it was reasonable not to include rituximab and cyclophosphamide as comparators. She responded that the company had a point in that cyclophosphamide is mostly used for neuropsychiatric lupus, severe vasculitis e.g. gut, severe cardiorespiratory involvement, and patients that fail mycophenolate mofetil (MMF) for some of these types of conditions and for renal lupus. In addition, the adverse event profile means that cyclophosphamide is avoided if possible and rituximab is used increasingly for severe manifestations where belimumab treatment criteria are not met or if belimumab has failed. The comparison with rituximab will be difficult according to our clinical expert because the evidence for rituximab is weaker

as the phase 3 trials were negative due to very stringent end-points (and different to those used for belimumab) and is mostly from registries. BILAG BR data cannot be used to compare them easily due to the different criteria for the use of rituximab and belimumab (See BSR guidelines for SLE).

3.4 Outcomes

The NICE final scope lists the following outcome measures:

- disease activity
- rate and duration of response
- rate and duration of remission
- incidence and severity of flares
- impact on disease manifestations
- incidence of long-term complications and/or organ damage
- corticosteroid use
- rate and duration of corticosteroid-free remission
- mortality
- health-related quality of life
- adverse effects of treatment

ERG comment: In the Table describing the decision problem (CS, Table 1, page 14) the company states that *"the rate and duration of remission were therefore not considered to be suitable outcomes in this submission*".[1] However, the company does not explain what the word 'therefore' is based on. In the description of the health condition (CS, Section B.1.3.5, page 21) the company states that *"To improve long-term patient outcomes, the overarching aim of treatment should be the remission of disease symptoms and signs, …"*.[1] Nevertheless, 'the rate and duration of remission' were not considered to be suitable outcomes in the company submission.

In the clarification letter, the ERG asked the company why the rate and duration of remission were not considered suitable outcomes in this submission (Clarification letter, Question A16).[47] The company responded that the rate and duration of remission was not directly measured in clinical trials of belimumab and at present, there is no universally accepted, validated definition of remission in SLE. This is despite the fact that a consensus framework for development of such definitions exists.[53] The EULAR guidelines define remission as the "absence of clinical activity with no use of glucocorticoids and immunosuppressive drugs", and at the same time acknowledge that remission defined in such a way is infrequent.[47] In parallel, clinicians have sought to define a "low disease activity state", which could be a more attainable target than remission for treat-to-target approaches; however, the proposed definitions[54, 55] also vary.[47] Our clinical expert added that as remission is very rare it is reasonable to exclude this as an outcome in this submission.

3.5 Other relevant factors

According to the company, belimumab is innovative because it is a biologic therapy, targeting the BLyS pathway associated with an immune response in SLE, that addresses a substantial unmet need in a chronic and potentially debilitating disease (CS, Section B.2.12).[1]

According to the CS, GlaxoSmithKline is proposing a patient access scheme (PAS) for both belimumab IV and SC. The PAS is in the form of a simple discount which means belimumab IV is offered at a discount of [redacted]% for 120 mg and 400 mg vials. The list price per prefilled pre-filled pen is currently [redacted]. The PAS price per prefilled device is £[redacted]. On 17 December 2020, the ERG received an email from NICE stating that 'the

PAS pricing of the SC formulation is still being discussed. However, the company have received confirmation from the DHSC that they have agreed to a list price for this formulation of "[redacted] for Benlysta (belimumab) 1 pack of 4 sub-cut pens"'.

This appraisal does not fulfil the End-of-Life criteria as specified by NICE because the life expectancy of patients eligible for belimumab is well beyond 24 months. Therefore, treatment is not indicated for patients with a short life expectancy (normally less than 24 months).

According to the company, no equality considerations have been identified (CS, Section B.1.4).[1] However, stakeholders (Prof. Lightstone on behalf of the Renal Association) commented that this is a disease predominantly affecting women and those from black, Asian and minority ethnic (BAME) backgrounds and that they are severely disadvantaged by this chronic disease.[56]

4. CLINICAL EFFECTIVENESS

4.1 Critique of the methods of review(s)

The company updated a previous systematic review (conducted to support the 2011 NICE submission) to identify evidence on '*the clinical effectiveness of treatment possibilities in SLE, specifically the efficacy (including quality of life [QoL]), safety and tolerability of belimumab and appropriate comparators in SLE*.'[57] Section 4.1 critiques the methods of the review including searching, inclusion criteria, data extraction, quality assessment and evidence synthesis.

4.1.1 Searches

Appendix D.1 of the CS details a systematic literature review (SLR) conducted to identify the clinical effectiveness of treatments for SLE, specifically the efficacy, safety and tolerability of belimumab and appropriate comparators in SLE. The searches were initially conducted in 2011 and subsequently updated in 2015. This report describes the latest update, finalised in January 2020. The previous SLR captured studies published from 1970 to August 2010, and the update covered literature from 2010 to January 2020. The search covered studies published from 1 February 2010 to 15 January 2020 to allow a six-month overlap in publication date. Searches were conducted on 15-19 January 2020. Separate searches were conducted in the MEDLINE and Embase databases to identify RCT and non-RCT evidence. A single search was used for the other resources included.

A summary of the sources searched is provided in Table 4.1. In addition to the database searches, the CS states that a manual check of the references of recent (i.e., published in the past three years) relevant SLRs, relevant practice guidelines and conference abstracts from the past two years was performed to supplement the database searches and ensure optimal and complete retrieval. The CS states that the searches were limited to English language studies only (CS, Appendix D; p.3), however the response to clarification[47] confirmed that this limit was applied during the inclusion screening stage.

Table 4.1: Data sources for the clinical effectiveness systematic review (as reported in CS)

	Resource	Host/source	Date ranges	Dates searched
Electronic databases	Embase	Ovid	2010 - 15/1/20 (RCTs) 2010 - 17/1/20 (non-RCTs)	15/1/20 (RCTs) 17/1/20 (non-RCTs)
	MEDLINE/MEDLINE In Process	Ovid	2010 - 15/1/20 (RCTs) 2010 - 17/1/20 (non-RCTs)	15/1/20 17/1/20 (non-RCTs)
	Cochrane Library (CDSR/CENTRAL)	Ovid	2010 - 15/1/20	15/1/20
Conference proceedings	ACR ELR BSR ASN Kidney Week ISPOR	Conference websites & Abstract books	Last two years	Not stated

CDSR = Cochrane Database of Systematic Reviews; CENTRAL = Cochrane Central Register of Controlled Trials; ACR = American College of Rheumatology; ELR = European League Against Rheumatism; BSR = British Society for Rheumatology; ASN = American Society of Nephrology; ISPOR = International Society for Pharmacoeconomics and Outcomes Research

ERG comment:

- Searches were undertaken to identify data published on the clinical effectiveness of treatments for SLE since 2010. The CS provided sufficient details for the ERG to appraise the literature searches. Several databases and a good range of conference proceedings were searched, and reference checking of recent publications was conducted. Searches were generally well documented, making them transparent and reproducible.
- Searches were undertaken in January 2020 for the CS in November 2020, so could now be considered rather out of date. Data published this year may therefore not have been identified by the searches.
- Separate searches were conducted in MEDLINE and Embase for RCT and non-RCT data. As in the 2010 searches, the RCT searches included terms for the intervention (belimumab) and its comparators. The non-RCT searches contained only belimumab search terms, and comparator terms were not included.
- Study design filters were comprehensive and included relevant search terms. They appear similar to those published by SIGN[58], but this was not explicitly stated in the CS. Given the low numbers of records retrieved by the non-RCT searches it may have been more appropriate not to limit by study design, as this runs the risk of losing potentially relevant records. No search filters were applied for searches of the Cochrane Library, as these databases are already limited by study design.
- Targeted searches were conducted for recent SLRs and relevant practice guidelines for reference checking. The CS states that these were for the last three years (App D; p. 3), however in the response to clarification[47] it states that only the last two years were covered.
- Update searches were not conducted on trials registers, although four clinical trials registers were included in the original 2010 searches.
- Additional synonyms could have been used in the strategies to increase recall, for example erythematodes visceralis, lupovisceritis and Libman Sacks disease for SLE, and the CAS registry numbers and additional trade names for belimumab and its comparators. 'Benlysta/' is incorrectly used as an EMTREE heading in the Embase searches, however this will not have affected recall.

4.1.2 Inclusion criteria

The eligibility criteria used in the search strategy for randomised controlled trials (RCTs) and non-RCTs is presented in Table 4.2.

Table 4.2: Eligibility criteria used in search strategy for RCT and non-RCT evidence

	Inclusion criteria	Exclusion criteria
Population	Adults (≥18 years) diagnosed with SLE	• Patients with only active lupus nephritis; included if just kidney involvement in SLE • ≥15% of patients have lupus nephritis • Patients without SLE • Paediatric patients <18 years • Patients with comorbid SLE and rheumatoid arthritis (rhupus)

	Inclusion criteria	Exclusion criteria
Interventions	**RCTS:** • Belimumab • Rituximab • Mycophenolate mofetil • Prednisolone and other steroids • Hydroxychloroquin and other anti-malarials • Azathioprine • Cyclophosphamide • Methotrexate **Non-RCTs:** • Belimumab • Rituximab	Study evaluates treatment other than those listed as interventions of interest in the inclusion criteria.
Comparators	• Rituximab • Cellcept® (mycophenolate mofetil) • Prednisolone and other steroids • Hydroxychloroquine and other anti-malarials • Azathioprine • Cyclophosphamide • Methotrexate • Placebo and mixed routine care (i.e., combination treatments)	Not applicable
Outcomes	**Efficacy:** • Change in SELENA-SLEDAI score • Change in BILAG score • Change in PGA scale • Change in SLICC/SDI score • Change in number/frequency of flares; the scale used to measure flares was recorded in the data extraction table • Quality of life • Mortality • Reduction in steroid use (including definition) • Medical resource utilisation • Fatigue (e.g., FACIT score) **Safety:** • Incidence and severity (grade) of all AEs reported • Withdrawals due to AEs • Incidence of SAEs	Studies that did not report at least one of the outcomes of interest listed in the inclusion criteria

	Inclusion criteria	Exclusion criteria
Study design	Randomised clinical trials, both parallel and crossover. Non-RCT designs, including non-randomised trials, single-arm trials, and observational designs: prospective and retrospective cohorts, cross-sectional, and case-control studies	Crossover designs that did not include adequate washout period (e.g., 7 days) and did not have statistical analysis taking paired design into account. Letters, case reports, editorials, reviews. Pooled studies.

Source: Table 6 of Appendix D[57]
AE= adverse event; BILAG = British Isles Lupus Activity Group; FACIT = Functional Assessment of Chronic Illness Therapy; PGA = Physician Global Assessment; RCT = randomised controlled trial; SAE = serious adverse event; SDI = Systemic Lupus International Collaborating Clinics Damage Index; SELENA = Safety of Estrogens in Lupus National Assessment; SLE = systemic lupus erythematosus; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index; SLICC = systemic Lupus International Collaborating Clinics.

ERG comment: The ERG notes inconsistencies regarding the included population. Patients with severe active lupus nephritis or CNS lupus were excluded from the included BLISS trials (see also Section 3.1 of this report). The company states that this decision was made due to no available evidence to support the use of belimumab in this population.[47] The company identifies a recently published trial, BLISS-LN, which focuses on patients with lupus nephritis, and notes that this population may be eligible for belimumab treatment in the future (Response to clarification, Question 11).[47] However, in the CS, the company states the population was not within the scope for the current appraisal.

In the inclusion criteria, the population was stated to include adults (≥18 years) diagnosed with SLE (see also Section 3.1 of this report). However, the decision problem included patients aged five years or more with SLE. The company confirmed that no searches were performed to identify studies in children and the majority of the clinical effectiveness data is focused on adults (Response to clarification, Question 17).[47]

The ERG attempted to seek clarification regarding the justification of the exclusion of studies if the main participant population included ≥ 15% with lupus nephritis. The company stated this exclusion criteria was based on internal discussion as no studies or selection of studies were identified to guide this criterion (Response to clarification, Question 19).[47]

The ERG attempted to seek transparency regarding the exclusion of studies. The company provided an Excel file indicating the excluded studies and reason for exclusion (Response to clarification, Questions 21 and 22).[47] One study was excluded due to not being published in English. In total 176 studies were excluded. Most study were excluded because the outcomes reported were not separable (N=53), the outcomes were not of interest (N=50), the population was not of interest (N=34) or the intervention was not of interest (n=31).

4.1.3 Critique of data extraction

There was no available information regarding the data extraction process. Appendix D noted that data extraction was used to collect evidence for clinical efficacy and safety, quality of life, and resource utilisation. However, Information regarding the number of authors involved in this stage and how discrepancies were addressed were lacking.

ERG comment: It is normally recommended that two reviewers are involved in data extraction to avoid bias and error. The lack of information regarding how data extraction was addressed presents issues regarding transparency.

4.1.4 Quality assessment

According to Appendix D of the CS, the quality assessment of the RCTs was completed using the using the University of York, Centre for Reviews and Dissemination guidelines.[59] The non-randomised LTE and RWE studies were assessed using the Downs and Black checklist.[60]

ERG comment: It is normally recommended that two reviewers are involved in the assessment of study quality to avoid bias and error.[61] Results of the company's quality assessment and the ERG's assessment are presented in Section 4.2.4.

4.1.5 Evidence synthesis

The company did not conduct a meta-analysis of the belimumab trials. However, they provided pooled results for the HDA subgroups from the two BLISS trials (BLISS-52 and BLISS-76).

ERG comment: The pooled results for BLISS-52 and BLISS-76 used a combined individual patient level dataset. The company stated that both trials were very similar with almost identical designs, inclusion and exclusion criteria and analysis methods. They were also run in similar time periods. All analyses were adjusted for study in the relevant regression model to account for the fact that there were two trials. The analysis of the primary endpoint, SRI-4 included a test for a treatment by study interaction, to evaluate if the treatment effect differed by study, and this showed no evidence of an interaction. The ERG considers that this was an appropriate analysis method.

4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

The CS presented three pivotal RCTs of belimumab (BLISS-52,[62] BLISS-76[63] and BLISS-SC[64]). Each of the trials had an extension study.[65-67] Of this evidence, only the three RCTs and one of the extension studies[66] were included in the economic model. All three of the RCTs provided evidence for the two HDA subgroups presented in the CS. The three RCTs will be the focus of this section.

The CS included an indirect cohort comparison study which compared BLISS-76 US LTE and the Toronto Lupus Cohort[29] to assess long-term organ damage in patients treated with belimumab.[19, 68] This is discussed in Section 4.3 of this report 'Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison'.

In addition to the previously mentioned LTE studies, the CS included a range of studies as supporting evidence. These studies did not assess HDA subgroups separately and were not included in the economic model. They are therefore discussed (along with the LTE studies) more briefly in Section 4.2.8 'Supporting evidence'.

As there is a large amount of evidence relevant to this submission, the ERG presents an overview in Table 4.3 including the location of each study in this report.

Table 4.3: Overview of the evidence in the CS

Study name	Description	Population	In the previous NICE submission?	Location in the current submission	Location in this report
BLISS-52[62] and BLISS-76[63]	Pivotal trials of IV belimumab	Total[&]*	Yes	Appendix L	Sections 4.2.1 to 4.2.7
		HDA-1[&]*	Yes	Document B Section 2.7	Sections 4.2.1 to 4.2.7
		HDA-2[&*]	No	Document B Section 2.7	Sections 4.2.1 to 4.2.7
BLISS-76 US LTE[66]	LTE study of US patients enrolled in BLISS-76	Total*	No	Document B Section 2.6 and Appendix M	Section 4.2.8
BLISS-52/76 non-US LTE[65]	LTE study of non-US patients enrolled in BLISS-52 or BLISS-76	Total	No	Document B Section 2.6 with further details in Appendix M	Section 4.2.8
BLISS-SC[64]	Pivotal trial of SC belimumab	Total*	No	Document B Section 2.6	Sections 4.2.1 to 4.2.7
		HDA-1*	No	Document B Section 2.7	
		HDA-2*	No	Document B Section 2.7	
BLISS-SC LTE[67]	Open-label extension for patients enrolled in BLISS-SC	Total	No	Document B Section 2.6 and Appendix M	Section 4.2.8
Phase 2B open-label, single-arm, repeat-dose study[69]	Study to evaluate the reliability of the SC autoinjector	Total	No	Appendix O	Section 4.2.8
LBSL02 Phase 2 trial[70]	Initial evidence on safety and efficacy of belimumab	Total	Yes	Appendix L	Section 4.2.8
LBSL02 LTE[71]	Data on long-term (up to 13 years) experience with belimumab	Total	Partially (further data available with additional follow-up)	Appendix M	Section 4.2.8
BASE (post-marketing)[72]	Safety study of mortality and adverse events of special interest	Total	No	Document B Section 2.6, 2.10 and App. F	Section 4.2.8

Study name	Description	Population	In the previous NICE submission?	Location in the current submission	Location in this report
Treatment holiday study (NCT02119156)[73]	A study of the effect of treatment holiday on belimumab efficacy	Total	No	Appendix O	Section 4.2.8
EMBRACE (post-marketing)[74]	Placebo-controlled trial of belimumab in people of black race	Total	No	Appendix O	Section 4.2.8
NCT01345253[75]	Placebo-controlled trial of belimumab in people from North-East Asia	Total	No	Appendix O	Section 4.2.8
PLUTO[76]	Belimumab in children and adolescents	Total	No	Appendix O	Section 4.2.8
BILAG-BR[77, 78]	UK-based registry of biologic therapy (including belimumab) for SLE	HDA-1 (belimumab data only)	No	Document B Section 2.7 and Appendix P	Section 4.2.8
OBSErve[79, 80]	A multi-country Evaluation Of use of Belimumab in clinical practice Settings	Total	No	Document B Section 2.6	Section 4.2.8
SLICC (ACR)/SDI Indirect Cohort Comparison Study (206347)	A PSM comparative analysis between BLISS-76 US LTE and the Toronto Lupus Cohort to assess long-term organ damage in patients treated with belimumab	Total*	No	Document B, Section 2.6 and Section B.3.3.6	Sections 4.3 and 4.4
ITC between SC and IV belimumab formulations[81]	To compare the efficacy of SC and IV belimumab formulations in patients with autoantibody-positive SLE with HDA	Total	No	Appendix O	Section 4.2.8

Source: Table 12 of the CS[1] and Appendix D[57]

[&] pooled across BLISS-52 and BLISS-76; *included in the economic model

4.2.1 Details of the included pivotal trials

The company submission focussed on three randomised controlled trials: BLISS-52[62], BLISS-76[63] and BLISS-SC[64].

The main evidence for the clinical effectiveness of belimumab was from two phase III clinical trials. The BLISS-52 (n=865) and BLISS-76 (n=819) trials were randomised, double-blind, placebo-controlled, parallel-group studies with follow-up at 52 weeks and 76 weeks respectively. In these trials, belimumab plus standard care was compared with placebo plus standard care. Standard care included: non-steroidal anti-inflammatory drugs, antimalarials, corticosteroids or other immunosuppressants (azathioprine, methotrexate and mycophenolate mofetil) either alone or in combination. Although each of the BLISS trials were three-arm trials (belimumab 10 mg/kg, belimumab 1 mg/kg and placebo), only results for the 10 mg/kg belimumab dose were presented in the company's submission because this is the dose covered by the marketing authorisation. BLISS-52 was undertaken mainly in Asia and South America while BLISS-76 patients mainly derived from North America and Europe. The purpose of the BLISS-52 and BLISS-76 trials was to investigate the IV formulation of belimumab, which was presented to NICE as the main evidence for TA397.

Adult patients (aged 18 years or older) who met the American College of Rheumatology criteria for systemic lupus erythematosus and had active autoantibody-positive disease and a SELENA-SLEDAI score of six or more at screening were eligible for enrolment in the BLISS trials. Patients with severe active lupus nephritis or central nervous system lupus were excluded from the trials. Of the patients in the standard care and belimumab 10 mg/kg arms (n=1,125), 52% (n=585) had disease that met the criteria for the marketing authorisation and 35% (n=396) had disease that met the criteria for the target population in TA397 (HDA-1); 47% (n=532) had disease that met the criteria for the target population in the current submission (HDA-2).

The BLISS-SC trial was presented in this submission to introduce the SC formulation of belimumab. The BLISS-SC trial is an international multicentre phase III randomised placebo-controlled trial lasting 52 weeks. Patients were randomised to belimumab 200 mg SC once weekly plus standard treatment (ST) or placebo plus ST.

All three trials used the same primary efficacy endpoint, which was the SRI-4 response rate at Week 52.

A summary of the methodology of the three main belimumab trials is presented in Table 4.4 below.

Table 4.4: Summary of the methodology of the pivotal belimumab trials

	BLISS-SC[64]	BLISS-52[62] and BLISS-76[63]
Trial design	Phase 3, multicentre, international, randomised, double-blind, placebo-controlled, 52-week study.	Phase 3, randomised, multicentre, international, double-blind, placebo-controlled, parallel-group study. BLISS-52 was 52 weeks and BLISS-76 was 76 weeks in duration.
Participant eligibility criteria	Patients with a clinical diagnosis of SLE according to the ACR criteria and clinically active SLE disease, defined as a SELENA SLEDAI disease activity score of ≥8 at screening. Patients with severe lupus kidney disease, severe active lupus nephritis or CNS lupus were excluded.	Patients with a clinical diagnosis of SLE according to the ACR criteria and clinically active SLE disease, defined as a SELENA SLEDAI disease activity score of ≥6 at screening. Patients with severe active lupus nephritis or CNS lupus were excluded.
Settings and locations where the data were collected	30 countries in North America, Central America, South America, Western Europe, and. Eastern Europe. The trial was run in the UK (6 patients enrolled in 3 centres). Locations were hospital settings, academic institutions (i.e. University hospitals), medical centres, rheumatology departments	BLISS-52: 13 countries in Latin America, Asia-Pacific and eastern Europe BLISS-76: 19 countries in North America and Europe (including the UK). Locations were hospital settings, academic institutions (i.e. University hospitals), medical centres, rheumatology departments.
Intervention	Belimumab 200 mg administered by SC injection on Day 0 and then weekly (i.e., every 7 days ± 1 day) through 51 weeks, plus ST (N=556).	BLISS-52: Belimumab 1 mg/kg (N=288) or belimumab 10 mg/kg (N=290) administered by IV infusion on Days 0, 14, 28 and every 28 days thereafter plus ST BLISS-76: belimumab 1 mg/kg (N=271) or belimumab 10 mg/kg (N=273) administered by IV infusion on Days 0, 14, 28 and every 28 days thereafter plus ST.
Comparator	Placebo administered by SC injection on Day 0 and then weekly (i.e., every 7 days ± 1 day) through 51 weeks, plus ST (N=280).	BLISS-52: Placebo (N=287) administered by IV infusion on Days 0, 14, 28 and every 28 days thereafter plus ST BLISS-76: Placebo (N=275) administered by IV infusion. on Days 0, 14, 28 and every 28 days thereafter plus ST.
Primary outcome	The primary efficacy endpoint was SRI-4 response rate at Week 52. SRI-4 response was defined as: • ≥4-point reduction from baseline in SELENA-SLEDAI score, AND: • No worsening (increase of <0.30 points from baseline) in PGA, AND:	

	BLISS-SC[64]	BLISS-52[62] and BLISS-76[63]
	• No new BILAG A organ domain score or 2 new BILAG B organ domain scores compared with baseline, at the time of assessment (i.e., at Week 52).	
Other outcomes used in the economic model / specified in the scope	• Disease activity: Change in PGA and SELENA-SLEDAI score • Rate and duration of response: SRI-4 response by visit, and at Week 52 (primary efficacy endpoint) • Incidence and severity of flares: Time to SFI flare/severe flare and rate of SFI flare/severe flare per 100 subject years • Incidence of long-term complications and/or organ damage: SELENA-SLEDAI and BILAG, scores by visit; SDI change at Week 52 • Corticosteroid use: Mean/median changes in steroid dose by visit; percent of patients whose average prednisone use reduced by ≥25% to ≤7.5 mg/day • Mortality: not assessed as an outcome, although included in safety reporting • HRQoL: FACIT-Fatigue Scale at Week 52 and by visit. • Adverse effects of treatment: monitoring of adverse events, clinical laboratory tests, vital signs, physical examinations, and immunogenicity.	• Disease activity: Change in PGA and SELENA-SLEDAI score • Rate and duration of response: SRI-4 response by visit, and at Week 52 (primary efficacy endpoint) • Incidence and severity of flares: Time to SFI flare, Time to first flare, number and rate of flares • Incidence of long-term complications and/or organ damage: Change in SELENA-SLEDAI, BILAG scores, and change in SDI at Week 52 • Corticosteroid use: Percent of patients whose average prednisone use reduced by ≥25% to ≤7.5 mg/day • Mortality: not assessed as an outcome, although included in safety reporting • HRQoL: FACIT-Fatigue Scale, SF-36, and EQ-5D at Week 52. • Adverse effects of treatment: monitoring of adverse events, clinical laboratory tests, vital signs, physical examinations, and immunogenicity • For BLISS-76, SRI-4 response rate at Week 76.

Source: CS Table 14[1]

ACR = American College of Rheumatology; BILAG = British Isles Lupus Assessment Group; CNS = central nervous system; EQ-5D = European Quality of Life-5 Dimensions; FACIT = Functional Assessment of Chronic Illness Therapy; HRQoL = health-related quality of life; IV = intravenous; PGA = Physician's Global Assessment; SC = subcutaneous; SDI = Systemic Lupus International Collaborating Clinics (SLICC)/ACR Damage Index; SELENA = Safety of Estrogen in Lupus Erythematosus National Assessment; SFI = SLE Flare Index; SLE = Systemic Lupus Erythematosus; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index; SRI-4 = SLE responder index-4; ST = Standard therapy; UK = United Kingdom; US =United States.

ERG comment: The same critique as in the original appraisal stands: The SLE population in BLISS-76 is more likely to resemble that in the UK than that in the BLISS-52; therefore the BLISS-76 results are probably more relevant to the decision problem than those from BLISS-52. Patients were required to be ≥ 18 years of age in both the BLISS-52 and BLISS-76 trials which therefore did not include any paediatric patients. There were no UK patients enrolled in BLISS-52. In BLISS-76, a total of 11 patients from the UK were enrolled, constituting 1.3% of the total trial population. Of those, six patients were randomised to placebo, four to the unlicensed 1 mg/kg belimumab dose and one to the licensed 10 mg/kg dose.

According to the company, the population enrolled in the BLISS trials is representative of patients with moderate to severe, active SLE in the UK (see Response to clarification, Question A28).[47] The baseline demographics of patients taking part in the BLISS trials were similar to those receiving belimumab in the UK and enrolled in BILAG-BR. Patients in both BLISS trials and BILAG-BR were predominantly females of working age, which is also consistent with the patterns of lupus incidence in general. The racial distribution was slightly different, with the BLISS trials including fewer African Heritage patients compared with the BILAG-BR. In addition, the proportion of White/Caucasian patients was similar between the UK BILAG-BR study and BLISS-SC but was higher in the pooled IV population. The proportion of Asian patients was similar. Disease activity (as described by SELENA SLEDAI/SLEDAI-2K and BILAG scores) and daily steroid dose both appeared higher in the BILAG-BR than the BLISS trials; however, this is likely to be a reflection of the BILAG-BR only collecting data on a subgroup of patients with high disease activity (HDA-1), whereas a broader population was enrolled in the BLISS trials.

As mentioned in Section 3.4 of this report, the rate and duration of remission were not considered suitable outcomes in the submission.

In the BASE trial it was found that belimumab had a possible link with depression. Patients with depression were not specifically excluded from BLISS-52 and BLISS-76. The BLISS-SC trial excluded patients at high risk of suicide. In addition the company stated (Response to clarification, Question A30): "In line with the summary of product characteristics, the risk of depression and suicide should be assessed before initiating treatment with belimumab and monitored during treatment".[47]

4.2.2 Statistical analyses of the pivotal trials

All three BLISS trials were superiority trials designed to demonstrate that belimumab 200 mg SC was superior to placebo in BLISS-SC and that belimumab 10 mg/kg IV was superior to placebo in BLISS-52 and -76. All three trials had the same primary efficacy endpoint of SRI-4 response at week 52 which was a binary outcome based on a composite of a reduction in SELENA-SLEDAI score, no worsening in PGA and no new BILAG A or two new BILAG B organ domain scores (see Table 4.3). All three trials also had long-term extension (LTE) phases including those patients who had completed the double-blind phase of the original trial, all patients received belimumab during the LTE.

The analysis of the primary endpoint used a logistic regression model which adjusted for treatment and the randomisation stratification factors. The analysis was performed on the intention to treat population defined as all patients who were randomised and received at least one dose of study treatment.

Table 4.5: Summary of statistical analyses in the pivotal trials of belimumab

	BLISS-SC[64]	**BLISS-52[62] and BLISS-76[63]**
Hypothesis objective	Demonstrate superiority of belimumab 200 mg SC over placebo when comparing the SRI-4 response at Week 52.	Demonstrate superiority of belimumab 10 mg/kg IV over placebo when comparing the SRI-4
Statistical analysis	The proportion of patients achieving a treatment response at Week 52 was compared between belimumab and placebo using a logistic regression model. The independent variables in the model included treatment groups, baseline SELENA-SLEDAI score, complement level and race. The analysed population was the same as BLISS-52 and BLISS-76, i.e. patients who were randomised and received ≥1 dose of study treatment. For the primary and the major secondary efficacy endpoints, a step-down sequential testing procedure was used to control the overall type 1 error rate. With this procedure, the primary and two major secondary endpoints were evaluated for statistical significance (2 sided alpha=0.05) based on a pre specified sequence for interpretation: (1) SRI-4 response rate at Week 52, (2) time to first severe SLE flare, and (3) percentage of patients with average prednisone dose that has been reduced by ≥25% from baseline to ≤7.5	The percentage of patients achieving a response at Week 52 was compared between belimumab 10 mg/kg and placebo using a logistic regression model. The independent variables in the model included treatment groups, baseline SELENA-SLEDAI score (≤9 vs ≥10), baseline proteinuria level (<2 g/24 hour vs ≥2 g/24 hour equivalent) and race (African descent or indigenous-American descent versus other). The population analysed was defined as for BLISS-SC.
Sample size, power calculation	The study aimed to randomise and treat approximately 816 patients, with a target of at least 544 patients in the belimumab arm and 272 patients in the placebo arm. This sample size provided at least 90% power at a 5% level of significance to detect a minimum of an evidence based 12% absolute improvement in the response rate for the belimumab group relative to the placebo group at Week 52.	Both BLISS-52 and BLISS-76 studies aimed to randomise approximately 810 patients (per study), with a target of at least 270 patients per treatment group (per study). This sample size provided at least 90% power at a 5% level of significance to detect a minimum of a 14% absolute improvement in the response rate in the 10 mg/kg belimumab group relative to the placebo group at Week 52.
Data management, patient withdrawals	Similar across BLISS-52/76/SC: For the SRI-4 endpoint and its components, any patient who was classified as a treatment failure was considered a non-responder for the primary efficacy analysis and the supportive analyses of the primary efficacy endpoint. A treatment failure was defined as any patient who: withdrew from the study prior to Week 52 and had no visit within ±28 days of Week 52, and/or received a protocol-prohibited medication or a dose of allowable (but protocol-restricted) medication that resulted in treatment failure designation prior to Week 52.	

	BLISS-SC[64]	BLISS-52[62] and BLISS-76[63]
Source: Table 20 of the CS[1] IV = intravenous; SELENA = Safety of Estrogen in Lupus Erythematosus National Assessment; SC = subcutaneous; SLE = systemic lupus erythematosus; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index; SRI-4 = SLE responder index-4.		

ERG comment: The analysis of the BLISS trials used appropriate statistical methods and the ERG has no concerns.

4.2.3 Baseline characteristics of the pivotal trials

Table 4.6 shows the characteristics of the participants in the BLISS trials that met the HDA-2 criteria (See Section 3.1 of this report for definitions of the HDA-1 and HDA-2 populations).

Briefly, the HDA-2 population of the BLISS-SC trial had a total of 437 participants and of the pooled BLISS-52 and BLISS-76 trials had 532 participants (HDA-2 population only). The mean age across all trials was approximately [redacted]. Female and male participants were included in all trials, though [redacted] of participants were female ([redacted] for BLISS-SC). All trials had participants from a diverse mix of ethnicities ([redacted]). All BLISS trials were conducted in multiple countries: BLISS-SC was conducted in 30 countries in the Americas and Europe (six patients from three centres in the UK); BLISS-52 was conducted in 13 countries in Latin America, Asia-Pacific and Eastern Europe; BLISS-76 was conducted in 19 countries in North America and Europe (including the UK). In all BLISS trials the patients had a mean disease duration of [redacted], with few patients [redacted] having no A or B BILAG organ domain involvement. All patients had a SELENA-SLEDAI category of 10-11 or 12+, with a mean score of around [redacted]. Most patients ([redacted]) had a high daily prednisolone dose (>7.5 mg/day), with few patients not on prednisolone at all (around [redacted]).

Table 4.6: Baseline characteristics in the pivotal trials of belimumab – HDA-2 population

Baseline characteristics	BLISS-SC[64]		Pooled BLISS-52 and BLISS-76 data[1]	
	Belimumab 200mg SC (n=296)	Placebo (n=141)	Belimumab 10 mg/kg IV (n=262)	Placebo (n=270)
Demographics				
Female, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Age (years), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Age ≤45 years, n (%)				
SCN1A mutation, n (%)				
Race, n (%)				
White	[redacted]	[redacted]	[redacted]	[redacted]
Asian	[redacted]	[redacted]	[redacted]	[redacted]
Black	[redacted]	[redacted]	[redacted]	[redacted]
Alaska Native or Am. Indian from North/ Central/South America	[redacted]	[redacted]	[redacted]	[redacted]
Native Hawaiian or Other Pacific Islander	[redacted]	[redacted]	[redacted]	[redacted]
Multiracial	[redacted]	[redacted]	[redacted]	[redacted]

Baseline characteristics	BLISS-SC[64]		Pooled BLISS-52 and BLISS-76 data[1]	
	Belimumab 200mg SC (n=296)	Placebo (n=141)	Belimumab 10 mg/kg IV (n=262)	Placebo (n=270)
Disease characteristics				
SLE disease duration (years), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
BILAG organ domain involvement, N (%)				
At least 1A or 2B	[redacted]	[redacted]	[redacted]	[redacted]
At least 1A	[redacted]	[redacted]	[redacted]	[redacted]
At least 1B				
No A or B	[redacted]	[redacted]	[redacted]	[redacted]
SELENA-SLEDAI category, N (%)				
≤ 9	[redacted]	[redacted]	[redacted]	[redacted]
10-11	[redacted]	[redacted]	[redacted]	[redacted]
≥12	[redacted]	[redacted]	[redacted]	[redacted]
SELENA-SLEDAI score, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
SLE Flare Index (SFI), N (%)				
At least 1 flare	[redacted]	[redacted]	[redacted]	[redacted]
At least 1 severe flare	[redacted]	[redacted]	[redacted]	[redacted]
Severe flare				
PGA Category, N (%)				
0–1				
<1				
1-<2				
≥2				
>1-2.5				
>2.5				
Missing				
PGA, N Mean (SD)				
SDI score, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
SDI score =0, N (%)				
SDI score =1, N (%)				
SDI score ≥2, N (%)				

Baseline characteristics	BLISS-SC[64]		Pooled BLISS-52 and BLISS-76 data[1]	
	Belimumab 200mg SC (n=296)	Placebo (n=141)	Belimumab 10 mg/kg IV (n=262)	Placebo (n=270)
Proteinuria category (g/24 h), N (%)				
≥2	[redacted]	[redacted]	[redacted]	[redacted]
Proteinuria level (g/24 h), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Clinical characteristics				
Low C3 and/or C4 n (%)				
No	[redacted]	[redacted]	[redacted]	[redacted]
Yes	[redacted]	[redacted]	[redacted]	[redacted]
Positive Anti-dsDNA n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Biomarker levels				
Anti-dsDNA (IU/mL), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
C3 (mg/dL)	[redacted]	[redacted]	[redacted]	[redacted]
C4 (mg/dL)	[redacted]	[redacted]	[redacted]	[redacted]
Medication usage				
Average daily prednisone dose, N (%)				
0 mg/day	[redacted]	[redacted]	[redacted]	[redacted]
>0–≤7.5 mg/day	[redacted]	[redacted]	[redacted]	[redacted]
>7.5 mg/day	[redacted]	[redacted]	[redacted]	[redacted]
Average daily prednisone dose (mg/day), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Number (%) of patients taking:				
Steroid only	[redacted]	[redacted]	[redacted]	[redacted]
Immunosuppressant only	[redacted]	[redacted]	[redacted]	[redacted]
Anti-malarial only	[redacted]	[redacted]	[redacted]	[redacted]
Steroid and immunosuppressant	[redacted]	[redacted]	[redacted]	[redacted]
Steroid and anti-malarial	[redacted]	[redacted]	[redacted]	[redacted]
Immunosuppressant and anti-malarial	[redacted]	[redacted]	[redacted]	[redacted]
Steroid and immunosuppressant and anti-malarial	[redacted]	[redacted]	[redacted]	[redacted]

Baseline characteristics	BLISS-SC[64]		Pooled BLISS-52 and BLISS-76 data[1]	
	Belimumab 200mg SC (n=296)	Placebo (n=141)	Belimumab 10 mg/kg IV (n=262)	Placebo (n=270)
NSAIDs				
Aspirin				

Source: Table 37 of the CS[1] and Response to Clarification (Question A23).[47]
a) Patients who checked more than 1 race category are counted under individual race category according to the minority rule as well as the multiracial category.
Note: Greyed boxes indicate that the category was not measured within the trial.
BILAG = British Isles Lupus Assessment Group; ITT = intention-to-treat; IV = intravenous; NSAID = non-steroidal anti-inflammatory drug; PGA = Physician's Global Assessment; SC = subcutaneous; SD = standard deviation; SDI = Systemic Lupus International Collaborating Clinics (SLICC)/ACR Damage Index; SELENA = Safety of Estrogen in Lupus Erythematosus National Assessment; SFI = SLE Flare Index; SLE = Systemic Lupus Erythematosus; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index.

ERG comment:

Baseline characteristics of HDA-2 group issues.

In all BLISS trials, the vast majority of patients were female: over [redacted]% in BLISS-52 and BLISS-76, and over [redacted]% in BLISS-SC. While this reflects the gender distribution of SLE, it limits statistical power to determine whether belimumab is as effective in men as in women, especially for the SC formulation. The ERG asked for the results of subgroup analyses by gender (Response to clarification, Question A29).[47] The interaction between treatment and gender was included in the logistic regression model for the primary endpoint, SRI-4. For BLISS-SC there was no significant interaction between treatment and gender but there were only [redacted] male patients included. In men the response rates were [redacted] ([redacted]%) for placebo and [redacted] ([redacted]%) for belimumab (OR [redacted], 95% CI [redacted]) and the corresponding response rates in women were [redacted] ([redacted]%) for placebo and [redacted] ([redacted]%) for belimumab (OR [redacted], 95% CI [redacted]). This indicates that belimumab was beneficial in women but there was a lack of evidence for men.

In the analysis of the pooled BLISS-52 and -76 data the treatment x gender interaction was also not significant in the model. SRI-4 response rates in men were [redacted] ([redacted]%) for placebo [redacted] ([redacted]%) for belimumab 10 mg (OR [redacted], 95% CI [redacted]). The corresponding results for women were [redacted] ([redacted]%) for placebo and [redacted] ([redacted]%) for belimumab 10 mg (OR [redacted], 95% CI [redacted]). Both sets of results indicate that belimumab was beneficial in women but there was a lack of evidence for men.

Further subgroup analyses of belimumab-SC showed little difference in effect by body weight or body-mass index (these analyses had reasonable power), and a subgroup analysis for age compared patients above and below 65 years but had limited statistical power, as only [redacted] participants were 65 years or older. Further subgroups analyses of belimumab-IV (10 mg/kg) showed little difference in effect between people with more or less than 15 U/mL baseline anti-Sm, more or less than 1.64 baseline ANA titre, baseline immunosuppressive use, or being younger or older than 45 years (all analyses had reasonable statistical power). Patients with White/Caucasian, Asian and Alaska Native/American Indian ethnicities had similar positive effects of belimumab-IV, though Black/African American patients did better on placebo, though the difference was not statistically significant.

Medication history was not collected in the BLISS trials, so the number of prior therapies, or lines of treatment received, is not available. As such, it is not possible to determine whether belimumab was first, second, or third line treatment for patients in any of the BLISS trials, and so whether the effectiveness of belimumab differs by line of treatment.

ERG comment: These results show that belimumab was effective in women but due to small numbers of men included in the BLISS trials there was a lack of evidence for men. However, it should be noted that the subgroup results apply to the full BLISS populations and not the HDA-2 subgroup.

4.2.4 Risk of bias assessment of the pivotal trials

The company assessed the quality of the three pivotal trials using the University of York, Centre for Reviews and Dissemination criteria.[59] Elements assessed were randomisation, allocation concealment, baseline comparability, care provider, participant and outcome assessor blinding, dropout imbalances, selective outcome reporting and use of intention to treat analysis. No information was provided on the number of reviewers who assessed the quality of included studies. The company concluded that all elements had been appropriately addressed in all three of the trials.

Table 4.7: Quality assessment of BLISS-52 and BLISS-76

BLISS-52 and BLISS-76	How is the question addressed in the study?	Company	ERG
Was randomisation carried out appropriately?	Patients who underwent all screening procedures and met the entry criteria were enrolled in the study and assigned to treatment by use of a central interactive voice response system. Patients were randomised in a 1:1:1 ratio to placebo, or belimumab 1 mg/kg or 10 mg/kg. Randomisation was stratified according to the SELENA-SLEDAI score (6–9 vs ≥10), proteinuria concentration (<2 g/24 h vs ≥2 g/24 h) at screening, and ethnic origin (African descent or indigenous American [Alaska Native or American Indian from North, South, or Central America] vs other).	Yes	Yes
Was the concealment of treatment allocation adequate?	An unmasked pharmacist prepared unmarked infusion bags for administration. Belimumab and placebo were both prepared as sterile and lyophilised vials (5 mL for belimumab 1 mg/kg; 20 mL for belimumab 10 mg/kg and placebo), and contained the same formulations, except without the active drug for placebo.	Yes	Yes
Were the groups similar at the outset of the study in terms of prognostic factors?	The three groups did not differ in any of the main baseline characteristics.	Yes	No
Were the care providers, participants, and outcome assessors blind to treatment allocation?	Patients, investigators, study coordinators, and sponsors were masked to treatment assignment during intravenous administration of the drug and assessment of the patients every 4 weeks during the 52-week trial until the database was locked.	Yes	Yes
Were there any unexpected imbalances in dropouts between groups?	The three groups did not differ in reasons for discontinuation of treatment.	No	Yes
Is there any evidence to suggest that the authors measured more outcomes than they reported?	The authors reported outcomes as specified in the study protocol.	No	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Analysis was done in a modified intention-to-treat population, defined as all randomly assigned patients who received a dose of the study drug. This was appropriate and appropriate methods for handling missing data were outlined in the clinical study report.	Yes	Yes

Source: Table 8 of the CS, Appendix D.[57]
NA = not applicable; RCT = randomised controlled trial; SELENA = Safety of Estrogens in Lupus National Assessment; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index.

Table 4.8: Quality assessment of BLISS-SC

BLISS-SC	How is the question addressed in the study?	Company	ERG
Was randomisation carried out appropriately?	Once patients had consented, undergone all screening procedures and been determined to be eligible for the study, they returned for the Day 0 visit to be randomly assigned (via an EDC-based interactive web response system [IWRS]) to 1 of 2 treatment groups (200 mg belimumab or placebo control) in a 2:1 ratio. The randomisation was stratified by screening SELENA-SLEDAI score (8-9 vs ≥ 10), complement level (C3 and/or C4 low vs other) and race (black vs other).	Yes	Yes
Was the concealment of treatment allocation adequate?	Study agent was supplied in blinded kits of prefilled syringes containing either belimumab or placebo. Following randomisation via the IWRS, the pharmacist received a specific kit number from the system.	Yes	Yes
Were the groups similar at the outset of the study in terms of prognostic factors?	Baseline disease activity was generally comparable between treatment groups, except for a slightly greater percentage of patients in the placebo group with at least 1A or 2B BILAG organ domain involvement, at least 1 flare (as measured by the SLE flare index), and proteinuria >2 g/24 h.	Yes	No
Were the care providers, participants, and outcome assessors blind to treatment allocation?	All study site personnel, the patient, the sponsor and the contract research organisation (CRO) remained blinded to the study agent received.	Yes	Yes
Were there any unexpected imbalances in dropouts between groups?	The three groups did not differ in reasons for discontinuation of treatment.	No	Yes
Is there any evidence to suggest that the authors measured more outcomes than they reported?	The authors reported outcomes as specified in the study protocol.	No	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Analysis was done in the intention-to-treat (ITT) population, defined as all patients who were randomised and treated with at least one dose of study treatment. The ITT analysis was performed according to the treatment that a patient was randomised to receive, regardless of the actual treatment received. This was appropriate and appropriate methods for handling missing data were outlined in the clinical study report.	Yes	Yes

Source: Table 8 of the CS, Appendix D.[57]
NA = not applicable; RCT = randomised controlled trial; SELENA = Safety of Estrogens in Lupus National Assessment; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index.

ERG comment:

- It is normally recommended that two reviewers are involved in the assessment of study quality to avoid bias and error.
- The ERG examined the clinical study reports (CSRs) for the three trials and assessed the trials against the above criteria. Randomisation and allocation concealment procedures appeared to be appropriate. Methods to ensure blinding of care providers, participants and outcome assessors also appeared to be appropriate. All outcomes appeared to be reported. Although the studies used a modified intention to treat analysis, this included almost all trial participants. In both BLISS-SC and the pooled BLISS-52 and BLISS-76 trials, the placebo groups had worse baseline prognostic factors: more BILAG organ domain involvement (75.0% vs 69.8% at least 1A or 2B in BLISS-SC, 62.8% vs 59.0% at least 1A or 2B in pooled BLISS-52 and BLISS-76), more chance of having at least one flare (20.4% versus 16.5% for BLISS-SC, 24.7% vs 20.4% for pooled BLISS-52 and BLISS-76), and for the BLISS-SC trial only, a higher chance of having proteinuria (7.1% vs 3.4% ≥2 g/24h) and a higher chance of being on immunosuppressants (48.9% vs 43.9%). In the BLISS-SC and BLISS-52 trials, more patients withdrew in the placebo group (N=66/280 [23.6%] and N=61/287 [21.2%], respectively) than the belimumab group (N=93/556 [16.7%] and N=49/290 [16.9%], respectively) although these were included in the analysis as non-responders. Therefore, the ERG agrees that the three trials were well conducted but notes that there may be a risk of bias from the placebo groups having worse baseline prognostic factors and larger withdrawal rates (for BLISS-SC and BLISS-52) than the belimumab groups.

4.2.5 Efficacy results of the pivotal trials

The results of the primary endpoint analysis and its components, major secondary endpoints, and further key endpoints of interest for both BLISS-SC and pooled BLISS-52 and BLISS-76 trials are presented in Table 4.9 for the HDA-2 population. The efficacy of belimumab was greater than compared with placebo for the primary endpoint and most of its individual components (apart from no new 1A/2B BILAG domain scores in BLISS-SC). BLISS-SC also showed no significant difference between belimumab and placebo in the time to the first SFI flare and prednisone reduction of 25% or more compared to baseline. The results were similar to those obtained for the overall ITT analyses of the BLISS trials.

ERG comment: As reported in the original ERG report for TA397, there was a relative lack of evidence for clinical effectiveness of belimumab seen in the BLISS-76 trial. The results favourable for belimumab submitted for the pooled population across trials were largely driven by BLISS-52 results. The SLE population in BLISS-76 is more likely to resemble that in the UK than that in the BLISS-52. Therefore, the BLISS-76 results are probably more relevant to the decision problem than those from BLISS-52, and results from the pooled population may overestimate the effectiveness of belimumab in the UK population.

Table 4.9: SRI-4 responder rate and individual components at Week 52 in the HDA-2 population

	BLISS-SC		Pooled BLISS-52 and BLISS-76 data	
	Placebo (n=141)	**Belimumab 200mg SC (n=296)**	**Placebo (n=270)**	**Belimumab 10 mg/kg IV (n=262)**
SRI-4a (Primary endpoint)				
Response, N (%)	[redacted]	[redacted]	[redacted]	[redacted]
OR (95% CI) vs. placebo	[redacted]	[redacted]	[redacted]	[redacted]
4-point reduction in SELENA-SLEDAI[a] (Primary endpoint component)				
Response, N (%)	[redacted]	[redacted]	[redacted]	[redacted]
OR (95% CI) vs. placebo	[redacted]	[redacted]	[redacted]	[redacted]
No worsening in PGA[b] (Primary endpoint component)				
Response, N (%)	[redacted]	[redacted]	[redacted]	[redacted]
OR (95% CI) vs. placebo	[redacted]	[redacted]	[redacted]	[redacted]
No new 1A/2B BILAG domain scores[c] (Primary endpoint component)				
Response, N (%)	[redacted]	[redacted]	[redacted]	[redacted]
OR (95% CI) vs. placebo	[redacted]	[redacted]	[redacted]	[redacted]
SELENA SLEDAI change from baseline at Week 52				
Mean (SD or SE)	[redacted]	[redacted]	[redacted]	[redacted]
LS mean (SE)[d]	[redacted]	[redacted]	[redacted]	[redacted]
Difference (95% CI) vs. placebo[d]	[redacted]	[redacted]		[redacted]
Time to first SFI flare[e]				
Patients with flare over 52 weeks, N (%)[f]	[redacted]	[redacted]	[redacted]	[redacted]
Median days (IQR or range)[g]	[redacted]	[redacted]	[redacted]	[redacted]
HR (95% CI) vs. Placebo[h]	[redacted]	[redacted]	[redacted]	[redacted]
Time to first severe SFI flare[e] (Major secondary endpoint)				

	BLISS-SC		Pooled BLISS-52 and BLISS-76 data	
	Placebo (n=141)	Belimumab 200mg SC (n=296)	Placebo (n=270)	Belimumab 10 mg/kg IV (n=262)
Patients with severe flare over 52 weeks, N (%)f	[redacted]	[redacted]	[redacted]	[redacted]
Median days (IQR or range)g	[redacted]	[redacted]	[redacted]	[redacted]
HR (95% CI) vs. Placebo[h]	[redacted]	[redacted]	[redacted]	[redacted]
FACIT-Fatigue Scale Score Change from Baseline				
Mean (SD or SE)	[redacted]	[redacted]	[redacted]	[redacted]
LS mean (SE)i	[redacted]	[redacted]	[redacted]	[redacted]
Treatment difference (95% CI) vs. placebo[i]	[redacted]	[redacted]	[redacted]	[redacted]
Prednisone reduction by ≥25% from baseline to ≤7.5 mg/day during weeks 40–52 in patients with baseline prednisone dose >7.5 mg/day (Major secondary endpoint)				
Patients with prednisone reduction to ≤7.5 mg/day n/N (%), n[j]	[redacted]	[redacted]	[redacted]	[redacted]
OR (95% CI) vs. Placebo[k]	[redacted]	[redacted]	[redacted]	[redacted]
EQ-5D UK Score change from baseline				
Mean (SE)	EQ-5D data not collected		[redacted]	[redacted]
LS mean (SE)[l]			[redacted]	[redacted]
Treatment difference (95% CI) vs. placebo[l]			[redacted]	[redacted]
p-value[l]			[redacted]	[redacted]

Source: Table 39 in the CS.[1]

Notes: Randomisation stratification factors were baseline SELENA SLEDAI score, race (black vs. other) and baseline proteinuria level (<2 g/24 hour vs. ≥2 g/24 hour equivalent) in BLISS-SC; and baseline SELENA SLEDAI score, baseline proteinuria level, race (African descent or indigenous-American descent vs. other) and study (BLISS-52 vs BLISS-76) in the pooled dataset.

[a]OR (95% confidence interval) and p-value are from a logistic regression model adjusting for treatment group and randomisation stratification factors. The pooled BLISS data was also adjusted for study.

	BLISS-SC		Pooled BLISS-52 and BLISS-76 data	
	Placebo (n=141)	**Belimumab 200mg SC (n=296)**	**Placebo (n=270)**	**Belimumab 10 mg/kg IV (n=262)**

[b]Baseline PGA score is also included in the model.
[c]Baseline BILAG domain involvement (at least 1A/2B versus at most 1B) is also included in the model.
[d]All statistics are from an analysis of covariance (ANCOVA) model comparing treatments adjusting for randomisation stratification factors and study (in the pooled BLISS analysis).
[e]Severe flares that were triggered only by an increase in SELENA SLEDAI score are reported as mild/moderate flares if the change from the previous visit was at ≥3 points and are excluded otherwise. Time to first flare is defined as (event date – treatment start date + 1).
[f]Only includes post-baseline flares.
[g]Statistics will be missing when the number of events is too low to estimate the value.
[h]From Cox proportional hazards model for the comparison between treatments adjusting for randomisation stratification factors Study was also included in the pooled BLISS data.
[i] ANCOVA model comparing treatments adjusting for randomisation stratification factors and baseline FACIT-Fatigue score. Pooled BLISS data was also adjusted for study.
[j]Includes only subjects with baseline prednisone > 7.5 mg/day. All corticosteroids are converted to a prednisone equivalent average daily dose (mg/day).
[k]Logistic regression model comparing treatments adjusting for randomisation stratification factors.
[l]From ANCOVA for the comparison between treatments, adjusted for the corresponding baseline EQ-5D score

ANCOVA = analysis of covariance; BILAG = British Isles Lupus Assessment Group; CI = confidence interval; HDA = high disease activity; HR = hazard ratio; IQR = interquartile range; NA = not available; OR = odds ratio; PGA = Physician's Global Assessment; SC = subcutaneous; SD = standard deviation; SE = standard error; SELENA = Safety of Estrogen in Lupus Erythematosus National Assessment; SFI = SLE Flare Index; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index; SRI-4 = SLE responder index-4.

ERG comment: The results were mostly favourable for belimumab in both BLISS-SC and the pooled BLISS-52 and -76 data for the HDA-2 subgroup.

In comparison to the results in the HDA-1 subgroup results seem slightly less favourable for belimumab. The primary endpoint SRI-4 response at week 52 for the pooled BLISS-52 and BLISS-76 population was OR = 2.7 (95% CI: 1.8, 4.1) for the HDA-1 subgroup compared with OR = [redacted] (95% CI: [redacted]) in the HDA-2 subgroup. Primary endpoint components at 52 weeks, such as '4-point reduction in SELENA-SLEDA' (OR = 2.6 (95% CI: 1.7, 3.9) for HDA-1 and OR = [redacted] (95% CI: [redacted]) for HDA-2), 'No worsening in PGA' (OR = 2.0 (95% CI: 1.3, 3.1) for HDA-1 and OR = [redacted] (95% CI: [redacted]) for HDA-2), and 'No new 1A/2B BILAG domain scores' (OR = 1.9 (95% CI: 1.2, 3.0) for HDA-1 and OR = [redacted] (95% CI: [redacted]) for HDA-2) showed a similar trend.

4.2.6 SLICC(ACR)/SDI Indirect Cohort Study

The company performed a propensity score matching (PSM) analysis which matched patients treated with belimumab (plus ST) in the BLISS-76 US LTE study (primary analysis) with patients from an external SLE cohort treated with ST, to enable a long-term comparative analysis of belimumab versus ST.[68] Further details are provided in Sections 4.3 and 4.4.

4.2.7 Adverse events

The safety of belimumab in patients with SLE has been evaluated in three pre-registration placebo-controlled IV studies (Phase 2 LBSL01 study (see Appendix L of the CS), BLISS-52 (Table 4.10 below) and BLISS-76 (Table 4.10 below)), one placebo-controlled SC study (BLISS-SC (Table 4.10 below)), and one post-marketing, placebo-controlled IV study (BASE - Table 4.11). Overall, adverse reactions were reported in 87% of belimumab treated patients and 90% of placebo-treated patients. The most frequently reported adverse reactions (≥5% of patients with SLE treated with belimumab plus ST and at a rate ≥1% greater than placebo) were viral upper respiratory tract infections, bronchitis, and diarrhoea. The proportion of patients who discontinued treatment due to adverse reactions was 7% for belimumab-treated patients and 8% for placebo-treated patients. Adverse events were not reported specifically for the HDA-2 population in the CS. Therefore, the ERG asked the company to provide specific adverse events for the HDA-2 population (Clarification letter, Question A37).[47] The are reported for the three BLISS trials in Table 4.10.

The BASE study was a double-blind, placebo-controlled, randomised (1:1), Phase IV safety study to evaluate all-cause mortality and adverse events of special interest (AESI) in adults with SLE receiving belimumab IV 10 mg/kg versus placebo over 52 weeks.[72] Differences in rates of mortality and other pre-specified AESI (malignancies, serious infections, opportunistic infections and other infections of interest, serious depression, suicidality, and serious infusion/hypersensitivity reactions) on-treatment (first to last dose +28 days) were assessed.[72] A total of 4,003 patients received at least one dose of trial medication. Overall rates of on-treatment AESIs were similar between groups, except for serious depression and serious infusion/hypersensitivity reactions, which were more frequently reported in the belimumab IV group (See Table 4.11).[72]

On-treatment deaths were most frequently caused by infection (three [0.15%] placebo versus nine [0.45%] belimumab); on-study deaths occurred in 22 (1.10%) placebo and 13 (0.65%) belimumab patients (difference [95% CI]: –0.45 [–1.03, 0.13]). Most fatal infections were observed during the first 20 weeks of treatment with belimumab.

On-treatment serious suicidal ideation/behaviour and self-injury events were reported for five (0.25%) placebo and 15 (0.75%) belimumab patients (difference [95% CI]: 0.50 [0.06, 0.94]); on-study suicidal

ideation/behaviour occurred in 39 (1.96%) placebo and 48 (2.43%) belimumab patients (difference [95% CI]: 0.47 [–0.44, 1.38]). No suicide-related deaths were reported.[72]

ERG comment: The ERG noted information regarding adverse events in the HDA-2 population was not provided in the initial submission and further clarification was requested. The experience of at least one adverse event was common among all the trial participants. However, of those that experienced a severe adverse event, it was unclear what the severe adverse event was as this was not further defined.

Table 4.10: Treatment-emergent AEs in the pivotal trials of belimumab (HDA-2 subgroup)

	BLISS-SC		BLISS-76 data			BLISS-52		
	Belimumab 200mg SC (n=296)	**Placebo (n=141)**	**Belimumab 1 mg/kg IV (n=124)**	**Belimumab 10 mg/kg IV (n=115)**	**Placebo (n=127)**	**Belimumab 1 mg/kg IV (n=131)**	**Belimumab 10 mg/kg IV (n=147)**	**Placebo (n=143)**
At least 1 AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
At least 1 related AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
At least 1 serious AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
At least 1 severe[a] AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
At least 1 serious/ severe[a] AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
At least 1 AE resulting in study discontinuation	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Death	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: Response to clarification, Question A37.[47]
[a]) Severe or life threatening. Only treatment-emergent adverse events (AEs) are summarized.

Table 4.11: BASE study: Pre-specified AESI endpoints (Total trial population)

	Placebo n=2001	Belimumab 10 mg/kg IV n=2002	Difference (%) versus placebo (95% CI)
Deaths, N (%)	8 (0.40)	10 (0.50)	0.10 (–0.31, 0.51)
Serious infections, N (%)	82 (4.10)	75 (3.75)	–0.35 (–1.55, 0.85)
Opportunistic infections and other infections of interest, N (%)	50 (2.50)	36 (1.80)	–0.70 (–1.60, 0.20)
Malignancies (excluding NMSC), N (%)	5 (0.25)	5 (0.25)	0 (–0.31, 0.31)
NMSC, N (%)	3 (0.15)	4 (0.20)	0.05 (–0.21, 0.31)
Serious depression, N (%)	1 (0.05)	7 (0.35)	0.30 (0.02, 0.58)
Suicidality[a] (C-SSRS), N (%)	23 (1.16)	28 (1.42)	0.26 (–0.44, 0.96)
Serious infusion, hypersensitivity reactions, N (%)	2 (0.10)	8 (0.40)	0.30 (–0.01, 0.61)

Source: CS, Table 50, page 109.[1]

[a]) Treatment-emergent suicidal ideation/behaviour.

AESI = adverse events of special interest; CI = confidence interval; C-SSRS = Colombia Suicide Severity Rating Scale; NMSC = non-melanoma skin cancer.

4.2.8 Included studies: Supporting evidence

Fourteen studies providing supporting evidence were highlighted in the CS – see Table 4.12. Only one of these was used to inform the economic model.[66] This was US data from the long-term extension of the BLISS-76 study. There were two further extension studies (BLISS-52/76 non-US data)[65] and (BLISS-SC LTE).[67]

The remainder of the supporting studies included a phase 2 belimumab trial[70] and its extension.[71] There were trials in specific populations: the PLUTO trial in children;[76] EMBRACE, a placebo-controlled trial of belimumab in people of black race;[74] and NCT01345253, a placebo-controlled trial of belimumab in people from North-East Asia.[75] Also included were studies of belimumab with a specific purpose: BASE, a safety study of mortality and adverse events of special interest;[72] a study to evaluate the reliability of the SC autoinjector[69] and NCT02119156, a treatment holiday study.[73] Two registries were included to provide real world evidence: BILAG-BR[77, 78] and OBSErve.[79, 80] Finally, there was an indirect comparison of the efficacy of SC and IV belimumab formulations.[81] All of these will be briefly mentioned in this section with more emphasis on those that add specific points to the main evidence.

Table 4.12: Overview of the supporting studies in the CS

Study name	Description	Population	Location in the current submission
BLISS-76 US LTE[66]	LTE study of US patients enrolled in BLISS-76	Total*	Document B Section 2.6 and Appendix M
BLISS-52/76 non-US LTE[65]	LTE study of non-US patients enrolled in BLISS-52 or BLISS-76	Total	Document B Section 2.6 with further details in Appendix M
BLISS-SC LTE[67]	Open-label extension for patients enrolled in BLISS-SC	Total	Document B Section 2.6 and Appendix M
Phase 2B open-label, single-arm, repeat-dose study[69]	Study to evaluate the reliability of the SC autoinjector	Total	Appendix O
LBSL02 Phase 2 trial[70]	Initial evidence on safety and efficacy of belimumab	Total	Appendix L
LBSL02 LTE[71]	Data on long-term (up to 13 years) with belimumab	Total	Appendix M
BASE (post-marketing)[72]	Safety study of mortality and adverse events of special interest	Total	Document B Section 2.6, 2.10 and Appendix F
Treatment holiday study (NCT02119156)[73]	A study of the effect of treatment holiday on belimumab efficacy	Total	Appendix O
EMBRACE (post-marketing)[74]	Placebo-controlled trial of belimumab in people of black race	Total	Appendix O
NCT01345253[75]	Placebo-controlled trial of belimumab in people from North-East Asia	Total	Appendix O
PLUTO[76]	Belimumab in children and adolescents	Total	Appendix O
BILAG-BR[77, 78]	UK-based registry of biologic therapy (including belimumab) for SLE	HDA-1 (bel data only)	Document B Section 2.7 and Appendix P
OBSErve[79, 80]	A multi-country Evaluation Of use of Belimumab in clinical practice Settings	Total	Document B Section 2.6
ITC between SC and IV belimumab formulations[81]	To compare efficacy of SC and IV belimumab formulations in patients with autoantibody-positive SLE with HDA	Total	Appendix O

Source: Table 12 of the CS[1] and Appendix D[57]
* Used in the economic model

Long-term extension studies

BLISS-76 US LTE was a multicentre continuation trial of belimumab in SLE patients who completed the Phase 3 BLISS-76 trial in the US.[82] In this trial, patients received belimumab IV every 28 days in the same dose as in the BLISS-76 trial, 1 mg/kg or 10 mg/kg.[82] If patients were originally assigned to the placebo group, they would then receive 10 mg/kg.[82] In time, patients who received 1 mg/kg of belimumab, then increased to 10 mg/kg. [82] The primary efficacy endpoint of the BLISS-76 US LTE trial was the SRI-4 response rate at each belimumab visit.[82] Additional efficacy endpoints of the trial included SELENA-SLEDAI, BILAG, PGA, SFI, and prednisone use.[82] In the placebo-to-belimumab group at Year 1, 10/60 were noted as SRI-4 responders, while 13/65 participants noted a greater than four-point reduction in SELENA-SLEDAI scores, and 67/84 noted no worsening in PGA.[1] When compared to Year 7 in this group, the SRI-4 responder rate was noted in 6/7 participants, while 7/7 participants noted a greater than four-point reduction in SELENA-SLEDAI, and 11/12 participants noted no worsening in PGA.[1] In the belimumab group at Year 1, SRI-4 responder was noted in 86/169 participants, while 91/169 participants observed a four-point reduction in SELENA-SLEDAI scores, and 166/177 participants noted no worsening in PGA.[1] When compared at Year 7, the SRI-4 responder rates were noted in 84/112 participants, while the four-point reduction in SELENA-SLEDAI scores were identified in 86/112 participants, and 108/115 participants noted no worsening in PGA.[1]

BLISS-52/76 non-US data was a multicentre, continuation trial of belimumab in SLE patients who completed the phase III BLISS-52 or BLISS-76 trials.[82] Patients received belimumab IV every 28 days in the same dose as in the BLISS-52 or BLISS-76 trials, 1 mg/kg or 10 mg/kg.[82] If patients were originally assigned to the placebo group, they would then receive 10 mg/kg.[82] In time, patients who received 1 mg/kg of belimumab, then increased to 10 mg/kg.[82] The efficacy endpoint in this trial was SDI.[82] In the placebo-to-belimumab group at Year 1 SDI worsening was observed in 11/220 participants, whereas at Year 8 this was observed in 0/5 participants.[1] In the belimumab group SDI worsening was identified in 28/496 participants, whereas at Year 8 this was observed in 8/60 participants.[1]

BLISS-SC LTE was a 6-month open-label extension phase of the BLISS-SC.[82] During this time patients received belimumab 200 mg SC on a weekly basis for six months.[82] If patients were initially in the placebo group, they were then switched to the belimumab SC group.[82] Patients who received belimumab 200 mg SC initially, continued to do so.[82] Efficacy assessments included SRI-4 responses.[82] In the placebo-to-belimumab 200 mg SC group results at week 24, SRI-4 responders were noted in 23/143 participants, whereas in the belimumab group, this was observed in 332/435 participants.[1]

Phase 2 trial and its extension

The purpose of the LBSL02 phase 2 study was to assess the safety, tolerability, biologic activity, and efficacy of belimumab in combination ST in patients with SLE.[83] Patients were randomly assigned to receive belimumab in 1, 4, or 10 mg/kg or placebo over the course of the 52-week study.[83] The co-primary endpoints of the study were the percent change in the SELENA-SLEDAI score at week 24 and the time to first SLE flare.[83] At Week 52, belimumab treatment resulted in significantly better responses than placebo for SELENA-SLEDAI score (–28.8% versus –14.2%; p=0.0435).[83] The combined belimumab group noted the median time to first SLE flare to be 67 days compared to the placebo group at 83 days.[83] However, during weeks 24-52, the median time to first SLE flare was significantly longer with belimumab treatment (154 versus 108 days; p=0.0361).[83]

LBSL02 LTE was the 24-week extension study of the LBSL02 phase 2 study.[82] In the extension study, patients who were in the placebo group switched to 10 mg/kg IV belimumab, while those who had

received belimumab originally continued to do so at the same dose.[82] Patients received belimumab for up to 13 years.[82] 32.8% of patients achieved an SRI-4 response in Year 1 and this increased to 75.6% by Year 12.[82] The percentage of patients who achieved a SELENA-SLEDAI score of ≤2 and a prednisone dose of ≤5 mg/day increased from 13.9% at Year 1 to 57.1% at Year 13.[82]

Trials in specific populations

PLUTO is a paediatric trial of IV belimumab compared with placebo (both in addition to ST),[76] with an ongoing long-term follow-up phase. PLUTO consists of three phases identified as Part A, Part B, and Part C.[84] Part A included a randomised, placebo-controlled, double-blind 52-week treatment phase.[84] Part B, which is ongoing, is a long-term belimumab open-label safety follow-up for any patient who completes Part A.[84] Part C, which is ongoing, is a long-term follow-up phase for patients who withdrew from Part A or Part B at any time.[84] Patients either received 10 mg/kg of belimumab or a placebo.[84] The primary endpoint was SRI-4 responses at Week 52.[84] By Week 52, the percentage of responders for the belimumab group was 52.8% compared with the 43.6% for the placebo group.[84] Results for Part B and Part C are not yet available.[84]

EMBRACE was a multicentre, double-blind, placebo-controlled post-marketing commitment study in patients of self-identified black race adults with active SLE.[84] Patients were randomised to either receive belimumab 10 mg/kg IV or placebo and standard therapy.[84] The primary endpoint was the difference in modified SRI-4-SLEDAI-2K response rate between placebo and belimumab.[84] At Week 52, greater proportions of belimumab patients were SRI-4-SLEDAI-2K responders than placebo.

NCT01345253[75] was a multicentre, randomised, double-blind, placebo-controlled study conducted in North East Asia.[84] Patients were randomised to receive either belimumab IV 10 mg/kg or placebo and standard therapy every four weeks until Week 48.[84] The primary endpoint of the study was the SRI-4 response rate measured at Week 52.[84] There were noted significant improvements in the belimumab group when compared to the placebo group (53.8% [240/446] vs 40.1% [87/217]; OR 1.99 [95% CI: 1.40, 2.82]; p<0.0001).[84]

Trials with a specific purpose

The BASE study was a post-marketing commitment, global, multicentre, randomised, double-blind, placebo-controlled, 52-week study used to assess the mortality and adverse events of special interest in adults with active, antibody-positive SLE treated with belimumab and standard therapy versus placebo and standard therapy.[85] The studied dose of belimumab was 10 mg/kg.[85] Eight deaths were observed in the placebo group, while 10 deaths were observed in the belimumab group.[85] Both groups reported five malignancies.[85] The placebo group identified 50 opportunistic infections, while the belimumab group identified 36 opportunistic infections.[85]

The Phase 2B study was an open-label, single-arm, multi-dose study conducted to assess the suitability of the autoinjector for self-administration of belimumab by patients with SLE.[84] Patients were required to be on an SLE treatment regimen including IV belimumab every 28 days for at least three 28-day cycles or have completed the BLISS-SC open-label LTE study.[84] The study's primary efficacy endpoint was the proportion of patients successfully able to self-administer their first and second doses.[84] Out of 736 patient-attempted injections, 720 injections were determined to be successful. Reasons for unsuccessful injections were identified as use error, device error, or both.[84]

The treatment holiday study (NCT02119156) was a post-marketing commitment study which investigated the effects of belimumab treatment holiday and reintroduction, and treatment discontinuation in patients with stable low disease activity.[1, 84] Adults with SLE received belimumab at

a dose of 10 mg/kg IV for ≥6 months were recruited to three arms including the treatment holiday, continuous belimumab, and long-term discontinuation.[84] The primary endpoint was the time to first SLE flare.[84] The flare rate was highest among participants in the long-term discontinuation arm, 2.1, compared to the 1.0 of the treatment holiday group, and 0.6 in the treatment control arm.[84]

Registries - BILAG-BR and OBSErve

The BILAG-BR included a prospective, multicentre, non-randomised, observational registry study of patients prescribed belimumab in clinical practice.[1] The purpose of the cohort study is to investigate the safety of biologics in SLE treatment. The study focused on three cohorts, which comprised of patients who received, belimumab, rituximab, and non-biologics.[1] Among belimumab patients disease activity was assessed using BILAG-2004, SLEDAI-2K, and the SDI.[1] [redacted]

The OBSErve registry is a series of non-randomised, single-arm retrospective, observational studies conducted in Argentina, Canada, Germany, Spain, Switzerland, and the US.[1] The primary endpoint was physician-assessed overall clinical response to belimumab at six months.[1] Of the 830 patients in the pooled dataset, 10 patients were identified as having a worse physician-assessed overall clinical response, 29 experienced no improvement, 104 patients experienced <20%, 288 patients experienced 20-49%, 292 patients experienced 50-79%, and 107 experienced ≥80% physician-assessed overall clinical response.[1]

Indirect comparison of SC and IV belimumab formulations

An indirect treatment comparison was performed in order to compare the efficacy of SC and IV belimumab formulations in patients with high SLE disease activity.[84] The 10 mg/kg IV formulations of belimumab were based on the BLISS-52, BLISS-76, and North-East Asia studies, while the 200 mg of SC belimumab were derived from the BLISS-SC trial.[84] Patients were included if they met either Criteria I, being low complement (C3 or C4) and anti-dsDNA positive, or Criteria II, being low complement (C3 or C4) or a SELENA-SLEDAI score equalling 10 or above.[84] A network meta-analysis was performed on SRI-4 response rates at Week 52.[84] SRI-4 response rates among belimumab IV patients in Criteria I were 313/596, while belimumab SC patients in Criteria I were 159/246.[1] In Criteria I, placebo was 188/530 were observed to be SRI-4 responders.[1] In Criteria II, 398/738 belimumab IV patients were observed to be SRI-4 responders, while 269/421 belimumab SC patients were observed to be responders.[1] 282/731 patients in the placebo group in Criteria II, were SRI-4 responders.[1]

4.2.9 Ongoing studies

The CS noted five ongoing studies intended to investigate the use of belimumab in patients with SLE. The details regarding ongoing studies are provided in Table 4.13. The company indicated that the BLISS-LN study, which evaluated the safety and efficacy of IV belimumab in patients with active lupus nephritis, has published results in the New England Journal of Medicine.[51] However, BLISS-LN was stated to be outside the scope of the current appraisal.

Table 4.13: Ongoing studies

Study Name	Study Type	Study Description	Belimumab formulation	Estimated study completion	Report expected
BLISS-BELIEVE	Phase IIIA Interventional Clinical Trial	Phase 3, 104- week, safety and efficacy study of belimumab-rituximab combination in patients with SLE.	SC	July 2021	[redacted]
SABLE	Observational registry establishing as a post-marketing commitment	Multicentre, prospective, observational cohort study to evaluate the incidence of AESI and effectiveness in SLE patients	Either	Jan 2025	[redacted]
114256	Pregnancy registry-post-marketing commitment	A registry to investigate the safety of belimumab in pregnancy. Due to the very slow recruitment, GSK is currently in discussions with the EMA around alternative relevant studies that could be conducted.	Either	Nov 2021	[redacted]
116559	Meta-analysis of the elderly SLE patients-post-marketing commitment	Meta-analysis conducted under study ID BEL116559 to assess belimumab efficacy and safety in elderly patients treated in selected belimumab studies. This is a post-marketing commitment with the EMA.	Either	Dec 2025	[redacted]
BASE	Phase IV Interventional Clinical Trial	Global, multicentre, placebo-controlled RCT to evaluate AESI in SLE patients treated with belimumab. Primary analysis of this study is now complete.	IV	Aug 2022	[redacted]

Source: CS, Table 53, page 113.
AESI = adverse events of special interest; EMA = European Medicines Agency; IV = intravenous; NA = not applicable; PD = pharmacodynamics; PK = pharmacokinetics; RCT = randomised controlled trial; SC = subcutaneous; SLE = systemic lupus erythematosus

4.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

As discussed in Section 3.3 of this report, the company did not include rituximab or cyclophosphamide as comparators. Therefore, no indirect treatment comparisons (ITC) with either rituximab or cyclophosphamide were performed.

4.3.1 Comparison to standard therapy

The company did perform a propensity score matching (PSM) analysis to compare belimumab with standard therapy (ST) although this was not reported as using the population adjusted indirect comparison methods recommended by NICE.

An SLR was performed to identify cohorts, registries or other databases that supported SLE. The objective was to identify a comparison cohort with population characteristics similar to the BLISS trial population with adequate sample size with complete clinical data and at least five years follow-up. In total 92 cohorts were identified of which 21 cohorts had at least 400 patients and from which data was extracted. Evaluation criteria included cohort size, ethnicity, age, duration of SLE, severity of disease activity, extent of organ damage, follow-up and scope of data collection and data availability. The Toronto Lupus Cohort (TLC)[19, 29] was selected as the preferred source of ST data for this post-hoc longitudinal PSM study, based on the size of the cohort, the extent of organ damage among the patients and the severity of their disease activity within the cohort.[68] Moreover, the scales for disease activity, organ damage progression and health-related quality of life were compatible with those from the BLISS studies.

The company used a SLR to identify publications that reported predictors of SLE organ damage and progression. Key predictors found in the literature were reviewed by a clinical expert and limited to those available in both the BLISS LTE studies and the TLC. This generated a list of 14 predictors which were used in the primary PSM analysis of the BLISS US LTE/TLC datasets.[68] The patient eligibility criteria from the BLISS trials were applied to the TLC cohort to select patients with similar characteristics.

The primary objective was to compare organ damage progression (SDI score) in patients treated with belimumab (plus ST) or ST alone, using PSM data from the BLISS-76 US LTE study and the TLC external cohort. Secondary objectives included comparing the time to organ damage progression and the magnitude of damage accrual. The time to organ damage progression analysis included all patients with >1 year of follow-up and excluded TLC patients with ≥15 years of follow-up.[68]

The primary endpoint was the difference in change in Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index (SDI) score from baseline to five years. A total of 259 patients from the BLISS-76 US LTE study[66] and 706 patients from the Toronto Lupus Cohort (TLC)[19, 29] were included in the PSM analysis. The PSM used a logistic regression model which started with all potential predictor variables (full model) and then used step-wise backwards elimination removing the least significant variables at each step until all the remaining predictors were significant at the 10% level (trimmed model). Patients from the BLISS US LTE study were matched to TLC patients on a 1:1 basis using a calliper of 20% (based on 20% of the SD for the distribution in the full sample) and 99 patients from each study were matched on a 1:1 basis.[68]

Over a five-year period, patients treated with belimumab experienced significantly less organ damage than patients treated with ST alone (Table 4.14). Further, patients receiving belimumab were 61% less likely to progress to a higher SDI score over any given year of follow-up compared with patients treated with ST (HR 0.391; 95% CI 0.253 to 0.605; p<0.001). A patient receiving belimumab had a 3.5% annual probability of organ damage progression compared with an 8.7% annual probability of progression with ST alone.[68]

When the magnitude of year-to-year organ damage progression was explored, it was found that of those patients treated with belimumab there were 33 instances of an SDI score increase of ≥1 compared with 72 instances in patients treated with ST.

A higher proportion of patients treated with ST experienced an SDI score increase ≥2 compared with patients treated with belimumab (p=0.006).[68]

Table 4.14: PSM analysis: Change in SDI from baseline to 5-years

	ST, N=99	Belimumab, N=99	Difference
5-year SDI change, mean (SE)	0.717	0.283	–0.434 (0.119)
95% CI	0.500 to 0.934	0.166 to 0.400	–0.667 to –0.201

Source: CS, Table 31, page 79.[1]
CI = confidence interval; PSM = propensity score matching; SD = standard deviation; SDI = Systemic Lupus International Collaborating Clinics (SLICC)/ACR Damage Index; ST = standard therapy.

4.3.2 Comparison to rituximab

The ERG asked the company to justify why they did not perform an indirect treatment comparison (ITC) between belimumab and rituximab. In the CS the company stated that there was no direct RCT evidence and that there were differences in the trial endpoints and patient populations between the EXPLORER trial of rituximab and the belimumab trials indicating that the rituximab patient population had more severe disease and SELENA-SLEDAI was not measured in the trial. The ERG asked the company to consider performing an ITC using the BLISS and EXPLORER trials (Clarification Letter, Question A42).[47] The mean disease duration was longer in the rituximab trial by approximately two years and the proportions of patients with BILAG of 1A or above was higher (53% vs. 14.8% and 16.9%). They also stated that the endpoints were too different to include in an ITC, in the BLISS trials the primary endpoint was SRI-4 response at week 52, a composite endpoint including reduction in SELENA-SLEDAI score, BILAG domains and no worsening in PGA from baseline. The primary endpoint in EXPLORER was achieving and maintaining a clinical response at week 52 assessed using each of the eight BILAG index organ system scores. Although BILAG was used in both trials it was used differently, as a worsening or an improvement endpoint. There were also differences in the use of steroids between the trials.

Additional data comparing belimumab and rituximab were available from the BILAG BR sub-study (CS, Appendix P). This was an analysis of the BILAG Biologic Register, an observational prospective cohort study of patients receiving hospital treatment for SLE in the UK. The eligibility criteria were defined by the NICE recommended subgroup of the current licensed population for belimumab and patients were included from October 2013 onwards who had high disease activity (anti-dsDNA positive, low Complement 3 or 4 level and SLEDAI-2K score ≥ 10). The analysis was performed by the University of Manchester. The analysis included [redacted] patients receiving belimumab and [redacted] receiving rituximab. The two primary endpoints for SLE were BILAG-2004 and SLEDAI-2K, a disease specific instrument (SLICC/ACR damage index) was also used and HRQoL was measured using both generic and disease-specific instruments. Patients were recruited from 39 centres across the UK and followed up at three, six and 12 months then every 12 months. At any point they could stop treatment and restart with a second round of the same treatment or restart with a different one. The analysis used multilevel repeated measures regression modelling of outcomes at three, six and 12 months. Potential confounding variables were identified before analysis and included as covariates in the models.

The results showed that there was [redacted] between belimumab and rituximab in SLEDAI-2K at 12 months of follow-up in the model adjusted for all covariates (mean difference [MD] [redacted], 95% CI [redacted]), [redacted] in BILAG total score (MD [redacted], 95% CI [redacted]).

4.4 Critique of the indirect comparison and/or multiple treatment comparison

4.4.1 Comparison to standard therapy

There are two major issues with the PSM analysis that means its use in calibrating the cost effectiveness model can severely bias the results in favour of belimumab.

The first issue is that any PSM analysis can only match on observed and measured variables, any important variables which were not measured in one or both studies cannot be included in the analysis. While the variables in the PSM model are reasonably free of bias for the 99 matched patients in each cohort, there are several important confounders that are not matched: of the predictors of SLE organ damage and progression the PSM study found in a literature review, household income, educational attainment, history of anti-phospholipid and anti-Ro autoantibodies, disease activity over time and physical functioning were not included in the matching.

The two cohorts are drawn from fundamentally different populations, as seen in the observed differences between the cohorts before matching, for example smoking, where 3.6% of BLISS US LTE patients smoked, compared with 23.7% of TLC patients. This indicates different levels of unobserved variables between the cohorts, particularly for deprivation and comorbidities, which are almost completely unaccounted for by matching. Given the observed variables (especially smoking, though also dislipidaemia and proteinuria rates), it is likely the TLC patients would have worse outcomes than the BLISS LTE patients even after matching (due to confounding from unmatched variables), making belimumab seem more effective than it is. This means that while the analysis is less biased than a non-PSM analysis, it is likely still heavily biased by confounding, and therefore it is not appropriate to treat the results of this analysis as if they had been estimated using an RCT, which is what calibrating a cost effectiveness model with these results does. The PSM analysis using the pooled BLISS LTE cohorts[86] is more confounded than the analysis using the US-only BLISS LTE, as smoking was not included as a matching variable due to the large differences in smoking rates between the cohorts.

The second issue is likely more serious. The withdrawal rate of patients in the BLISS LTE studies[87] was high: of the 1,749 patients who started in BLISS trials (1,184 any belimumab dose, 565 placebo), 1,333 patients [76%] completed the trial (921 belimumab [78%], 412 placebo [73%]), and 738 patients [42%] enrolled in BLISS-LTE studies (504 belimumab [43%], 234 placebo [41%]). The PSM analysis using the pooled BLISS studies[86] states that 592 patients had five years of follow-up [34% of all patients in BLISS trials. This means that at least 66% of patients originally starting the BLISS trials were not included in the PSM analysis due to not entering the BLISS LTE studies (595 patients [34%], 417 [70%] on belimumab) or withdrawing from them (416 patients [24%], and at least 146 from BLISS LTE studies before five years of follow-up [8%]).

Reasons for withdrawal from the BLISS-76 and BLISS-52 trials[88, 89] for belimumab patients included: 46 patient requests (4.1% of all belimumab patients); 72 adverse events (6.4%); 53 withdrawals due to lack of efficacy (4.7%); 6 withdrawals due to lack of compliance (0.5%); 21 patients lost to follow-up (1.9%); 17 protocol violations (1.5%); 12 investigator decisions (1.1%); and 24 other reasons (2.1%). Reasons for withdrawal from the LTE studies (at any time), included: 182 patient withdrawals (18.1% of all BLISS LTE patients); 94 adverse events (9.3%); 91 other reasons, usually withdrawal of consent (9.1%); 53 physician decisions (5.3%); 16 withdrawals due to lack of compliance (1.6%); 34 patients lost to follow-up (3.4%); 20 withdrawals due to lack of efficacy (2.0%); and five protocol deviations (0.5%).

Patients with five years of follow-up of belimumab are therefore more likely to either have slowly progressing SLE, had a favourable response to belimumab with no or limited adverse effects, or a combination, compared with patients who either withdrew before five years had elapsed on belimumab, or chose not to enter the BLISS LTE studies. The CS states in table 64 that "*patients who had not demonstrated a sufficient response with belimumab during the Phase 3 studies would unlikely have continued into the extension study*",[1] emphasising this point.

Because a maximum of 34% of the original BLISS trial patients were included in the PSM, the size of this bias could be high. As TLC participants are drawn from an SLE clinic and therefore are unlikely to be highly selected, it is likely that even if belimumab had no effect, there would be a large difference in the five-year outcomes solely because the BLISS LTE patients have been highly selected for either low disease activity or high belimumab response. A further issue regarding the generalisability of these results to the UK population is that they were based on analyses of patients from the US and Canada and not the UK.

As both major biases likely make belimumab seem more effective than it is, the calibration factor derived from fitting the cost effectiveness model to the PSM analysis likely biases the model to make belimumab seem more cost effective than it is. This bias could potentially be large, due both to the large differences between the BLISS LTEs and TLC, and the large number of BLISS trial patients who never reached five years on belimumab.

4.4.2 Comparison to rituximab

Although the company stated that an ITC between belimumab and rituximab was not possible using trial data, there was comparative data available in the BILAG BR sub study. This was observational data collected from patients in hospitals in the UK and included patients with high disease activity corresponding to HDA-1. Although the analysis was based on small numbers and used observational data this does provide a comparison in a UK population relevant to the NICE decision problem. The outcome measures were different to the SRI-4 measure used by the BLISS trials but it did include BILAG scores and SLEDAI-2K (which were also measured in BLISS). The CS stated that there is a high likelihood of confounding and selection bias so that these data are not appropriate for comparing treatment efficacy.

4.5 Additional work on clinical effectiveness undertaken by the ERG

No further additional work on clinical effectiveness was undertaken by the ERG.

4.6 Conclusions of the clinical effectiveness section

The population defined in the scope is: "People aged 5 years or more with active, autoantibody-positive systemic lupus erythematosus with a high degree of disease activity despite standard therapy".[49] The scope does not provide a definition for 'a high degree of disease activity'. In the CS, the company provides two definitions:

- High Disease Activity Subgroup-1 (HDA-1): Patients with a SELENA SLEDAI score ≥10 AND low complement AND positive anti-dsDNA (current NICE guidance population; TA397)
- High Disease Activity Subgroup-2 (HDA-2): patients with a SELENA-SLEDAI score ≥10 AND at least one of the following serological features: low complement OR positive anti-dsDNA – The Base case

The current appraisal is different from the original appraisal (TA397[44]) in three ways:

1. The definition of 'high disease activity' (i.e. HDA-1 versus HDA-2, see above)

2. Age: in TA397, belimumab was approved for adults only. This appraisal includes people aged five years or more.
3. Formulation: The original appraisal included an intravenous (IV) formulation only. The current appraisal also includes a new subcutaneous (SC) formulation in the form of a pre-filled pen.

In the response to clarification (Clarification letter, Question A11),[47] the company acknowledged that patients with severe active CNS lupus were excluded from the BLISS trials and no evidence to support the use of belimumab is available in this population. Patients with lupus nephritis (LN) were also excluded from the BLISS trials. In addition, the company confirmed that 'no searches were performed in people over the age of five as the CS focuses on an adult population with SLE as does the economic modelling. The majority of clinical effectiveness data available on belimumab is for adult patients aged 18 years and older with SLE.' (Clarification letter, Question A17).[47]

This submission refers to the previously appraised IV formulation and introduces a new subcutaneous (SC) formulation in the form of a pre-filled pen.

The description of the comparators in the NICE scope is as follows: "Standard therapy alone. For people in whom it is considered appropriate: rituximab plus standard therapy or cyclophosphamide plus standard therapy".[49]

The company only included one comparator: Standard therapy alone. The company provides several reasons why rituximab was not included as a comparator: lack of data to perform an (indirect) comparison, and the recently published NHS England commissioning policy for rituximab in the treatment of SLE, which states that belimumab should be considered prior to rituximab in the treatment pathway.[1] Cyclophosphamide was not included as a comparator because "it is largely reserved for the treatment of lupus nephritis or CNS lupus, both of which are outside of the current marketing authorisation for belimumab".[1]

In the original appraisal (TA397), the Committee considered that, "because rituximab is provided through individual funding requests and its use in the NHS is likely to be limited, it should not be considered to be the main comparator in routine practice (although it had been specified in the scope for the appraisal). The Committee therefore concluded that standard care should be the main comparator for belimumab, as included in the final scope and in the manufacturer's decision problem" (FAD committee papers, page 29).[52] In addition, "The Committee was aware that cyclophosphamide was also included as a comparator in the scope for the appraisal, but noted the manufacturer's justification that it was largely used for lupus nephritis, which was a different population to the one included in the trials of belimumab and covered by the marketing authorisation for belimumab. Furthermore, it heard from clinical specialists that cyclophosphamide is used infrequently because of side effects" (FAD committee papers, page 29).[52]

We asked our clinical expert whether she thought it was reasonable not to include rituximab and cyclophosphamide as comparators. She responded that the company had a point in that cyclophosphamide is mostly used for neuropsychiatric lupus, severe vasculitis e.g. gut, severe cardiorespiratory involvement, and patients that fail mycophenolate mofetil (MMF) for some of these types of conditions and for renal lupus. In addition, the adverse event profile means that cyclophosphamide is avoided if possible and rituximab is used increasingly for severe manifestations. The comparison with rituximab will be difficult according to our clinical expert because the evidence for rituximab is weaker as the phase III trials were negative due to very stringent end-points (and different to those used for belimumab) and is mostly from registries. BILAG BR data cannot be used to

compare them easily due to the different criteria for the use of rituximab and belimumab (See BSR guidelines[39]).

The CS presented three pivotal RCTs of belimumab (BLISS-52,[62] BLISS-76[63] and BLISS-SC[64]). Each of the trials had an extension study.[65-67] Of this evidence, only the three RCTs and one of the extension studies[66] were included in the economic model. All three of the RCTs provided evidence for the two HDA subgroups presented in the CS.

The main evidence for the clinical effectiveness of belimumab was from two phase III clinical trials. The BLISS-52 (n=865) and BLISS-76 (n=819) trials were randomised, double-blind, placebo-controlled, parallel-group studies with follow-up at 52 weeks and 76 weeks respectively. In these trials, belimumab plus standard care was compared with placebo plus standard care.

The BLISS-SC trial was presented in this submission to introduce the SC formulation of belimumab. The BLISS-SC trial is an international multicentre phase III randomised placebo-controlled trial lasting 52 weeks. Patients were randomised to belimumab 200 mg SC once weekly plus standard treatment (ST) or placebo plus ST. All three trials used the same primary efficacy endpoint, which was the SRI-4 response rate at Week 52.

The results from the main trials (BLISS-SC and the pooled BLISS-52 and -76 data) were mostly favourable for belimumab in the HDA-2 subgroup. In comparison to the results in the HDA-1 subgroup results seem slightly less favourable for belimumab. The primary endpoint SRI-4 response at week 52 for the pooled BLISS-52 and BLISS-76 population was OR = 2.7 (95% CI: 1.8, 4.1) for the HDA-1 subgroup compared with OR = [redacted] (95% CI: [redacted]) in the HDA-2 subgroup. Primary endpoint components at 52 weeks, such as '4-point reduction in SELENA-SLEDA' (OR = 2.6 (95% CI: 1.7, 3.9) for HDA-1 and OR = [redacted] (95% CI: [redacted]) for HDA-2), 'No worsening in PGA' (OR = 2.0 (95% CI: 1.3, 3.1) for HDA-1 and OR = [redacted] (95% CI: [redacted]) for HDA-2), and 'No new 1A/2B BILAG domain scores' (OR = 1.9 (95% CI: 1.2, 3.0) for HDA-1 and OR = [redacted] (95% CI: [redacted]) for HDA-2) showed a similar trend.

The same critique as in the original appraisal stands: The SLE population in BLISS-76 is more likely to resemble that in the UK than that in the BLISS-52; therefore the BLISS-76 results are probably more relevant to the decision problem than those from BLISS-52. Patients were required to be ≥ 18 years of age in both the BLISS-52 and BLISS-76 trials which therefore did not include any paediatric patients. There were no UK patients enrolled in BLISS-52. In BLISS-76, a total of 11 patients from the UK were enrolled, constituting 1.3% of the total trial population. Of those, six patients were randomised to placebo, four to the unlicensed 1 mg/kg belimumab dose and one to the licensed 10 mg/kg dose.

Based on the total population included in the BLISS trials (that is, not just the target population) over 90% of patients in each arm experienced one or more adverse events. The most frequent (occurring in more than 10% of patients) events were headache, upper respiratory tract infection, arthralgia, nausea, urinary tract infection, diarrhoea and fatigue. Of these events, only diarrhoea and nausea occurred slightly more frequently in the belimumab groups than in the standard care groups. Serious adverse events were experienced by 17% in the 10 mg/kg belimumab group, compared with 16% in the standard care group. Across the double-blind treatment periods, 14 people died, including three (0.4%) in the standard care group, five (0.7%) in the 1 mg/kg group and six (0.9%) in the 10 mg/kg belimumab group. Four deaths were infection-related: one in the standard care group, one in the 1 mg/kg belimumab group and two in the 10 mg/kg belimumab group. Infection may have contributed to the deaths of two further patients (one in the 1 mg/kg belimumab group and one in the 10 mg/kg belimumab group). There were

two suicides, both in patients receiving belimumab (one in the 1 mg/kg group and one in the 10 mg/kg group), and one cancer-related death in a patient receiving 1 mg/kg belimumab.

The company performed a propensity score matching (PSM) analysis which matched patients treated with belimumab (plus ST) in the BLISS-76 US LTE study (primary analysis) with patients from an external SLE cohort treated with ST, to enable a long-term comparative analysis of belimumab versus ST.[68] Using the PSM analysis in calibrating the cost effectiveness model can severely bias the results in favour of belimumab. The largest issue is that the withdrawal rate of patients in the BLISS LTE studies was high. Patients with five years of follow-up of belimumab are more likely to either have slowly progressing SLE, had a favourable response to belimumab with no or limited adverse effects, or a combination, compared with patients who either withdrew before five years had elapsed on belimumab, or chose not to enter the BLISS LTE studies. Because a maximum of 34% of the original BLISS trial patients were included in the PSM, the size of this bias from using the calibration factor could be high, and almost certainly biases the model to make belimumab seem more cost-effective than it is. Additionally, by matching between BLISS non-US LTE and the Toronto Lupus Cohort (TLC), the data from the belimumab arm of the PSM analysis is not generalisable to the BLISS non-US LTE and TLC cohorts, but only to the subgroup of patients who match between the cohorts; it is therefore unknown how well the data will generalise to a UK cohort.

Additional data comparing belimumab and rituximab was available from the BILAG BR sub-study (CS, Appendix P). This was an analysis of the BILAG Biologic Register, an observational prospective cohort study of patients receiving hospital treatment for SLE in the UK. The eligibility criteria were defined by the NICE recommended subgroup of the current licensed population for belimumab and patients were included from October 2013 onwards who had high disease activity (anti-dsDNA positive, low Complement 3 or 4 level and SLEDAI-2K score ≥ 10). The analysis was performed by the University of Manchester. The analysis included [redacted] patients receiving belimumab and [redacted] receiving rituximab. The two primary endpoints for SLE were BILAG-2004 and SLEDAI-2K, a disease specific instrument (SLICC/ACR damage index) was also used and HRQoL was measured using both generic and disease-specific instruments. Patients were recruited from 39 centres across the UK and followed up at three, six and 12 months then every 12 months. At any point they could stop treatment and restart with a second round of the same treatment or restart with a different one. The analysis used multilevel repeated measures regression modelling of outcomes at three, six and 12 months. Potential confounding variables were identified before analysis and included as covariates in the models. However, BILAG BR data cannot be used to make a reliable comparison of the effectiveness of belimumab versus rituximab due to the different criteria for the use of rituximab and belimumab (See BSR guidelines for SLE[39]).

5. COST EFFECTIVENESS

5.1 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 (5.1.1) 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.

5.1.1 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 an SLR which was conducted to identify economic evaluations of belimumab against any other comparator.

A previous search for relevant cost effectiveness studies to support the NICE submission for belimumab did not identify any relevant economic studies. For this update, resources were searched between 28 January 2020 and 19 February 2020, and no publication date limits were applied. Conference abstracts were restricted to 2017 onwards. The CS states that the searches were limited to English language studies only (CS; p.121), however the response to clarification[47] confirmed that this limit was applied during the inclusion screening stage.

A summary of the sources searched is provided in Table 5.1.

Table 5.1: Data sources for the cost effectiveness systematic review (as reported in CS)

	Resource	Host/source	Date range	Date searched
Electronic databases	Embase	Ovid	Inception - 28/1/20	28/1/20
	PubMed	PubMed	Inception - 28/1/20	28/1/20
	DARE NHS EED HTA Database	CRD website	Inception - 28/1/20	28/1/20
	EconLit	Ovid	Inception - 31/1/20	31/1/20
Conference Proceedings	ISPOR	Conference websites & abstract books	2017+	19/2/20
	ACR			
Additional resources	ClinicalTrials.gov	Web search	Not stated	28/1/20
	RePEc			28/1/20

DARE = Database of Abstracts of Reviews of Effects: NHS EED = NHS Economic Evaluation Database; HTA Database = Health Technology Assessment Database; ISPOR = International Society for Pharmacoeconomics and Outcomes Research; ACR = American College of Rheumatology; RePEc = Research Papers in Economics;

ERG comment:

- Searches were undertaken to identify data published on economic evaluations of belimumab against any other comparator. The CS provided sufficient details for the ERG to appraise the literature searches. Several databases, conference proceedings and other resources were

searched, and reference checking of recent publications was conducted. ClinicalTrials.gov was also searched for ongoing trials. Searches were generally well documented, making them transparent and reproducible.

- Searches were undertaken in January/February 2020 for the CS in November 2020, so could now be considered rather out of date. Data published this year may therefore not have been identified by the searches.
- Most search strategies combined synonyms for belimumab with a filter designed to identify cost and economic evaluation studies. The filters employed included search terms that did not appear to be based on any recognised published filter. Additional synonyms may have helped improve recall, and relevant studies may have been missed from searches of these databases. Searches of the CRD databases, RePEc, EconLit and ISPOR did not include a cost/economics filter however, so this may have helped mitigate against this loss of recall.
- Searches did not appear to be based on the 2010 strategies; however, all databases were searched from inception and therefore did not seem to be intended as update searches to the 2010 study.
- The CS states that 'additional records were identified via grey literature searches and references identified in SLR studies'. In response to clarification,[47] the company said that SLRs were identified through the searches of the databases listed above. The ERG notes, however, that as many of the searches employed cost/economic filters, they may not have retrieved all relevant SLR studies.
- The 'HGS1006'/'HGS 1006' synonym of belimumab is used inconsistently through the searches.
- Line #28 of the PubMed search was not entered correctly and so retrieved 0 hits.

Section B.3.4.4 of the CS states that no formal systematic review searches for HRQoL or resource use data were conducted. It refers to Sections 6.4.5 and 6.5 of the previous submission (TA397) which found that comprehensive systematic literature reviews were not feasible because of the breadth of organ systems that would need to be searched for.

Appendix H/I of the CS outlines the methods used to update cost and utility data of organ systems which were needed to populate the economic model. Formal systematic review searches were not conducted; instead, an update of the 'targeted literature review' was conducted which consisted of a two-staged search approach. The NICE website was initially searched for relevant health technology assessments containing cost and utility data. Where no relevant data were found, PubMed was searched combining general health economics/quality of life terms with disease-specific search terms.

ERG comment:

- As the previous model included searches conducted up to 14 December 2010, publications published after 14 December 2010 were included in this targeted literature review.
- The date the searches were conducted was not supplied, and no numbers of records retrieved were provided for either the NICE website or PubMed searches. No additional databases or other resources were searched.
- The terms used in PubMed for the health economics/quality of life search facet were given in Appendix H/I. They did not appear to be based on any recognised filter and included only MeSH subject heading terms and subheadings. No free text terms were used in the filter for additional synonyms, so the searches were therefore limited in scope and may not have retrieved potentially relevant records.
- In response to clarification,[47] the Company provided a list of the specific disease search terms used in the searches of PubMed and the NICE website. Many of the PubMed searches included only MeSH subject heading terms, and few synonyms were used for any of the disease topics.

It can therefore be concluded that the searches were very restrictive, and potentially relevant records are likely to have been missed.

- Given the nature of the methodology used and the limited documentation provided, the ERG is unable to adequately appraise the quality or appropriateness of the searches for health-related quality of life and resource use studies.

5.1.2 Inclusion/exclusion criteria used in the study selection

The in- and exclusion criteria used for the selection of studies in the cost-effectiveness section of the submission is reflected in Table 5.2.

Table 5.2: In- and exclusion criteria for the use of cost effectiveness studies

	Inclusion criteria	Exclusion criteria
Population	Patients with SLE	Studies which included more than 25% of patients with significant renal involvement (lupus nephritis) or CNS involvement (central nervous system lupus). Belimumab is not currently licensed for the management of lupus nephritis or CNS lupus.
Intervention	Belimumab	No reference to belimumab
Comparators	Standard Therapy alone; belimumab; cyclophosphamide; rituximab	
Outcomes	Total costs; Summary health outcomes (Quality-adjusted Life Years (QALYs)); Incremental cost effectiveness ratio (ICER).	
Study Design	Cost utility analysis Cost effectiveness analysis	Case reports Case studies News Comments Editorials Letters Budget impact, cost comparisons
Limits	Reported in English language Conference abstracts published from 2017 onwards.	Non-English language studies. Full text unavailable Duplicate studies ERG report on the original NICE submission Published in error and withdrawn. Societal perspective analysis Conference papers published before 2017.

Source: Table 54 of the CS[1]
HTA = health technology assessment; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year; SLR = systematic literature review.

5.1.3 Identified studies

Three studies which fit the criteria described above were identified in the review. Specchia et al. (2014)[90] reported on a cost utility analysis performed on belimumab in patients with SLE in an Italian setting. Pierottie et al. (2015),[91] sponsored by GSK, provided detailed reporting on the analysis which had been presented by Specchia et al. (2014). This study was submitted to NICE as part of TA397. CADTH (2012)[92] reported limited details on an economic analysis which was performed as part of the decision for the Canadian Agency for Drugs and Technologies in Health (CADTH).

5.1.4 Interpretation of the review

As no detailed new models, other than the one already used for the past submission, were found the previous model which was constructed for TA397 was re-used for this submission.

ERG comment: The literature review for the search of cost effectiveness studies is appropriate.

5.2 Summary and critique of company's submitted economic evaluation by the ERG

Table 5.3: Summary of the company's economic evaluation (with signposts to CS)

	Approach	Source/Justification	Signpost (location in CS)
Model	Micro-simulation	To be in line with previous NICE TA397 and capture the complex nature of SLE	B.3.2
States and events	Disease activity, organ damage, treatment discontinuation, death	To be in line with previous NICE TA397	B.3.2
Comparators	Standard therapy	Rituximab not a comparator according to current guidelines and insufficient data for comparison. Cyclophosphomide not appropriate comparator as largely reserved for the treatment of lupus nephritis or CNS lupus	B.3.2
Population	HDA2, with (SELENA-SLEDAI) score ≥10 AND at least one of the following serological features: low complement AND/OR positive anti-dsDNA	Subgroup of total SLE population; extension beyond HDA-1 population for which there is a recommendation with justification that the new subgroup will be clinically more relevant	B.3.2
Treatment effectiveness	Based on BLISS, BLISS-LTEs, Johns Hopkins cohort, Toronto Lupus cohort, BILAG-BR	Including long-term effectiveness for belimumab and standard therapy, and patient weight from BILAG-BR	B.3.3
Adverse events	Not included	To be in line with previous NICE TA397	B.3.3

	Approach	Source/Justification	Signpost (location in CS)
Health related QoL	Utilities were estimated based on disease activity using regression analysis on EQ-5D-3L measurements collected during the BLISS-52 and BLISS-76 Phase 3. Utility multipliers from the literature were used to estimate the impact of organ damage on HRQoL.	To be in line with previous NICE TA397	B.3.4
Resource utilisation and costs	Drug acquisition and administration costs, disease activity related costs and organ damage related costs based on multiple sources.	Unit prices were based on National Health Service (NHS) reference prices and inflated using the Personal Social Services Research Unit (PSSRU), consistent with NICE reference case.	B.3.5
Discount rates	Discount of 3.5% for utilities and costs	Consistent with NICE reference case	
Subgroups	No subgroups as company presents mainly the HDA-2 subgroup; HDA-1 is presented in scenario		
Sensitivity analysis	DSA, PSA and scenario analyses were performed.		B.3.8

CS = company submission; DSA = deterministic sensitivity analysis; eMIT = Drugs and pharmaceutical electronic market information tool; EQ-5D-3L = European Quality of Life-5 Dimensions 3 levels; HDA = high disease activity; HRQoL = health-related quality of life; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; LTE = long term extension; PSA = probabilistic sensitivity analysis; PSSRU = Personal Social Services Research Unit; SLE = systemic lupus erythematosus; TA = technology appraisal

5.2.1 NICE reference case checklist (TABLE ONLY)

Table 5.4: NICE reference case checklist

Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether *de novo* evaluation meets requirements of NICE reference case
Population	As per NICE scope	Narrower than NICE scope	The population falls within the NICE scope.
Comparator(s)	Therapies routinely used in the National Health Service (NHS), including technologies regarded as current best practice	Partly	Only standard therapy was included as comparator, the NICE scope listed rituximab and cyclophosphamide

Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether *de novo* evaluation meets requirements of NICE reference case
Type of economic evaluation	Cost effectiveness analysis	Yes	
Perspective on costs	NHS and Personal Social Services (PSS)	Yes	
Perspective on outcomes	All health effects on individuals	Yes	
Time horizon	Sufficient to capture differences in costs and outcomes	Yes	
Synthesis of evidence in outcomes	Systematic literature review (SLR)	Partly	Exclusion of data sources for long-term extrapolation not fully justified
Measure of health effects	Quality adjusted life years (QALYs)	Yes	
Source of data for measurement HRQoL	Described using a standardised and validated instrument	Yes	
Source of preference data for valuation of changes in HRQoL	Time-trade off or standard gamble	Yes	
Discount rate	An annual rate of 3.5% on both costs and health effects	Yes	
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes	
Sensitivity analysis	Probabilistic modelling	Yes	

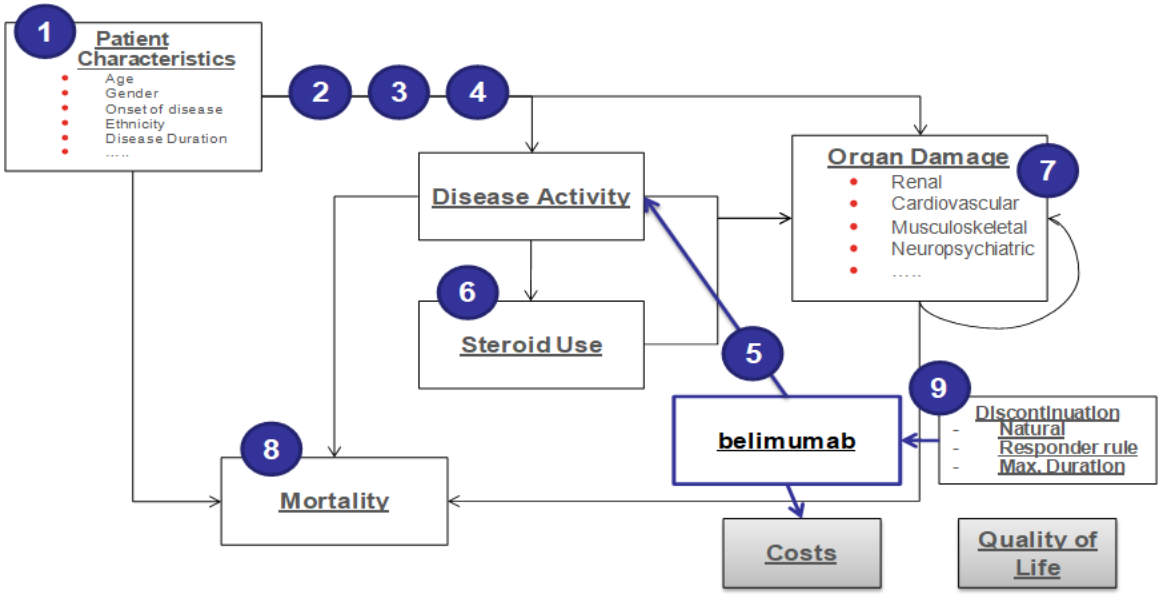
HRQoL = health-related quality of life; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; PSS = Personal Social Services; QALY = quality-adjusted life year; SLR = systematic literature review

5.2.2 Model structure

The company continued to use the de novo model developed for TA397,[44] as no new cost effectiveness analyses were identified. According to the company, the original model structure remained unchanged (Figure 5.1). The model uses a cycle length of one year with a lifetime horizon, with the justification that this best captures the changes in overall disease activity (as measured by SELENA-SLEDAI scores)

and the accumulation of organ damage (as measured by SLICC/ACR Damage Index). A half cycle correction was not included.

Figure 5.1: Depiction of model structure (as presented in TA 397)



ERG comment: The ERG's concerns relate to a) the chosen cycle length, and b) flares being excluded from the model.

a) The ERG questioned whether the annual cycle length was appropriate to reflect the progression of SLE patients. In response to clarification question B6, the company argued that, due to the long-term nature of the disease, an annual cycle length is likely sufficient to capture progression in disease activity and organ damage. This issue is likely not influential.

b) The company also conceded that flares were not captured by their model structure and noted that no data were available to inform modelling of incidence and severity of flares. The ERG accepts that it would be difficult to incorporate flares in the model structure but considers that flares may have been implicitly included through BLISS trial data. In particular the SELENA-SLEDAI (SS) scores measured on a four-weekly basis would have likely captured flares according to the ERG clinical expert. The significant average improvement of SS scores in year 1 regardless of treatment arm, as presented by the company in Figures 6.7 and 6.11 of the original CS[1] would support this, that is, point to regression to the mean, i.e. patients may have had higher than their average SS scores at baseline, possibly due to flares. However, the ERG clinical expert also added that large placebo effects are common in SLE trials and may also be caused by patient care in trials being better than in general clinical practice, which improves adherence. In addition, the ERG clinical expert stated that: *"many patients in trials are treated with steroids, as it is considered unethical not to provide any therapy for flares at time of recruitment, and this may also account for the fast improvement in all groups initially".*[93] It is difficult to assess the impact of the issue of flares not being explicitly modelled, but the ERG agrees that it would have been challenging to incorporate this functionality in the model, also considering the absence of evidence.

5.2.3 Population

TA397 recommends the use of add-on belimumab in a high disease activity (HDA) subpopulation of the total pooled SLE patient population from BLISS-52 and BLISS-76, defined as SELENA-SLEDAI score ≥10 AND low complement AND positive anti-dsDNA, referred to as HDA-1. The company proposed a new, and in their view more clinically relevant, subgroup, which is defined as: SELENA-SLEDAI score ≥10 AND at least one of the following serological features: low complement AND/OR positive anti-dsDNA. This subgroup is referred to as HDA-2. The rationale for the new subgroup is discussed in more detail in Section 3.1 of this report. The company used the HDA-2 subgroup for its new base-case but also presented results for the HDA-1 subgroup. The company also focused its submission on adults.

ERG comment: The ERG's concerns relate to a) limiting the scope of the submission to the HDA-2 subgroup; and b) the focus on the adult population.

a) The relevance of the HDA-2 subgroup has been confirmed by the ERG's clinical expert. It is to be noted that the most significant change compared with TA397 in terms of impact on ICER (the inclusion of the Toronto cohort for estimating long-term disease activity and organ damage) is not specific to any subgroup and hence the effect of its inclusion would move cost effectiveness in both HDA-2 and HDA-1 subgroups in the same direction (decreasing the ICER). However, other updates made to the model increased ICERs, meaning that cost effectiveness of belimumab in the HDA-1 subgroup may differ compared to before. The ERG presents results for the HDA-1 subgroup in scenarios.
b) The ERG wishes to highlight that the company confirmed in their response to clarification question B2[47] that the focus of this submission is solely on adults, due to the limited belimumab data in paediatric SLE patients, particularly in the HDA subgroups.

5.2.4 Interventions and comparators

In TA397,[44] the company presented the belimumab IV formulation. The current submission introduces a SC formulation of belimumab, administered via a pre-filled pen (autoinjector device). The company expects the new formulation to provide comparable efficacy results with the IV formulation, advantages for patients for whom travelling to the hospital to receive a monthly IV infusion is difficult or poses a burdensome interruption to their everyday lives, a reduction in burden on NHS resources (no clinic time needed), and potentially an improvement in equality of access to treatment. The IV formulation is dosed by patient weight, whilst the SC formulation is not. The level of compliance for both is assumed to be 100%. However, the company stated that patients in the BLISS-SC trial had an exposure of 97%. No vial sharing is assumed for the IV formulation.

The dosing for the IV formulation is modelled as follows: belimumab 10 mg/kg is administered as an IV infusion over a one-hour period on days 0, 14 and 28, and at four-week intervals thereafter in addition to standard therapy in a clinic centre by trained nurses. In Year 1 there are 14 administrations and from Year 2 onwards there are 13 administrations per year

The dosing for the SC formulation is modelled as follows: belimumab 200 mg solution for injection in pre-filled pen administered via SC route each week, with 53 doses in the first year, and 52 doses each year thereafter. A one-off training for the self-administration is necessary and included in the model (up to one hour with a specialist nurse).

The final scope for the current appraisal includes standard therapy alone, rituximab plus standard therapy, and cyclophosphamide plus standard therapy as comparators. The company only presents cost

effectiveness analyses using standard therapy as the comparator. Justifications for exclusion of rituximab plus standard therapy and cyclophosphamide plus standard therapy from the economic analyses are discussed in Section 3.3 of this report and in the original submission respectively.

Standard therapy continues to include the use of antimalarials (i.e. hydroxychloroquine), NSAIDs, corticosteroids and immunosuppressants such as azathioprine, methotrexate and mycophenolate mofetil. Many of the treatments used for SLE are unlicensed, with only hydroxychloroquine, and corticosteroids (not azathioprine, although widely used) licensed for use in SLE.

ERG comment: The ERG's concerns relate to a) the exclusion of rituximab plus standard therapy as comparator and b) the new SC formulation.

a) KEY ISSUE. The ERG's concern relates to omitting a comparator named in the scope. According to the ERG's clinical expert it was likely difficult to undertake this comparison (see Section 3.3), but nevertheless the comparison may be relevant.

Table 5.5: Key issue 6: Rituximab excluded as comparator

Report section	Section 4.2.6 and 5.2.6
Description of issue and why the ERG has identified it as important	Rituximab was excluded as a comparator but may be a relevant comparator and was mentioned in the scope.
What alternative approach has the ERG suggested?	Include rituximab as comparator
What is the expected effect on the cost effectiveness estimates?	Unknown
What additional evidence or analyses might help to resolve this key issue?	The ERG understands that this comparison will be difficult.

b) The new subcutaneous formulation is administered at a fixed dose. According to our clinical expert, it may be slightly less effective than the IV formulation, but it offers other advantages in that it is easier to use and cheaper to administer. It is also generally well tolerated, although according to the ERG clinical expert, there are some patients that do not tolerate SC injections well or prefer less frequent IV treatment. The ERG furthermore considers that the company should have included both formulations in one model to enable direct comparison.

Table 5.6: Key issue 7: IV and SC formulations are not compared

Report section	Section 4.2.6 and 5.2.6
Description of issue and why the ERG has identified it as important	IV and SC formulations are not compared with each other. Two separate model files are provided.
What alternative approach has the ERG suggested?	Include both formulations in one model file and enable comparison of IV and SC formulations (full incremental analysis).
What is the expected effect on the cost effectiveness estimates?	Cost effectiveness will not be affected but this would enable a comparison of cost effectiveness of IV and SC.

What additional evidence or analyses might help to resolve this key issue?	

5.2.5 Perspective, time horizon and discounting

An NHS and personal social service perspective for the analysis is adopted and discounting is applied at 3.5% for both costs and benefits.

ERG comment: Perspective and discounting are appropriate.

5.2.6 Treatment effectiveness and extrapolation

Patient baseline characteristics

Baseline characteristics of the HDA-2 patient subgroups for both the IV and SC models were drawn from the total pooled SLE patient population recruited into the two Phase III clinical trials: BLISS-52 and BLISS-76, (excluding the belimumab 1 mg/kg treatment arm). Baseline weight distribution for belimumab IV patients was obtained from the BILAG biologics registry (BILAG-BR). Characteristics for the SC model were drawn from the full study population who participated in the Phase III BLISS-SC clinical trial.

Aligned with the previous submission, an individual organ damage item score was drawn from a multinomial distribution with each category having the probability as outlined in Table 5.7. This reflects the baseline SLICC/ACR Damage Index (SDI) item score occurrences observed in the HDA-2 subgroup.

Table 5.7: Individual SLICC item scores simulated at baseline for the HDA-2 population

SLICC damage item	IV model HDA-2 subgroup					SC model HDA-2 subgroup				
	Score 0	Score 1	Score 2	Score 3	Score 4	Score 0	Score 1	Score 2	Score 3	Score 4
Cardiovascular	94.0%	5.1%	0.9%	0.0%	0.0%	96.8%	2.7%	0.5%	0.0%	0.0%
Diabetes	97.6%	2.4%	0.0%	0.0%	0.0%	98.4%	1.6%	0.0%	0.0%	0.0%
Gastrointestinal	96.4%	3.4%	0.2%	0.0%	0.0%	96.8%	3.2%	0.0%	0.0%	0.0%
Malignancy	99.6%	0.4%	0.0%	0.0%	0.0%	99.1%	0.9%	0.0%	0.0%	0.0%
Musculoskeletal	87.4%	8.8%	3.2%	0.4%	0.2%	90.6%	7.6%	1.6%	0.2%	0.0%
Neuropsychiatric	88.9%	9.2%	1.5%	0.4%	0.0%	93.6%	6.2%	0.2%	0.0%	0.0%
Ocular	93.6%	6.2%	0.2%	0.0%	0.0%	90.2%	9.4%	0.5%	0.0%	0.0%
Peripheral vascular	94.4%	5.1%	0.4%	0.2%	0.0%	95.9%	3.4%	0.7%	0.0%	0.0%
Premature gonadal failure	98.9%	1.1%	0.0%	0.0%	0.0%	98.2%	1.8%	0.0%	0.0%	0.0%
Pulmonary	97.0%	2.8%	0.2%	0.0%	0.0%	97.0%	2.7%	0.2%	0.0%	0.0%
Renal	97.4%	2.6%	0.0%	0.0%	0.0%	98.4%	1.6%	0.0%	0.0%	0.0%
Skin	92.1%	7.1%	0.6%	0.2%	0.0%	94.1%	5.3%	0.7%	0.0%	0.0%
Source: Table 60 of the CS[1]										

One-year treatment effects

One-year treatment effects were included in the model based on disease activity (SS scores) as observed in the relevant BLISS trials (BLISS-52 and BLISS-76 in the IV model, and BLISS-SC in the SC model). This was unchanged from the previous submission (TA397;[44] for the IV model).

The company stated that the methodology used to determine a patient's change in SS score at week 52 was consistent with the previous submission. A linear regression model was fitted to the pooled BLISS IV trial data (or BLISS-SC study for SC) to explain the difference between the SS score at baseline and week 52, depending on baseline SS score combined with a treatment indicator variable, and a "response" indicator variable identifying whether or not patients are classified as satisfying the treatment continuation rule at week 24 with belimumab (see Table 5.8).

Table 5.8: Linear regression explaining change in SS score after 52 weeks compared to ST for the HDA-2 population

Parameter	IV model – HDA-2				SC model – HDA-2			
	Estimate	Std Error	t-value	p-value	Estimate	Std Error	t-value	p-value
SS0 ST	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SS0 all belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SS0 belimumab responders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Note "responders" are patients on belimumab who satisfy the treatment continuation rule.
Source: Table 61 of the CS[1]

Treatment continuation

The company stated that reasons for treatment discontinuation in the current submission remain consistent with the reasons for treatment discontinuation provided in the previous submission: natural discontinuation, and no longer deriving clinical benefit from treatment. In both the IV model and the SC model for both the HDA-1 and HDA-2 subpopulations, patients on belimumab had to satisfy the treatment continuation criterion, defined as demonstrating a SS score decrease of four points or greater at week 24. Patients on belimumab who did not satisfy the treatment continuation criterion at week 24 remain in the belimumab arm of the model but continue to receive ST treatments after this time-point and assume the average ST level of disease activity for the remainder of the model horizon.

This was modelled by the company by using the probability of treatment continuation after week 24 in the BLISS trials (i.e. only for patients with a reduction of ≥ 4 points in SS score) stratified by baseline SS score. Hence, the company estimated the percentage of patients defined as a "responder" for each of the baseline SS scores. In the model, treatment continuation is determined for belimumab patients by using a Bernoulli distribution with a probability corresponding to the probabilities based on baseline SS scores.

To derive year 1 natural discontinuation rates for patients receiving belimumab, an analysis for HDA-1 and HDA-2 populations was conducted on the relevant pivotal Phase III BLISS trials for each formulation. As no long-term randomised controlled trial exists beyond 76 weeks for the IV formulation and 52 weeks for the SC formulation, data to calculate the natural discontinuation probability in years

subsequent to year 1 were derived from an analysis of LTE study data. Continuation and discontinuation rates can be found in Table 5.9.

Table 5.9: Summary of percentage belimumab continuations and natural discontinuation for HDA-2

	IV model HDA-2 subgroup	**SC model HDA-2 subgroup**
% belimumab patients satisfying treatment continuation rule at 24 weeks	[redacted]	[redacted]
Natural discontinuation	**Patients satisfying treatment continuation at 24 weeks**	
KM estimate week 76 IV, week 52 SC	[redacted]	[redacted]
Daily hazard rate (wk24-wk76 IV, wk 24-52 SC)	[redacted]	[redacted]
Year 1	[redacted]	[redacted]
Subsequent years	[redacted]	[redacted]
Source: Table 62 of the CS[1]		

Extrapolation to long-term SLE outcomes

The Johns Hopkins (JH) cohort was used to develop a natural history model (NHM) for patients with SLE. In both the IV and SC models, rather than using SS scores to reflect disease severity over time, the scores are used to calculate the Adjusted Mean SLEDAI (AMS) score. This approach remains unchanged compared to TA397.[44]

Organ damage reduction on belimumab

The original IV cost effectiveness model presented in TA397 was populated using up to 1.5 years of observed effectiveness data derived from the Phase 3 BLISS-52 and BLISS-76 clinical studies.

In the current STA, the long-term extension studies BLISS-52 and BLISS-76 were used to extrapolate long-term effects on disease progression (e.g. organ damage and mortality). To this extent a propensity score matched (PSM) analysis has been undertaken to estimate the long-term comparative effectiveness of belimumab plus ST compared with ST from a matched population. The primary analysis of the PSM was conducted using the BLISS-76 US open label extension study population to compare organ damage progression (SDI score) from baseline (defined as first exposure to belimumab) to Year 5 in patients treated with belimumab or ST.[68]

In the absence of a control arm, BLISS LTE patients were propensity score matched post-hoc 1:1 to the Toronto Lupus Cohort (TLC) to obtain comparative evidence on organ damage progression compared with ST alone. The primary endpoint of the PSM comparative analysis was the difference in change of total SDI score from baseline to five years between patients on belimumab compared with those on ST from the TLC in patients with ≥5 years of follow-up.

Results of the PSM analysis demonstrated that over a five-year period, patients treated with belimumab experienced a five-year SDI change of 0.283 (95% CI 0.166 to 0.400), which represented less organ damage compared with patients treated with ST alone (who had a five-year SDI change of 0.717 [95% CI 0.500 to 0.934]).

Calibrating the model using the PSM analysis

The cost-effectiveness model was validated by comparing the modelled long-term organ damage progression results to the observed five-year SDI progression data for belimumab and ST. To ensure comparability of the simulated model results with the long-term evidence the baseline characteristics of the model population were re-adjusted to reflect the BLISS LTE population.

As the IV model captures the observed pooled analysis results from the pooled BLISS LTE studies, it was decided that the validation exercise of the deterministic model should be simulated as a five-year increase in SDI score (further from the baseline duration of 1.5 years). The model starts at the beginning of the BLISS trial, hence the period from 1.5 to 6.5 years from the model was chosen to compare with the PSM analysis results. This simulated an SDI score increase of 0.568 in the belimumab arm and 0.611 in the ST arm, respectively (see Table 5.10).

The company argued that, compared with the results from the PSM analysis, it was apparent that the existing cost effectiveness model overestimated SDI progression in the belimumab arm and underestimated SDI progression in the ST arm.

Table 5.10: 5-year SDI increases, modelled versus real world data

5-year SDI increase	Belimumab + ST	ST
Cost effectiveness model; matched LTE ITT population	0.568	0.611
Propensity score-matched analysis	0.283	0.717

Source: Table 65 of the CS[1]

To account for the difference in the model's predicted SDI progression and the results from the long-term evidence, a calibration factor was derived and applied to allow for adjustment of the existing natural history model in the cost effectiveness model. To derive the calibration factor, the model was simulated several times with varying calibration factors, until the model's results matched the observed results from the PSM up to 3 decimals (see Table 5.11).

The model calibrations resulted in the amendment of the original organ damage probabilities in the time-to-event risk equations in the model. For ST, this implied that the annual risk of organ damage for ST was adjusted upwards with 18.6%, in order to reflect the observed long-term organ damage progression after five years with ST. For belimumab, this implied that the annual risk of organ damage for belimumab was adjusted downwards with 50.9%. However, the company only used the calibration factor for the belimumab arm, not the ST calibration factor.

Table 5.11: Calibrated five-year increase in SDI score

5-year SDI increase*	Belimumab + ST	ST
Model results with no calibration	0.568	0.611
Observed 5-year SDI increase from PSM	0.283	0.717
Ratio of observed vs. current SDI value	0.498	1.173
Calibration factors	0.491	1.186
Model results with calibration factors	0.283	0.717

* SDI increase between t=1.5 and t=6.5.
Source: Table 66 of the CS[1]

Mortality

Using the JH cohort data, a Weibull survival model was developed explaining the risk of death based on AMS. In order to avoid an underestimation of mortality in the model a correction was required to increase mortality risk at older ages using mortality estimates for the general population (see Table 6.12 of the original CS in TA397[44]). According to the company a Weibull distribution was chosen to extrapolate existing mortality data due to goodness of fit according to the Akaike information criteria. For the new submission an update was conducted to reflect the most recent UK values. The model does not include the incidence and severity of flares in the disease activity, nor organ damage. The approach to calculate patient mortality as in the economic analysis described in TA397[44] remains unchanged.

ERG comment: The ERG's concerns relate to a) the use of the calibration factor to adjust long-term organ damage estimates for the belimumab arm; b) implementation of 24-week response and treatment continuation in the model; c) an error in modelling of non-responder disease activity at 52 weeks; d) modelling of belimumab non-responder disease activity after 52 weeks; e) correlations between patient baseline characteristics has not been taken into account in the simulation of patients in the model; f) some patient baseline characteristics were taken from different studies compared to the effectiveness estimates; g) concerns surrounding the choice of survival models to assess mortality; h) long-term corticosteroid sparing effect of belimumab; i) the exclusion of the BILAG-BR study to estimate long-term effectiveness;

a) KEY ISSUE. As highlighted in Section 4.2.6 of this report the ERG questions the validity of the calibration factor. The company used results of the PSM study as a justification that the cost-effectiveness model overestimated SDI progression in the belimumab arm and underestimated SDI progression in the ST arm and hence employed a calibration factor to adjust the existing NHM for belimumab to match PSM results. The use of this calibration factor is a major driver of cost effectiveness outcomes. For example, in response to CQ12f,[47] the company provided a scenario without using the calibration exercise, resulting in an ICER of £47,872 per QALY gained for the IV model with the HDA-2 population and an ICER of £56,277 per QALY gained for the SC model with the HDA-2 population. The ERG main concerns regarding the use of the calibration factor include:
    - Since the company only used the calibration factor on the belimumab arm of the model, it was not necessary to use the PSM to obtain the calibration factor. The PSM reduces the number of patients included with the aim to obtain a comparable set of patients to the TLC. As no comparative estimates for treatment effectiveness or long-term outcomes are derived, the purpose of using this reduced (matched) set of patients is questionable, in particular because the resulting set of patients is not necessarily more generalisable to the UK setting than the original sample.
    - It is further questionable whether results of the PSM are applicable to the HDA-2 population. SDI change from baseline was estimated using the PSM on the total patient population in the LTE, without restrictions in terms of SS score or complement levels. In response to CQ12, the company argued that if a restriction would have been applied to only patients who meet the HDA-2 subgroup criteria (or any subgroup that restricts the numbers of patients as compared to the total BLISS LTE population), patient numbers would be small and therefore limit the power required for analyses to be conducted robustly.
    - Methodological issues, highlighted in Section 4.2.6, related to 1) unobserved differences between the BLISS LTE and TLC studies; and 2) the fact that the BLISS LTE studies followed only people continuing to use belimumab and who (if still in the cohort after five years and thus still using belimumab) are more likely to have any/all of the following: 1)

little disease progression, 2) a continuous response to belimumab, 3) no/few adverse effects from using belimumab. Using only the five-year BLISS-LTE estimates to obtain long-term organ damage estimates will likely be biased and result in an over-estimate of belimumab effectiveness.

- Calibrating the existing model to one observed time-point (five years) is problematic, especially given the attrition of patients treated with belimumab in the BLISS LTE studies mentioned above and in Section 4.2.6.
- The pre-calibration model already assumes a long-term treatment effect based largely on clinical expert opinion. In TA397, it is stated that *"In the simulation model, an assumption was made that the additional absolute effect of belimumab on disease activity reduction remains constant after one year. This is a key model assumption and was discussed with Professor Petri who has observed patients on belimumab in her clinic for a number of years as part of the Phase 2 open-label extension study"*.[93] The ERG clinical expert considers this a reasonable assumption.
- The pre-calibration model already assumes that the absolute SS value is adjusted in comparison to that predicted by JH. In TA397,[44] it is stated that *"The adjusted JH model was used to predict the SLEDAI score of a patient treated with SoC after one year. The model allows for the selection of the original and adjusted JH model. For the base case analysis, the adjusted JH model has been used"*. In the final appraisal document of TA397, the committee also noted that "*that there was uncertainty about whether the equations derived from the Johns Hopkins data could be reliably applied to the target population because of differences in study populations*".[93] Given that the adjustment favours the effectiveness of belimumab, it is likely that this adjustment causes the ICER to be underestimated.

In conclusion, the ERG considers the use of the calibration factor inappropriate; it unfortunately does not resolve any uncertainty about comparative long-term organ damage estimates of belimumab versus standard therapy. The ERG advises that the calibration factor should not be used and removed it in its base-case. In scenarios, the ERG explores 1) the use of the company's calibration factor; 2) the use of the company's calibration factors for both treatment arms; and 3) the use of the unadjusted JH model.

Table 5.12: Key issue 8: Use of the calibration factor

Report section	Section 4.2.6 and 5.2.6
Description of issue and why the ERG has identified it as important	The calibration factor lacks validity and should not be used. It does not resolve uncertainty about long-term organ damage in patients treated with belimumab versus standard therapy due to issues with methodology, the BLISS-LTE evidence and the PSM study. The use of this calibration factor is a major driver of cost-effectiveness outcomes. For example, the company provided a scenario without using the calibration exercise, resulting in an ICER of £47,872 per QALY gained for the IV model with the HDA-2 population and an ICER of £56,277 per QALY gained for the SC model with the HDA-2 population.
What alternative approach has the ERG suggested?	The ERG considers it most appropriate to remove the calibration factor in its base-case.
What is the expected effect on the cost effectiveness estimates?	The use of a calibration factor likely results in an overestimation of treatment effect of belimumab and hence underestimates the ICER.

What additional evidence or analyses might help to resolve this key issue?	Empirical evidence is lacking to validate the calibration factor. The issues with the long-term belimumab data are unlikely to be resolved.

b) KEY ISSUE. In the model, the probability of being a responder is based on baseline SS score, which is linked to the responder criteria applied to patients in the BLISS trials (i.e. only for patients with a reduction of ≥ 4 points in SS score). Hence, it is estimated at baseline in the model and not directly linked to the actual improvement in SS score in the model. In turn, actual SS scores are estimated based on a regression model where response is an independent variable, given that a 24-week time point does not exist in the model. As a result, a large proportion of patients is classed as non-responders but experiences >4 points reduction in SS at 52 weeks. This could lead to under-estimation of costs in the model compared to clinical practice as patients with no response do not continue belimumab. A structural adjustment to the company's model is necessary to resolve this issue.

Table 5.13: Key issue 9: Implementation of 24-week response and treatment continuation in the model

Report section	Section 5.2.6
Description of issue and why the ERG has identified it as important	The probability of being a responder is based on the baseline SS score, which is linked to the responder criteria applied to patients in the BLISS trials (i.e. only for patients with a reduction of ≥ 4 points in SS score). Hence, it is estimated at baseline in the model and not directly linked to the actual improvement in SS score in the model. In turn, actual SS scores are estimated based on a regression model where response is an independent variable, given that a 24-week time point does not exist in the model. As a result, a large proportion of patients is classed as non-responder but experiences >4 points reduction in SS at 52 weeks.
What alternative approach has the ERG suggested?	An adjustment of the model to align reduction in SS at 52-weeks with the defined criteria for being responder/non-responder (i.e. >4 points reduction in SS).
What is the expected effect on the cost effectiveness estimates?	This could lead to under-estimation of belimumab costs in the model compared to clinical practice as patients with no response do not continue belimumab.
What additional evidence or analyses might help to resolve this key issue?	Not applicable

c) KEY ISSUE. The ERG noted an error in the modelling of non-responder disease activity at 52 weeks. In the CS, the company mentions that *"the methodology used to determine a patient's change in SS score at week 52 is consistent with the previous submission"*.[1] In the CS it is mentioned that a linear regression explaining change in SS score after 52 weeks compared to ST for the HDA-2 population was used (table 61 of the CS). Using this function, the company mentions that disease activity at one year is calculated using the baseline SS score combined with a treatment indicator variable, and a "response" indicator variable identifying whether or not patients are classified as satisfying the treatment continuation rule at week 24 with belimumab. In the CS of TA397 (Table 6.5), an example calculation is provided on how to use the regression function, which shows that for SoC a coefficient of -0.379 should be used, for belimumab non-responders a coefficient of -0.327, and for belimumab responder a coefficient of -0.327 + -0.318 should be used. However, the coefficient for belimumab non-responders is

implemented wrongly due to an error in the coding which results in belimumab non-responders to be assigned the same coefficient as the SoC group (this is caused by the function "*Public Sub UpdSSFirstHalfYear()*" in the company model; line "*If mInput.respAnalysis = True And curPat.responder = False Then curPat.mTx = SoC*", which assigns SoC treatment to non-responders, preceding the 52 week analysis, such that belimumab non-responders are modelled as if they had been allocated to SoC at the start. The ERG has therefore moved this treatment reassignment to after the disease activity is updated in the first year – see Figure 5.2). For example, patient 5 in the belimumab arm is a non-responder with baseline SS score of 16; consequently her disease activity at one year is estimated to be: 16+(-0.379*16)=9.936. However, the ERG believes, in line with the company's presentation in the CS, that it should be: 16+(-0.327*16)=10.768. In addition, Figure 6.5 in the CS of TA397[44] also seems to indicate that belimumab non-responders have a smaller reduction in disease activity compared to SoC. This difference leads to an overestimation of treatment effectiveness in the belimumab arm (in particular the belimumab non-responders).

Figure 5.2: Changes to CS model by ERG to overcome an error in modelling of non-responder disease activity at 52 weeks

```
Private Sub UpdSS(firstyear As Boolean)

    Dim tmp As Double, mSS_Bel As Double, fact_eff As Double
    tmp = curSS

    If firstyear Then
        mSS_SoC = mSS_SoC + mSS_B * (mInput.SSRed(1))

        If curPat.mTx = Belimumab Then
            mSS_Bel = mSS_SoC + diffwk52
            curSS = mSS_Bel
            ' ERG_1: moved the next 3 lines here from first half year calculations
            If mInput.respAnalysis = True And curPat.responder = False Then
                  curPat.mTx = SoC
            End If

        Else
            curSS = mSS_SoC
        End If
```

Table 5.14: Key issue 10: Error in calculation of belimumab non-responder disease activity at 52 weeks

Report section	Section 5.2.6
Description of issue and why the ERG has identified it as important	Due to a modelling error, belimumab non-responders have the same reduction in disease activity at 52 weeks as patients in the standard therapy arm.
What alternative approach has the ERG suggested?	The ERG proposes a correction of the model to align reduction in SS at 52-weeks for non-responders with the regression function mentioned in the CS.
What is the expected effect on the cost effectiveness estimates?	This difference leads to an overestimation of treatment effectiveness in the belimumab arm at 52 weeks (in particular the belimumab non-responders).
What additional evidence or analyses might help to resolve this key issue?	Correction of the company's modelling

d) The ERG considered that the company's assumption that belimumab non-responders beyond 52 weeks have equal disease activity to the average patient on standard therapy might be questionable. However, the ERG's clinical expert confirmed that this was a reasonable assumption. Even though response failures on belimumab in the first year may be confounded with response failures on standard therapy, which could explain the differences between belimumab non-responders and patients treated with standard therapy in the first year, the ERG clinical expert considers that in the longer run, i.e. from one year onwards, belimumab non-responders would experience similar disease activity to the average patient (and therefore those in the standard therapy arm).
e) No correlation has been taken into account in the simulation of patient baseline characteristics. Despite the potential correlation between baseline characteristics, they are sampled independently. Bootstrapping (i.e. sampling from the trial data) was considered but would underestimate the actual heterogeneity when simulating 50,000 patients. However, in response to CQ15,[47] the company argued that although it is possible to generate implausible patient profiles, this probability of this happening in the model is extremely low. The ERG agrees that this is likely only a minor issue.
f) Patient baseline characteristics were taken from the Phase 3 BLISS-52 and BLISS-76 clinical studies except for the baseline weight distribution for belimumab IV patients which was obtained from the BILAG-BR. Given that the BILAG-BR study was not used to assess long-term outcomes, the ERG is concerned that baseline characteristics of the patients do not match the effectiveness estimates. However, the ERG acknowledged that the impact of this difference is likely to be minor given that the differences between baseline characteristics between the short-term and long-term studies does not appear to be large (e.g. see Table 17 and Table 19 of the CS). Hence, the ERG explored a scenario in which the mean weight from the BLISS IV trials was used instead of BILAG-BR.
g) In the CS, it is stated that *"using the JH cohort data, a Weibull survival model was developed explaining the risk of death with AMS included and SS item involvement effects removed*".[1] In response to clarification question B10,[47] the company provided goodness of fit statistics (AIC only) of three other common survival models (i.e. exponential, Gompertz, and log-logistic). The ERG would have liked to see goodness of fit statistics using a log-normal model (as another commonly used model) and BIC goodness of fit for all models. Although the likely impact is small, the ERG considers that the choice of the best fitted survival model could have been made more carefully by the company. Moreover, the generalisability of the JH cohort to the UK is questionable given differences between both populations (e.g. black ethnicity of 38% in the JH cohort compared to 7.9% for the HDA-2 population in the model; see CS table 71[1]).
h) Long-term corticosteroid sparing effect of belimumab remains unchanged in the current submission compared to TA397. Hence, the concerns of the ERG in TA397 remain similar in that *"the corticosteroid sparing effect, together with other belimumab benefits such as reduced flare frequency, would reduce the development of organ damage and would therefore translate into long-term benefit. However, data are only available for six years, which indicates that there is a substantial degree of uncertainty over whether the effects observed in the data would translate into longer-term effects"*.[44] This was also mentioned in the FAD, stating that *"the committee concluded that these data suggested, but were not definitive proof of, a reduction in corticosteroids associated with belimumab treatment. However, the Committee understood the importance of reduction in corticosteroid dose for patients and recognised the positive indications of these findings"*.[44]
i) Results of the BILAG-BR registration were only used for a limited number of parameters in the model (e.g. BILAG-BR-captured weight distribution for belimumab patients in the IV model). Although it is mentioned in the CS[1] that the BILAG-BR results were not used because

"there is a high likelihood of confounding, including selection bias in the treatment groups, so that the data captured is not suitable to test the causal efficacy of the treatment or compare treatment efficacy", this is also the case for the PSM analysis. The ERG questioned whether results of the BILAG-BR study could also have been used to validate long-term effectiveness estimates of the model (instead of the PSM), however the ERG clinical expert highlighted that there was variation in baseline demographic features, and the non-biologics group had much shorter disease duration. It is, also worth noting that this comparison would not resolve the bias within the BLISS LTE studies caused by high withdrawal rates, as discussed in Section 4.4 of this report.

5.2.7 Adverse events

Consistent with the economic model provided as part of TA397, adverse events (AEs) continue not to be included in the IV and SC models.

ERG comment: The ERG clinical expert highlighted that there is evidence in the BLISS-52 and BLISS-76 LTE studies that allowed non-US patients with SLE to continue belimumab treatment, in order to evaluate its long-term safety and tolerability including organ damage accrual.[87] The ERG considers that the company could have explored how this data source could have been used for including adverse events in the model.

5.2.8 Health-related quality of life

Health-related quality of life (HRQoL) in the cost effectiveness model is based on calculations relating to 1.) a 'clean' utility equation including the SS score to incorporate the main symptoms and 2.) utility multipliers to incorporate dis-utility from the organ damage sustained. HRQoL was calculated in the same manner for the belimumab SC and IV models. Adverse reactions were not included as it was argued in Section B.3.4.5 of the CS[1] that there were limited differences in the adverse events experienced by the treatment groups in the BLISS trials and that therefore adverse reactions would not be an important utility differentiation between the arms of the economic models.

For 1), based on EQ-5D-3L measurements collected during the BLISS-52 and BLISS-76 Phase III studies a linear regression analysis was used to determine a patient's utility free from organ damage items. The equation for this "clean" utility was:

$$U = 1.275 - 0.14 * \log e\,(age) - 0.036\ * black - 0.009 * SS$$

where 'age' stands for the age of the patient, 'black' stands for African ethnicity, and 'SS' stands for the SELENA-SLEDAI score during the yearly cycle (based on CS Page 159).[1]

For 2), organ damage was included in the form of utility multipliers. A literature review was conducted in order to update the utility values of key organ damage systems in addition to the SLICC/ACR Damage Index (SDI). No update was conducted for utility values of diabetes gastrointestinal, ocular and premature gonadal failure and skin organ systems, for which utility multipliers thus remained unchanged from TA397. Of the remaining seven systems, six systems were updated (cardiovascular, neuropsychiatric, pulmonary, malignancy, peripheral vascular). Out of these, only utilities for organ damage of the renal system were not updated. Different types of damage to the respective organ system were weighted according to their incidence divided by the total number of events in the organ system. The utility multipliers for year 1, 2 and subsequent years is presented in Table 5.15. To avoid underestimating utility, only the damaged organ system which resulted in the largest utility loss was then multiplied with the 'clean' utility (equation above).

Table 5.15: Utility multipliers for organ damage

Organ Damage System	Disutilities		
	Year		
	1	2	Subsequent
Cardio-vascular	0.779	0.806	Same as year 2
Diabetes	0.91	0.91	Same as year 2
Gastro-intestinal	0.79	0.91	Same as year 2
Malignancy	0.837	0.837	Same as year 2
Musculo-skeletal	0.655	0.729	Increasing
Neuro-psychiatric	0.713	0.772	Same as year 2
Ocular	0.97	0.99	Same as year 2
Peripheral vascular	0.863	0.873	Same as year 2
Source: CS[1] Table 68			

ERG comment: The ERG's concerns relate to a) the impact of SLE on HRQoL may not be appropriately captured; b) a lack of transparency and incorrect calculation of utility estimates; d) the utility equation not being re-estimated based on the BLISS-SC trial and c) uncertainty about organ damage multipliers.

a) HRQoL was valued through measurements of the EQ-5D. Measures which are more specific to the quality of life of patients suffering from SLE could be recommended in this case as the EQ-5D does not have items pertaining to body image and fatigue which are relevant according to measures such as the LupusQoL. The use of the EQ-5D to value health-related quality of life is, however, in line with the NICE reference case. It may be that the full impact of the disease on HRQoL is not captured.
   To estimate individual patients' HRQoL in the model, the SS score, a physician-assessed disease activity measure was used as the only SLE-related predictor. Physician-assessed disease-activity as measured by the BILAG-2004 has been demonstrated to not be a good predictor of HRQoL as measured by the LupusQoL or the SF-36.[94] This may be explained through the lack of measurement of more subjective items such as body image and fatigue which are relevant according to measures such as the LupusQoL but not measured by physician-assessed disease-activity measures. Additionally, the ERG's clinical expert noted that the SS-score "does not cover some of the rarer and severe manifestations or scores them relatively low".[39] The ERG understands that it was difficult to incorporate better predictors for HRQoL in the model but wishes to reiterate that the impact of SLE on HRQoL may not be appropriately captured in the model. The direction of any bias is unknown.
b) KEY ISSUE. A transparency issue was found in the calculation of utilities described in section B.3.4.2 of the economic model: according to the CS the equation described above was used to calculate utilities. It is unclear where the coefficients used in this equation come from as the values could not be found again in the submission text or the model. In the model, the following equation has been used, which differs in its coefficients:

$$U = 1.279 - 0.145 * \log e(AGE) - 0.054 * BLACK - 0.009 * SS$$

   These coefficients used in the model result from the use of the "reduced model" from Table 6.20 in the original submission.[95] However, organ damage covariates were excluded in this without re-estimating the remaining coefficients, *"to reduce complexity in calculating utilities"*.[1] The exclusion of organ damage coefficients to simplify the utility function renders

the calculation incorrect as coefficients which were considered statistically significant are eliminated without re-estimating the other coefficients. The ERG would prefer the use of re-estimated coefficients after excluding the organ damage covariates.

Table 5.16 Key issue 11: Violation in utility estimation

Report section	Section 5.2.8
Description of issue and why the ERG has identified it as important	The SLE-related utility estimate excludes key organ damage covariates without adjusting the remaining coefficients.
What alternative approach has the ERG suggested?	The ERG suggests re-estimating the utility model coefficients after excluding the organ damage covariates.
What is the expected effect on the cost effectiveness estimates?	Probably minor, direction unknown
What additional evidence or analyses might help to resolve this key issue?	The ERG suggests re-estimating the utility model coefficients after excluding the organ damage covariates and providing an explanation regarding the discrepancy between model coefficients presented in the CS and the ones used in the model.

c) KEY ISSUE. According to the original submission (Table 6.20), the influence of organ damage measures on HRQoL was not significant in the estimates based on BLISS data. The reduced model including organ damage coefficients also seems to have only limited validity due to the positive ocular damage coefficient, implying an improvement HRQoL when ocular damage occurs. The company excluded these coefficients but included organ damage multipliers from the literature which were applied to the equation above with the argument that *"(...)no statistically meaningful relationships with quality of life were found. However this does not imply that in reality there is no quality of life impairment associated with these organ damage manifestations"* (Page 235-236, original submission).[95] Organ damage items not having a statistically significant influence on HRQoL could be explained by the SS score having a higher explanatory value. However, some uncertainty remains as the utility multipliers which were used have a larger influence than the coefficients used for estimating the impact of disease activity. According to the clinical expert, the weighting of organ damage utilities may be inaccurate, under-estimating the weight of some organ damage incidences while over-estimating others. For example, the clinical expert did not agree with the very low incidence of musculoskeletal organ damage item avascular necrosis, nor with the incidence of cranial or peripheral neuropathy being larger than that of stroke in the neuro-psychiatric organ damage items, and questioned the sources. This may lead to biases in the overall organ damage utilities, where the direction is unknown. This adds to the uncertainty about the implementation of organ damage multipliers. The company attempted to mitigate any over-estimation of the impact of organ damage utility multipliers on HRQoL by only using one utility multiplier even where more than one organ had been damaged. The ERG performed scenarios excluding the impact of organ damage on HRQoL and found that the ICER of belimumab versus standard therapy was reduced compared to the ERG base-case (from ERG base-case £52,891 to £48,347 in the IV model) whilst it increased when it was compared with the company's base-case conditional on an error being fixed (from company base-case with ERG fixing error 1 fixed £31,695 to £32,775 in the IV model). This difference in how the impact of organ damage on HRQoL affects the ICER is related to patients' organ damage duration, which is always longer in the belimumab arm because belimumab, in the model, extends patients' lives. With the calibration

factor, organ damage duration for belimumab patients is reduced because organ damage is delayed, but still remains slightly longer compared with ST patients. The slight increase in the ICER when the impact of organ damage on HRQoL is removed and the calibration factor is used is then likely explained by, on average, less severe organ damage occurring for belimumab patients compared with ST patients.

Table 5.17: Key issue 12: Uncertainty about organ damage utility multipliers

Report section	Section 5.2.8
Description of issue and why the ERG has identified it as important	Uncertainty about organ damage utility multipliers – this may over-estimate the impact of organ damage on HRQoL as the utility estimation function may capture this to a certain extent.
What alternative approach has the ERG suggested?	To explore this uncertainty, use a scenario in which the organ damage utility multipliers are disabled
What is the expected effect on the cost effectiveness estimates?	When the company's calibration factor is used, the ICER increases upon removal of organ damage utility multipliers. When the calibration factor is removed, the ICER decreases upon removal of organ damage utility multipliers.
What additional evidence or analyses might help to resolve this key issue?	The company could investigate whether the weighting of organ damage items corresponds with the latest evidence and consult expert opinion on the magnitude of the organ damage utility multipliers.

d) The SLE-based utility model continues to be estimated based on data from BLISS-52 and BLISS-76 and the IV related utilities are therefore also used in the SC model. The direction in which this influences the ICER is not clear.

5.2.9 Resources and costs

Resource use included in the model can be structured into the three parts: 1) treatment cost, 2) disease-related activity cost and 3) key organ damage cost. Standard therapy costs as well as costs for adverse events were not included as the company argued that there would be no meaningful difference between treatment groups.

A difference was made between the two belimumab formulations – IV and SC. The company proposes a patient access scheme for both formulations [redacted]. The IV formulation requires 14 infusions in year 1 and 13 in every subsequent year, with administration cost of £154 per infusion. Two different list prices for the SC formulation have been recorded. The SC list price model received by the ERG puts the list price at [redacted] per pen ([redacted] per four pen package). A recent update on the list price puts the list price at [redacted] per pen ([redacted] per four pen package). Further, the model assumed vial wastage and the cost for the IV population to be weight-dependent. In comparison, the company assumes only one hour of specialist nurse time to teach the patient how to self-administer the drug. As the formulation is not weight-dependent, no vial wastage is assumed. This results in drug-related costs of [redacted] for the first year and [redacted] for subsequent years for the IV formulation and [redacted] and [redacted] for the SC formulation (with the patient access scheme and administration cost included).

Disease activity related costs were based on an analysis conducted for the original submission which was in turn based on the belimumab Phase 2 trial. A linear regression analysis was used to relate the resource use to disease activity represented by the SS score. To define the disease activity maximum SS-score was used as a proxy for disease activity. The resulting costs which were based on 2005/2006

NHS reference costs were then inflated to 2018/2019 values using the consumer price index PSSRU from 2019. The inflated cost per SS-score can be found in Table 5.18.

Table 5.18: Disease activity related cost based on SS Score

SELENA-SLEDAI Score	Yearly Costs (2018/2019)
0	£1,294.53
1	£1,444.42
2	£1,594.30
3	£1,700.50
4	£1,762.95
5	£1,825.40
6	£1,887.86
7	£1,950.31
8	£2,012.76
9	£2,085.65
10	£2,168.92
11	£2,252.19
12	£2,335.46
13-20	£2,418.76
Source: CS[1] Table 70	

Cost of organ damage was calculated by multiplying the frequency of each medical condition by the cost incurred for the condition and adding it to a patient with the related organ damage. These costs were updated from the previous submission by applying new costs when they could be found through a targeted literature search and inflating them to 2018/2019 using the consumer price index PSSRU from 2019 when no new reference could be found.

ERG comment: a) disease activity was defined through the most severe point of disease activity over six months with uncertainty of whether this is the best measure b) resource use was taken from the previous submission and only adjusted for inflation.

a) To estimate the costs related to the SS score, the company defined severity of disease activity through the most severe point of disease activity over six months. Other measures such as the mean or median disease activity, or time-adjusted mean SLEDAI as in the JH cohort, are imaginable, plausible and to the best of the ERGs understanding could have been explored. The ERG considers that insufficient justification was provided for that choice.
b) To include related costs, the company only adjusted belimumab related care cost used in the previous company submission for inflation. The possibility of related cost changing beyond inflation through a change of belimumab-related care was investigated by asking a clarifying question to the company (CQ B22)[47] and adding a question about this in the inquiry to a clinical expert. The company responded saying that while they did not find evidence of belimumab-related care changing the inquiry was reasonable. The company therefore implemented a scenario analysis in which related care costs were doubled and halved to see what the possible influence on the ICER would be. With care costs doubled, the ICER would increase to £31,421

and with care costs halved the ICER would decrease to £30,139 per QALY gained. While the clinical expert responded that general SLE-related SoC had changed since the original submission, she gave no indication that belimumab-related care specifically had changed.[39] She did however note that updated cost were available.[96] As SoC costs are not included in this model, there would therefore be no change in the model. In summary, as there was no evidence for a change in belimumab-related care and the company's scenario analysis made clear that the impact on the ICER would be minor even in the case that belimumab-related care had changed, the issue was not pursued further.

6. COST EFFECTIVENESS RESULTS

6.1 Company's cost effectiveness results

Including the patient access scheme offered by the company belimumab SC compared to standard treatment in the HDA-2 population results in incremental costs of [redacted], [redacted] additional life year, [redacted] additional QALYs, resulting in an ICER of £30.566 per QALY gained. For the IV formulation compared to standard treatment, the economic evaluation results in incremental cost of [redacted], [redacted] additional life-years and [redacted] additional QALYs, resulting in an ICER of £30,001 per QALY gained. The ICERs for the HDA-1 population are £29,910 and £28,361 per additional QALY for the SC formulation and the IV formulation, respectively.

Table 6.1: HDA-2 population results for IV and SC belimumab formulations

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
IV model - [redacted]							
ST	£160,470	16.90	9.81				
Belimumab IV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£30,001
SC model - [redacted]							
ST	£151,999	17.12	10.06				
Belimumab SC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£30,566

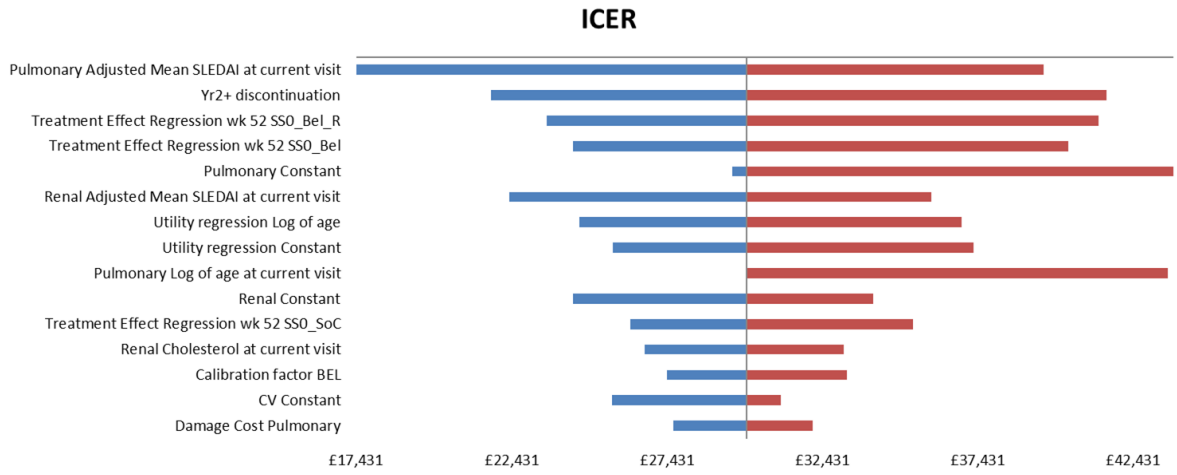
Source: CS[1] Table 74
All model outcomes presented are discounted.
Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years

6.2 Company's sensitivity analyses

The company performed deterministic, and probabilistic sensitivity analysis as well as scenario analyses. The probabilistic sensitivity analysis (1,000 iterations) varied the same set of model parameters to understand the influence of imprecision on the cost effectiveness results. This resulted in an ICER of £31,629 per QALY for the IV formulation and £29,264 per QALY for the SC formulation, with a probability of approximately 40% for the IV formulation to be cost effective and 50% for the SC formulation for the commonly cited £30,000 WTP-threshold.

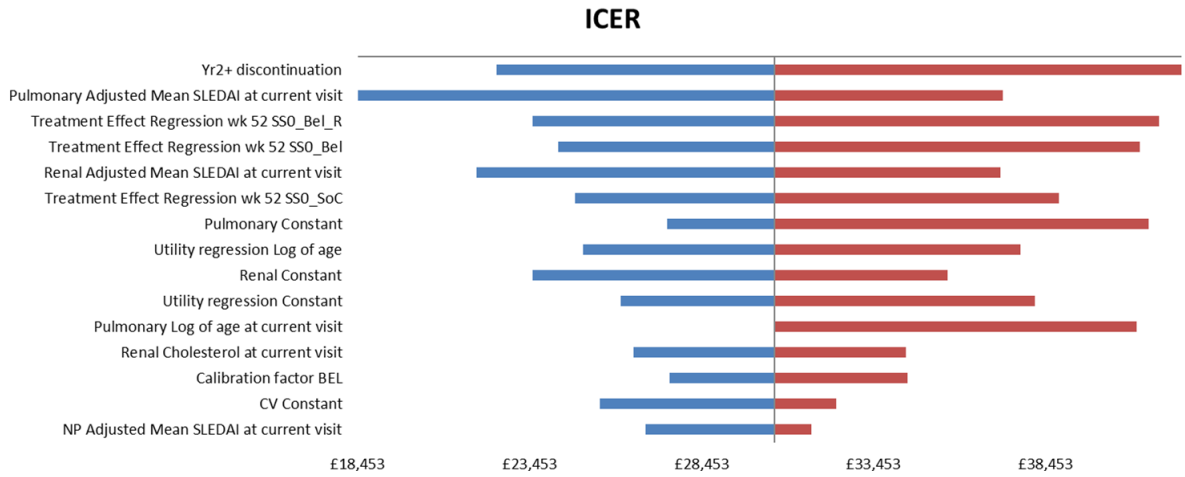
The deterministic sensitivity analysis was performed by varying individual parameters using a 95% confidence interval to see which had the greatest effect. For the IV model, the most influential parameters were "pulmonary adjusted mean SLEDAI at current visit", "Yr2+ discontinuation" and "Treatment Effect Regression wk52 SS0_BEL_R". For the SC formulation the most influential parameters were the "Yr2+ discontinuation", the "Pulmonary Adjusted Mean SLEDAI at current visit" and the "Treatment Effect Regression week 52 SS0_Bel_R". In both formulations in all three of the referenced DSA analyses the commonly cited WTP-threshold of £30,000 was far exceeded (Figures 6.1 and 6.2).

Figure 6.1: Tornado diagram ICER impact of selected factors belimumab IV formulation



Source: CS[1] Figure 22

Figure 6.2: Tornado diagram ICER impact of selected factors of belimumab SC formulation



Source: CS[1] Figure 25

Further, scenario analyses were conducted. All scenarios benefit the ICER of belimumab versus standard therapy, as the results range between £29,095 and £19,818 for the IV formulation and £24.396 and £20,241 for the SC formulation. The scenario which leads to the largest decrease of the ICER is applying differential discount rates of 1.5% for the benefits and 3.5% for the costs.

ERG comment: a) a larger number of iterations in the PSA would have been beneficial, and b) incorporation of uncertainty related to the calibration factor in the PSA.

a) The company conducted an analysis with 1,000 iterations. While higher numbers of iterations would have been beneficial, the ERG understands that this was difficult due to time and computational limitations.
b) It is unclear to the ERG in which way the company has reflected uncertainty around the calibration factor in its uncertainty assessment. While, in answer to clarification question B24[47] the company provided a table and highlighted the factors to indicate that it was included in the PSA, no distribution was indicated – it appears that uncertainty was calculated using an assumed standard error.

6.3 Comparison to analyses for TA397

The company, in response to clarification question B5b[47] helpfully provided an overview of impact of changes made to their model in the HDA-1 population, IV formulation (Table 6.2). The addition of the calibration factor to the belimumab treatment arm in the model had the largest impact on the ICER.

Table 6.2: Impact of reverting eight key parameter input updates on IV model for the HDA-1 patient subgroup provided as part of the current submission, to values used in final base case of TA397

Parameters	Base Case Value In Current Model	Previous Value applied from TA397	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
IV HDA-1 model (base)	-	-	[redacted]	[redacted]	[redacted]	£28,362
Calibration factors	Added	None	[redacted]	[redacted]	[redacted]	£48,604
Annual Long-Term Organ Damage costs	Updated literature search	Previous literature search	[redacted]	[redacted]	[redacted]	£37,219
SLEDAI Annual costs	HCHS inflation 2018/19 values	2009/10	[redacted]	[redacted]	[redacted]	£28,307
Patient mortality table	ONS 2016-18 values	ONS 2007-09 values	[redacted]	[redacted]	[redacted]	£28,235
Organ damage disutilities	Updated literature search	Previous literature search	[redacted]	[redacted]	[redacted]	£28,007
Patient weight source	BILAG Registry	Trial	[redacted]	[redacted]	[redacted]	£25,984
Natural discontinuation rates	[redacted]	Y1 8.0% / Y2 11.7%	[redacted]	[redacted]	[redacted]	£23,521
Drug cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Source: Response to clarification question B5b[47] Table 11						

6.4 Model validation and face validity check

The company has performed several exercises in an effort to validate their model: model convergence checks were undertaken to assess stability of results, long-term outcome predictions were internally and externally validated, and internal validation was performed by model developers and independent health economists. The latter did not highlight any important errors.

ERG comment: a) insufficient model validation, b) internal validation of organ damage estimates, c) external validation of organ damage estimates), d) modelling of responders not in line with response rule in clinical practice, and e) noise introduced by same patient following different trajectories based on allocation to treatment arm.

The ERG remains concerned that no full validation exercise was provided by the company. In response to clarification questions, the company provided further validation exercises, that helped the ERG in identifying some concerns with model validity.

a) The company did not provide a validation of their entire model as requested in clarification question B4b, arguing that their model remained identical to their previous model, with the exception of calibration factors (longer term evidence from the Toronto cohort). This representation is not quite accurate, as more changes have been implemented than just the addition of calibration factors. Table 11 in the company's response to clarification questions shows an overview of these changes.[47] The ERG would have liked to see this validation exercise, which can help in establishing that the model performs as expected.
b) Internal validation of predicted versus observed organ damage at years one, five and 10 in the JH cohort (response to B4a).[47] Overall, the ERG considers that the model performs as expected, predicting organ damage as observed in the JH cohort with relatively high accuracy.
c) External validation of predicted versus observed organ damage at years one, five and 10 in the Toronto cohort (response to B4a).[47] Some differences are highlighted in pulmonary malignancy and CNS damage, which tend to be over predicted compared to the Toronto cohort, and skin damage, which is slightly under predicted.
d) The ERG questions the plausibility of patient profiles and trajectories over time. In response to clarification question B4c),[47] the company provided a model file that enables the creation of patient logs. These show the baseline patient characteristics of each simulated patient, whether they are a responder and what their disease activity levels and organ damage are at baseline and over time. There are questions about disease activity SS scores: there appears to be a contradiction in that patients that have been classified as non-responders can have a reduction of ≥ 4 points in SS score from year 0 to 1. For example patient 5 has a reduction from a SS score of 16 in year 0 to 10 in year 1, despite being classed as non-responder (and assumed to discontinue treatment at 24 weeks). Similar reductions in scores (≥4) occur for patients 8 and 9; only patient 10 (out of the first 10 simulated patients) was classed as non-responder and had a reduction of <4 at year 1 (in fact it was 3.8). Essentially, this may mean that in the model patients discontinue belimumab treatment even though they were eligible for continuation. Since SS scores are informing HRQoL, this mis-match between SS scores and treatment continuation has the potential to bias model outcomes. This mis-match may be owed to the fact that the primary outcome of response was assessed at 52 weeks in the BLISS trials and that the probabilities of response at 24 weeks is, in fact, not based on this response criterion. For completeness, it should be mentioned that SS score reductions are identical for those modelled patients that are classed as non-responders irrespective of treatment allocated to them. In conclusion, HRQoL may be over-estimated in the model overall or proportions of responders

and therefore belimumab-associated costs may be under-estimated, and the ERG therefore considers that the model results are to be interpreted with extreme caution. In the ERG's view, whether a patient is classed as a responder in the model should depend on their SS score reduction.

e) KEY ISSUE. An additional issue is that the same simulated patient appears to follow entirely different trajectories in terms of organ damage and death, when the only difference should be the allocation to treatment and organ damage and death caused by that allocation: e.g. the same patient can experience damage in completely different organs when all that differs is treatment allocation. For example, looking at the company's submitted file, patient 18 has a SLICC score of 2 at baseline, with SLICC items affected musculoskeletal and skin. When treated with belimumab, organ damage occurs in: musculoskeletal (year 0), neuropsychiatric (year 8), peripheral vascular (year 0), and skin (year 0). When treated with ST, cardiovascular (year 3), musculoskeletal (year 0), ocular (year 5), pulmonary (year 11), and skin (year 0). It is not clear why the same patient develops damage to different organs, when the only difference is treatment. So, at the end of year 0, a patient treated with belimumab + ST has organ damage in peripheral vascular that did not occur for the same patient treated with ST alone. This suggests that, unless belimumab caused the damage in peripheral vascular, the way the company models trajectories includes noise, which may make it difficult to assess the benefit of belimumab over belimumab + ST. Another example is that patient 8 receives treatment with belimumab or ST for 32 years in the belimumab arm, but for 62 years in the ST arm. Presumably, treatment stops because of death (unfortunately death is not reported in these patient logs). Patient 5 also appears to die much sooner in the belimumab than in the ST arm (year 2 versus 19). The ERG suggests to amend the model in a way that enables identical patients allocated to different treatment arms to follow exactly identical trajectories, with the exceptions of treatment-related changes (i.e. disease activity, and organ damage). The impact of the company's way of modelling is probably limited to the creation of noise and makes validation more difficult – which means that the ERG is uncertain whether the model indeed performs as expected.

Table 6.3: Key issue 13: Sampling of organ damage and death occurs after allocation to treatment

Report section	Section 6.3
Description of issue and why the ERG has identified it as important	In the VBA, first, a simulated patient is allocated to a treatment and organ damage and death are only sampled within the treatment arm. This leads to the same simulated patient (same age, gender, SS score) experiencing differential organ damage and times of death only dependent on allocation to treatment arm but not caused by this allocation (so just because of sampling). This induces noise in the model and makes validation difficult.
What alternative approach has the ERG suggested?	A structural model adjustment in which organ damage items involved and death are sampled from before treatment allocation
What is the expected effect on the cost effectiveness estimates?	Unknown. May be no effect if all this did was induce noise.
What additional evidence or analyses might help to resolve this key issue?	A structural model adjustment in which organ damage items involved and death are sampled from before treatment allocation

7. EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

7.1 Exploratory and sensitivity analyses undertaken by the ERG

Based on all considerations in Section 5.2 (summarised in Table 7.1), 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):[97]

- Fixing errors (correcting the model where the company's submitted model was unequivocally wrong)
- Fixing violations (correcting the model where the ERG considered that the NICE reference case, scope or best practice had not been adhered to)
- Matters of judgement (amending the model where the ERG considers that reasonable alternative assumptions are preferred)

These model amendments were made in both models for the SC and IV treatment formulation.

7.1.1 Fixing errors

1. First year corrected reductions in SS score for belimumab non-responders.

7.1.2 Fixing violations

NA

7.1.3 Matters of judgement

2. Addition of calibration factor was not viewed as appropriate by ERG (Section 5.2.6). The ERG removed the calibration factor in its base-case.

Results are presented in Tables 7.2 and 7.3.

7.1.4 Additional sensitivity analyses

Additional sensitivity analyses were performed to examine the potential impact of alternative assumptions on the cost effectiveness estimates. Exploratory analyses conditional on the ERG base-case included (again presented for both SC and IV treatment formulations):

1. Use unadjusted JH model
2. Use company's calibration factor
3. Use calibration factor on both treatment arms
4. Remove impact of organ damage on utility estimation
5. Use patient weight based on trial rather than BILAG-BR
6. HDA-1 subgroup

Results are presented in Tables 7.4 and 7.5.

Table 7.1: Overview of key issues related to the cost effectiveness (conditional on fixing errors highlighted in Section 5.1)

Key issue	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in ERG base-case[b]	Required additional evidence or analyses
		Transparency, Methods, Imprecision, bias & indirectness, unavailability			*Yes/No/Partly/Explored*	
1) Rituximab + standard therapy was excluded as comparator.	5.2.4	Unavailability	Difficult due to lack of data	+/-	No	May not be possible
2) IV and SC formulations are not compared with each other, as two separate model files are provided.	5.2.6	Methods	Include both formulations in one model	+/-	No	
3) Use of calibration factor	5.2.6	Unavailability of long-term comparative data, bias & indirectness due to differences in patient population between JH and BLISS, and bias in BLISS long-term follow-up	Not using the calibration factor	+, when removed ICER increased by approximately £21,000 and £28,000 for IV and SC formulation	Yes	Uncertainty about long-term effectiveness is unlikely to be resolved with evidence available

Key issue	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in ERG base-case[b]	Required additional evidence or analyses
4) Implementation of 24-week response and treatment continuation in the model	5.2.6	Methods of modelling	Structural model adjustment	+	No	Not applicable
5) Error in calculation of belimumab non-responder disease activity at 52 weeks	5.2.6	Methods	Fix error	+, when error was fixed the ICER increased by approximately £1,700 and £2,000 (IV and SC formulations)	Yes	
6) Violation in utility estimation	5.2.8	Methods: it is incorrect to remove co-variates without re-estimating the model	Re-estimate coefficients when organ damage covariates are excluded	+/-, likely minor impact	No	The company could re-estimate utility model
7) Uncertainty about organ damage utility multipliers	5.2.8	Bias & indirectness as utility model may include organ damage to certain extent and there	Exclude impact of organ damage on utilities	- conditional on ERG base-case (when removed ICERs decreased by approximately	Explored, only in scenario as impact of organ damage on HRQoL uncertain	Consult expert opinion and further literature on weights of organ

Key issue	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in ERG base-case[b]	Required additional evidence or analyses
		may be double-counting		£5,000 and £4,000 for IV and SC), + conditional on company's base-case		damage incidence and utility multipliers
8) Sampling of organ damage and death occurs after allocation to treatment	6.3	Methods	Model adjustment	+/-, may not have an impact	No	

[a] Likely conservative assumptions (of the intervention versus all comparators) are indicated by '-'; while '+/-' indicates that the bias introduced by the issue is unclear to the ERG and '+' indicates that the ERG believes this issue likely induces bias in favour of the intervention versus at least one comparator; [b] Explored
ERG = Evidence Review Group; FE = Fixing errors; FV = fixing violations; MJ = matters of judgement; ICER = incremental cost effectiveness ratio

7.2 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

In Section 7.1 the ERG base-case assumptions were presented. Tables 7.2 and 7.3 show how individual changes impact the results plus the combined effect of all changes simultaneously for the IV and SC formulations respectively. Table 7.4 shows probabilistic results for both formulations. The exploratory scenario analyses are presented in Tables 7.5 and 7.6 for the IV and SC formulations respectively. These are all conditional on the ERG base-case. The submitted model file contains technical details on the analyses performed by the ERG, with pointers to what changes were made in the Visual Basic.

7.2.1 ERG base-case

Table 7.2: Deterministic ERG base-case for the IV formulation (HDA-2 subgroup, PAS price)

Technologies	Total costs	Total LYs	Total QALYs	Incremental costs	Incremental QALYs	ICER (/QALY)
CS base-case						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£30.000
Standard therapy	£160.470	16.900	9.809			
Fixing errors 1: 1st year: SS reduction for belimumab non-responders						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£31,695
Standard therapy	£160,470	16.900	9.809			
Matter of judgement 2: Calibration factor removed conditional on FE1 = ERG base-case						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£52,891
Standard therapy	£160,470	16.900	9.809			

Table 7.3: Deterministic ERG base-case for the SC formulation (HDA-2 subgroup, PAS price)

Technologies	Total costs	Total LYs	Total QALYs	Incremental costs	Incremental QALYs	ICER (/QALY)
CS base-case						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£30,566
Standard therapy	£151,999	17.122	10.056			
Fixing errors 1: 1st year: SS reduction for belimumab non-responders						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£32,617
Standard therapy	£151,999	17.122	10.056			
Matter of judgement 2: Calibration factor removed conditional on FE1 = ERG base-case						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£61,057
Standard therapy	£151,999	17.122	10.056			

Table 7.4: Probabilistic ERG base-case (HDA-2 subgroup, PAS price)

Technologies	Incremental costs	Incremental QALYs	ICER (/QALY)
IV formulation			
Belimumab	[redacted]	[redacted]	£56,894
Standard therapy			
SC formulation			
Belimumab	[redacted]	[redacted]	£62,367
Standard therapy			

7.2.2 ERG scenarios

Table 7.5: ERG scenarios for the IV formulation (HDA-2 subgroup, PAS price)

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (/QALY)
ERG base-case					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£52,891
Standard therapy	£160,470	9.809			
Scenario 1: Use unadjusted JH model					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£63,951
Standard therapy	£161,467	10.798			
Scenario 2: Use calibration factor					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£31,695
Standard therapy	£160,470	9.809			
Scenario 3: Use calibration factors on both arms					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£24,847
Standard therapy	£167,261	9.669			
Scenario 4: Remove impact of organ damage					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£48,347
Standard therapy	£160,470	11.941			
Scenario 5: Patient weight based on trial					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£50,451
Standard therapy	£160,470	9.809			
Scenario 6: HDA-1 subgroup					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£48,849
Standard therapy	£166,658	10.216			
Scenario 7: HDA-1 subgroup conditional on company's base-case with FE1					

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (/QALY)
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£28,265
Standard therapy	£166,658	10.216			

Table 7.6: ERG scenarios for the SC formulation (HDA-2 subgroup, PAS price)

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (/QALY)
ERG base-case					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£61,057
Standard therapy	£151,999	10.056			
Scenario 1: Use unadjusted JH model					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£68,909
Standard therapy	£151,873	11.036			
Scenario 2: Use calibration factors					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£32,617
Standard therapy	£151,999	10.056			
Scenario 3: Use calibration factors on both arms					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£25,418
Standard therapy	£158,791	9.916			
Scenario 4: Remove impact of organ damage					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£56,901
Standard therapy	£151,999	12.082			
Scenario 6: HDA-1 subgroup					
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£60,241
Standard therapy	£156,692	10.476			
Scenario 7: HDA-1 subgroup conditional on company's base-case with FE1					

Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	£31,706
Standard therapy	£156,692	10.476			

7.3 Conclusions of the cost effectiveness section

The company's health economic model mostly addresses the scope, with the exception of comparators included (Issue 1, as detailed and numbered in Table 7.1). The company have provided justification for excluding those comparators and the ERG agreed that it would have been challenging to model the comparison with rituximab and that cyclophosphamide may not be an appropriate comparator. The company's cost effectiveness estimates rest crucially on assumptions surrounding long-term treatment effectiveness and impact on organ damage. The resulting uncertainty was not resolved with additional evidence and modelling. Crucially, the ERG does not consider the application of the calibration factor to be appropriate (Issue 3). The ERG identified issues in the modelling that may result in bias in the company's cost effectiveness estimates and that can be resolved by adjustments to the model; Issue 2) IV and SC formulations were not included in the same model, thus hampering a direct comparison between the two formulation; Issue 4) an inconsistency in the implementation of 24-week response and treatment continuation in the model; Issue 5) an error in the calculation of belimumab non-responder disease activity at 52 weeks; Issue 6) a violation in the estimation of utilities; and Issue 8) a violation in sampling of patient events after they are allocated to treatment. Further uncertainty remains around the impact of organ damage on patients' HRQoL and the way it has been included in the modelling (Issue 7). This could potentially be partially addressed by exploring expert opinion and the literature on the weights of organ damage items and utility multipliers used.

In conclusion, with the current model, cost effectiveness estimates of belimumab compared with standard therapy are uncertain and likely contain bias. Even when the modelling issues are addressed, substantial uncertainty remains about long-term treatment effectiveness of belimumab. Some uncertainty also remains over whether the impact of the disease on HRQoL is accurately captured by the model.

8. REFERENCES

[1] GlaxoSmithKline. *Belimumab for treating active autoantibody-positive systemic lupus erythematosus [ID1591]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Document B - Company evidence submission*: GlaxoSmithKline, 2020. 218p.

[2] Manson JJ, Rahman A. Systemic lupus erythematosus. *Orphanet J Rare Dis* 2006;1:6.

[3] Gergianaki I, Bortoluzzi A, Bertsias G. Update on the epidemiology, risk factors, and disease outcomes of systemic lupus erythematosus. *Best Pract Res Clin Rheumatol* 2018;32(2):188-205.

[4] Morgan C, Bland AR, Maker C, Dunnage J, Bruce IN. Individuals living with lupus: findings from the LUPUS UK Members Survey 2014. *Lupus* 2018;27(4):681-687.

[5] Rees F, Doherty M, Grainge M, Davenport G, Lanyon P, Zhang W. The incidence and prevalence of systemic lupus erythematosus in the UK, 1999-2012. *Ann Rheum Dis* 2016;75(1):136-41.

[6] Johnson AE, Gordon C, Palmer RG, Bacon PA. The prevalence and incidence of systemic lupus erythematosus in Birmingham, England. Relationship to ethnicity and country of birth. *Arthritis Rheum* 1995;38(4):551-8.

[7] Somers EC, Thomas SL, Smeeth L, Schoonen WM, Hall AJ. Incidence of systemic lupus erythematosus in the United Kingdom, 1990-1999. *Arthritis Rheum* 2007;57(4):612-8.

[8] Yee CS, Su L, Toescu V, Hickman R, Situnayake D, Bowman S, et al. Birmingham SLE cohort: outcomes of a large inception cohort followed for up to 21 years. *Rheumatology (Oxford)* 2015;54(5):836-43.

[9] Rahman A, Isenberg DA. Systemic lupus erythematosus. *N Engl J Med* 2008;358(9):929-39.

[10] Lam NC, Ghetu MV, Bieniek ML. Systemic lupus erythematosus: primary care approach to diagnosis and management. *Am Fam Physician* 2016;94(4):284-94.

[11] Petri M, Orbai AM, Alarcón GS, Gordon C, Merrill JT, Fortin PR, et al. Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. *Arthritis Rheum* 2012;64(8):2677-86.

[12] Aringer M. EULAR/ACR classification criteria for SLE. *Semin Arthritis Rheum* 2019;49(3S):S14-S17.

[13] Fava A, Petri M. Systemic lupus erythematosus: diagnosis and clinical management. *J Autoimmun* 2019;96:1-13.

[14] Kaul A, Gordon C, Crow MK, Touma Z, Urowitz MB, van Vollenhoven R, et al. Systemic lupus erythematosus. *Nat Rev Dis Primers* 2016;2:16039.

[15] Farinha F, Freitas F, Águeda A, Cunha I, Barcelos A. Concerns of patients with systemic lupus erythematosus and adherence to therapy - a qualitative study. *Patient Prefer Adherence* 2017;11:1213-1219.

[16] Lau CS, Mak A. The socioeconomic burden of SLE. *Nat Rev Rheumatol* 2009;5(7):400-4.

[17] Yelin E, Trupin L, Katz P, Criswell L, Yazdany J, Gillis J, et al. Work dynamics among persons with systemic lupus erythematosus. *Arthritis Rheum* 2007;57(1):56-63.

[18] Ugarte-Gil MF, Acevedo-Vásquez E, Alarcón GS, Pastor-Asurza CA, Alfaro-Lozano JL, Cucho-Venegas JM, et al. The number of flares patients experience impacts on damage accrual in systemic lupus erythematosus: data from a multiethnic Latin American cohort. *Ann Rheum Dis* 2015;74(6):1019-23.

[19] Gladman DD, Urowitz MB, Rahman P, Ibañez D, Tam LS. Accrual of organ damage over time in patients with systemic lupus erythematosus. *J Rheumatol* 2003;30(9):1955-9.

[20] Lopez R, Davidson JE, Beeby MD, Egger PJ, Isenberg DA. Lupus disease activity and the risk of subsequent organ damage and mortality in a large lupus cohort. *Rheumatology (Oxford)* 2012;51(3):491-8.

[21] Gladman D, Ginzler E, Goldsmith C, Fortin P, Liang M, Urowitz M, et al. The development and initial validation of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology damage index for systemic lupus erythematosus. *Arthritis Rheum* 1996;39(3):363-9.

[22] Guidelines for referral and management of systemic lupus erythematosus in adults. American College of Rheumatology Ad Hoc Committee on Systemic Lupus Erythematosus guidelines. *Arthritis Rheum* 1999;42(9):1785-96.

[23] Segura BT, Bernstein BS, McDonnell T, Wincup C, V MR, Giles I, et al. Damage accrual and mortality over long-term follow-up in 300 patients with systemic lupus erythematosus in a multi-ethnic British cohort. *Rheumatology (Oxford)* 2020;59(3):524-533.

[24] Swaak AJ, van den Brink HG, Smeenk RJ, Manger K, Kalden JR, Tosi S, et al. Systemic lupus erythematosus: clinical features in patients with a disease duration of over 10 years, first evaluation. *Rheumatology (Oxford)* 1999;38(10):953-8.

[25] Cervera R, Khamashta MA, Font J, Sebastiani GD, Gil A, Lavilla P, et al. Morbidity and mortality in systemic lupus erythematosus during a 10-year period: a comparison of early and late manifestations in a cohort of 1,000 patients. *Medicine (Baltim)* 2003;82(5):299-308.

[26] Goldblatt F, Chambers S, Rahman A, Isenberg DA. Serious infections in British patients with systemic lupus erythematosus: hospitalisations and mortality. *Lupus* 2009;18(8):682-9.

[27] Campbell R, Jr., Cooper GS, Gilkeson GS. Two aspects of the clinical and humanistic burden of systemic lupus erythematosus: mortality risk and quality of life early in the course of disease. *Arthritis Rheum* 2008;59(4):458-64.

[28] Roman MJ, Shanker BA, Davis A, Lockshin MD, Sammaritano L, Simantov R, et al. Prevalence and correlates of accelerated atherosclerosis in systemic lupus erythematosus. *N Engl J Med* 2003;349(25):2399-406.

[29] Urowitz MB, Gladman DD, Tom BD, Ibañez D, Farewell VT. Changing patterns in mortality and disease outcomes for patients with systemic lupus erythematosus. *J Rheumatol* 2008;35(11):2152-8.

[30] Bernatsky S, Boivin JF, Joseph L, Rajan R, Zoma A, Manzi S, et al. An international cohort study of cancer in systemic lupus erythematosus. *Arthritis Rheum* 2005;52(5):1481-90.

[31] Parikh-Patel A, White RH, Allen M, Cress R. Cancer risk in a cohort of patients with systemic lupus erythematosus (SLE) in California. *Cancer Causes Control* 2008;19(8):887-94.

[32] Bernatsky S, Ramsey-Goldman R, Isenberg D, Rahman A, Dooley MA, Sibley J, et al. Hodgkin's lymphoma in systemic lupus erythematosus. *Rheumatology (Oxford)* 2007;46(5):830-2.

[33] Bernatsky S, Boivin JF, Joseph L, Manzi S, Ginzler E, Gladman DD, et al. Mortality in systemic lupus erythematosus. *Arthritis Rheum* 2006;54(8):2550-7.

[34] Alarcón GS, McGwin G, Jr., Bastian HM, Roseman J, Lisse J, Fessler BJ, et al. Systemic lupus erythematosus in three ethnic groups. VII [correction of VIII]. Predictors of early mortality in the LUMINA cohort. LUMINA Study Group. *Arthritis Rheum* 2001;45(2):191-202.

[35] Molad Y, Borkowski T, Monselise A, Ben-Haroush A, Sulkes J, Hod M, et al. Maternal and fetal outcome of lupus pregnancy: a prospective study of 29 pregnancies. *Lupus* 2005;14(2):145-51.

[36] Rees F, Doherty M, Grainge MJ, Lanyon P, Davenport G, Zhang W. Mortality in systemic lupus erythematosus in the United Kingdom 1999-2012. *Rheumatology (Oxford)* 2016;55(5):854-60.

[37] Pearce FA, Rutter M, Sandhu R, Batten RL, Garner R, Little J, et al. BSR guideline on the management of adults with systemic lupus erythematosus (SLE) 2018: baseline multi-centre audit in the UK. *Rheumatology (Oxford)* 2020;Epub ahead of print.

[38] Gladman DD, Urowitz MB. Prognosis, mortality and morbidity in systemic lupus erythematosus. In: Wallace DJ, Hahn BH, editors. *Dubois' lupus erythematosus*. 7th ed. Philadelphia: Lippincott Williams & Wilkins, 2007: 1333-53.

[39] Gordon C, Amissah-Arthur MB, Gayed M, Brown S, Bruce IN, D'Cruz D, et al. The British Society for Rheumatology guideline for the management of systemic lupus erythematosus in adults. *Rheumatology (Oxford)* 2018;57(1):e1-e45.

[40] Gordon C, Isenberg D, Lerstrøm K, Norton Y, Nikaï E, Pushparajah DS, et al. The substantial burden of systemic lupus erythematosus on the productivity and careers of patients: a European patient-driven online survey. *Rheumatology (Oxford)* 2013;52(12):2292-301.

[41] Baker K, Pope J. Employment and work disability in systemic lupus erythematosus: a systematic review. *Rheumatology (Oxford)* 2009;48(3):281-4.

[42] Kent T, Davidson A, Newman D, Buck G, D'Cruz D. Burden of illness in systemic lupus erythematosus: results from a UK patient and carer online survey. *Lupus* 2017;26(10):1095-1100.

[43] Amissah-Arthur MB, Gordon C. Contemporary treatment of systemic lupus erythematosus: an update for clinicians. *Ther Adv Chronic Dis* 2010;1(4):163-75.

[44] National Institute for Health and Care Excellence. *Belimumab for treating active autoantibody-positive systemic lupus erythematosus [TA397] [Internet]*. London: NICE, 2016 [accessed 9.12.20] Available from: https://www.nice.org.uk/guidance/ta397

[45] National Institute for Health and Care Excellence. *Systemic lupus erythematosus: oral mycophenolate. Evidence summary [ESUOM36] [Internet]*. London: NICE, 2014 [accessed 9.12.20] Available from: https://www.nice.org.uk/advice/esuom36/chapter/Key-points-from-the-evidence

[46] Fanouriakis A, Kostopoulou M, Alunno A, Aringer M, Bajema I, Boletis JN, et al. 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. *Ann Rheum Dis* 2019;78(6):736-745.

[47] GlaxoSmithKline. *Belimumab for treating active autoantibody-positive systemic lupus erythematosus [ID1591] – Response to request for clarification from the ERG*: GlaxoSmithKline, 2020. 96p.

[48] Kalunian K, Merrill JT. New directions in the treatment of systemic lupus erythematosus. *Curr Med Res Opin* 2009;25(6):1501-14.

[49] National Institute for Health and Care Excellence. *Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (review of TA397) [ID1591]: Final scope*. London: NICE, 2020. 5p.

[50] GlaxoSmithKline. *Benlysta (belimumab). Summary of product characteristics [Internet]*, 2020 [accessed 9.12.20] Available from: https://www.ema.europa.eu/en/documents/product-information/benlysta-epar-product-information_en.pdf

[51] Furie R, Rovin BH, Houssiau F, Malvar A, Teng YKO, Contreras G, et al. Two-year, randomized, controlled trial of belimumab in lupus nephritis. *N Engl J Med* 2020;383(12):1117-1128.

[52] National Institute for Health and Care Excellence. *Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus: FAD committee papers [ID416/TA397] [Internet]*. London: NICE, 2013 [accessed 9.12.20] Available from: https://www.nice.org.uk/guidance/ta397/documents/committee-papers

[53] van Vollenhoven R, Voskuyl A, Bertsias G, Aranow C, Aringer M, Arnaud L, et al. A framework for remission in SLE: consensus findings from a large international task force on definitions of remission in SLE (DORIS). *Ann Rheum Dis* 2017;76(3):554-561.

[54] Franklyn K, Lau CS, Navarra SV, Louthrenoo W, Lateef A, Hamijoyo L, et al. Definition and initial validation of a Lupus Low Disease Activity State (LLDAS). *Ann Rheum Dis* 2016;75(9):1615-21.

[55] Polachek A, Gladman DD, Su J, Urowitz MB. Defining low disease activity in systemic lupus erythematosus. *Arthritis Care Res (Hoboken)* 2017;69(7):997-1003.

[56] National Institute for Health and Care Excellence. *Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (Review of TA397) [ID1591] [Internet]*. London: NICE, [accessed 12.01.21] Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10626/documents

[57] GlaxoSmithKline. *Belimumab for treating active autoantibody-positive systemic lupus erythematosus [ID1591]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix D. Identification, selection and synthesis of clinical evidence*: GlaxoSmithKline, 2020. 23p.

[58] SIGN. Search filters [Internet]. [accessed 08.12.20]. Available from: https://www.sign.ac.uk/what-we-do/methodology/search-filters/

[59] 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

[60] Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health* 1998;52(6):377-84.

[61] 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/

[62] GlaxoSmithKline. *A phase 3, multi-center, randomized, double-blind, placebo-controlled, 52-week study to evaluate the efficacy and safety of belimumab (HGS1006, LymphoStat-B™), a fully human monoclonal anti-BLyS antibody, in subjects with systemic lupus erythematosus (SLE). Data on file. 21 January 2010.*

[63] GlaxoSmithKline. *A phase 3, multi-center, randomized, double-blind, placebo-controlled, 76-week study to evaluate the efficacy and safety of belimumab (HGS1006, LymphoStat-B™), a Fully human monoclonal anti-BLyS antibody, in subjects with systemic lupus erythematosus (SLE). Data on file. 28 September 2010*

[64] GlaxoSmithKline. *Clinical Study Report: A phase 3, multi-center, randomized, double-blind, placebo-controlled, 52-week study to evaluate the efficacy and safety of belimumab (HGS1006) administered subcutaneously (SC) to subjects with systemic lupus erythematosus (SLE) - double-blind endpoint analysis. Data on file. 8 March 2016*

[65] GlaxoSmithKline. *A multi-center, continuation trial of belimumab (HGS1006, LymphoStat-B), a fully human monoclonal anti-BLyS antibody, in subjects with systemic lupus erythematosus (SLE) who completed the phase 3 protocol HGS1006-C1056 or HGS1006-C1057. Data on file. 31 July 2017.*

[66] GlaxoSmithKline. *A multi-center, continuation trial of belimumab (HGS1006, LymphoStat-B), a fully human monoclonal anti-BLyS antibody, in subjects with systemic lupus erythematosus (SLE) who completed the phase 3 protocol HGS1006-C1056 in the United States. Data on file. 2 December 2015.*

[67] GlaxoSmithKline. *A phase 3, multi-center, randomized, double-blind, placebo controlled, 52-week study to evaluate the efficacy and safety of belimumab (HGS1006) administered subcutaneously (SC) to subjects with systemic lupus erythematosus (SLE) - open-label phase - end of study report. Data on File. 26 September 2016*

[68] Urowitz MB, Ohsfeldt RL, Wielage RC, Kelton KA, Asukai Y, Ramachandran S. Organ damage in patients treated with belimumab versus standard of care: a propensity score-matched comparative analysis. *Ann Rheum Dis* 2019;78(3):372-379.

[69] GlaxoSmithKline. *An open-label single-arm study to evaluate the reliability of an autoinjector that administers belimumab subcutaneously in subjects with systemic lupus erythematosus (SLE). 23 October 2015.*

[70] Wallace DJ, Stohl W, Furie RA, Lisse JR, McKay JD, Merrill JT, et al. A phase II, randomized, double-blind, placebo-controlled, dose-ranging study of belimumab in patients with active systemic lupus erythematosus. *Arthritis Rheum* 2009;61(9):1168-78.

[71] Wallace DJ, Ginzler EM, Merrill JT, Furie RA, Stohl W, Chatham WW, et al. Safety and efficacy of belimumab plus standard therapy for up to thirteen years in patients with systemic lupus erythematosus. *Arthritis Rheumatol* 2019;71(7):1125-1134.

[72] Sheikh S, Scheinberg M, Cheng-Chung Wei J, et al. LB0012 Headline results for a phase 4, 52-week, randomised, double-blind, placebo-controlled study to assess adverse events of special interest (AESI) in adults with active, autoantibody-positive systemic lupus erythematosus (SLE) receiving belimumab. *Ann Rheum Dis* 2019;78(Suppl 2):266.

[73] Bae S, Dimelow R, Ji B, et al. Results of the open-label, non-randomized 52-week study to evaluate treatment holidays and rebound phenomenon after treatment with belimumab in patients with SLE. *Arthritis Rheumatol* 2019;71(suppl 10).

[74] Sergio Guedes Barbosa L, Saxena A, Ginzler E, et al. Efficacy of belimumab in patients of black race with systemic lupus erythematosus and high disease activity or renal manifestations. *Arthritis Rheumatol* 2019;71(suppl 10).

[75] Zhang F, Bae SC, Bass D, Chu M, Egginton S, Gordon D, et al. A pivotal phase III, randomised, placebo-controlled study of belimumab in patients with systemic lupus erythematosus located in China, Japan and South Korea. *Ann Rheum Dis* 2018;77(3):355-363.

[76] Brunner HI, Abud-Mendoza C, Viola DO, Calvo Penades I, Levy D, Anton J, et al. Safety and efficacy of intravenous belimumab in children with systemic lupus erythematosus: results from a randomised, placebo-controlled trial. *Ann Rheum Dis* 2020;79(10):1340-1348.

[77] Bruce IN, British Isles Lupus Assessment Group. *BILAG Biologics Prospective Cohort: the use of novel biological therapies in the treatment of systemic lupus erythematosus (SLE). BILAG BR Protocol Version 1.4 07/03/2011. Data on file*

[78] GlaxoSmithKline. *Reporting and analysis plan (RAP) for the UK Benlysta prospective, observational BILAG British Registry (BILAG-BR) sub-study. Data on file. 6 April 2020*

[79] Collins CE, Dall'Era M, Kan H, Macahilig C, Molta C, Koscielny V, et al. Response to belimumab among patients with systemic lupus erythematosus in clinical practice settings: 24-month results from the OBSErve study in the USA. *Lupus Sci Med* 2016;3(1):e000118.

[80] GlaxoSmithKline. *Pooled analysis of the real-world effectiveness of belimumab using multi-country data from the OBSErve studies. 19 March 2020.*

[81] Ramachandran S, Parks D, Kurtinecz M, Roth DA, Alfonso-Cristancho R. An indirect comparison of intravenous and subcutaneous belimumab efficacy in patients with SLE and high disease activity. *J Comp Eff Res* 2018;7(6):581-593.

[82] GlaxoSmithKline. *Belimumab for treating active autoantibody-positive systemic lupus erythematosus [ID1591]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix M: Further methodology and results of belimumab LTE studies*: GlaxoSmithKline, 2020. 20p.

[83] GlaxoSmithKline. *Belimumab for treating active autoantibody-positive systemic lupus erythematosus [ID1591]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix L: Further methodology and results of belimumab RCTs*: GlaxoSmithKline, 2020. 49p.

[84] GlaxoSmithKline. *Belimumab for treating active autoantibody-positive systemic lupus erythematosus [ID1591]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix O: Additional studies: further details*: GlaxoSmithKline, 2020. 16p.

[85] GlaxoSmithKline. *Belimumab for treating active autoantibody-positive systemic lupus erythematosus [ID1591]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix F. Adverse events*: GlaxoSmithKline, 2020. 13p.

[86] Urowitz MB, Ohsfeldt RL, Wielage RC, Dever JJ, Zakerifar M, Asukai Y, et al. Comparative analysis of long-term organ damage in patients with systemic lupus erythematosus using belimumab versus standard therapy: a post hoc longitudinal study. *Lupus Sci Med* 2020;7(1).

[87] van Vollenhoven RF, Navarra SV, Levy RA, Thomas M, Heath A, Lustine T, et al. Long-term safety and limited organ damage in patients with systemic lupus erythematosus treated with belimumab: a phase III study extension. *Rheumatology (Oxford)* 2020;59(2):281-291.

[88] Furie R, Petri M, Zamani O, Cervera R, Wallace DJ, Tegzová D, et al. A phase III, randomized, placebo-controlled study of belimumab, a monoclonal antibody that inhibits B lymphocyte stimulator, in patients with systemic lupus erythematosus. *Arthritis Rheum* 2011;63(12):3918-30.

[89] Navarra SV, Guzmán RM, Gallacher AE, Hall S, Levy RA, Jimenez RE, et al. Efficacy and safety of belimumab in patients with active systemic lupus erythematosus: a randomised, placebo-controlled, phase 3 trial. *Lancet* 2011;377(9767):721-31.

[90] Specchia ML, de Waure C, Gualano MR, Doria A, Turchetti G, Pippo L, et al. Health technology assessment of belimumab: a new monoclonal antibody for the treatment of systemic lupus erythematosus. *Biomed Res Int* 2014;2014:704207.

[91] Pierotti F, Palla I, Treur M, Pippo L, Turchetti G. Assessment of the economic impact of belimumab for the treatment of systemic lupus erythematosus in the italian setting: a cost-effectiveness analysis. *PLoS One* 2015;10(10):e0140843.

[92] Canadian Agency for Drugs and Technologies in Health. *CDEC Final Recommendation. Belimumab (Benlysta – GlaxoSmithKline Inc.). Indication: systemic lupus erythematosus. April 25,*

2012. 6p. Available from: https://www.cadth.ca/media/cdr/complete/cdr_complete_Benlysta_April-27-12_e.pdf

[93] Personal communication: Prof. Caroline Gordon, 15 January 2021.

[94] McElhone K, Abbott J, Sutton C, Mullen M, Lanyon P, Rahman A, et al. Sensitivity to change and minimal important differences of the LupusQoL in patients with systemic lupus erythematosus. *Arthritis Care Res (Hoboken)* 2016;68(10):1505-13.

[95] Connock M, Cummins E, Sutcliffe P, Court R, Johnson R, Grove A, et al. *Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus: a single technology appraisal*. Warwick: Warwick Evidence, 2011. 166p. Available from: https://www.nice.org.uk/guidance/ta397/documents/systemic-lupus-erythematosus-active-belimumab-evidence-review-group-report3

[96] Barber MRW, Hanly JG, Su L, Urowitz MB, St Pierre Y, Romero-Diaz J, et al. Economic evaluation of damage accrual in an international systemic lupus erythematosus inception cohort using a multistate model approach. *Arthritis Care Res (Hoboken)* 2020;72(12):1800-1808.

[97] 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.

National Institute for Health and Care Excellence

Centre for Health Technology Evaluation

ERG report – factual accuracy check and confidential information check

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (Review of TA397) [ID1591]

'*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 Thursday 28 January 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.

Contents

Issue 1 Patient switching to SC during the pandemic.

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
1. On page 16 of the ERG report, it reads: *"However, our clinical expert suggested that the majority of UK patients have switched to SC during the pandemic and so there might be a higher proportion of patients wanting SC in future if the disease appears to be as well controlled on SC as IV."*	Amend section to acknowledge that majority of patients in the UK have not switched to the SC formulation.	GSK sales data shows that a majority of patients have not switched to SC during the pandemic. Based on sales distribution data GSK estimate that a maximum of [redacted] patients in the UK have been switched to SC during the pandemic.	Not a factual error. This was the opinion of our clinical expert.

Issue 2 Impact of peak incidence on loss of work

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
1. 'It should also be noted that as the peak incidence in both men and women are close to the historical retirement age in the UK, it would be useful to know how many years of work people with SLE in the	Amend section to acknowledge that more working years are likely to be impacted in women than in men.	As stated in the submission, peak incidence in women is between the ages of 40-49.[1]	Not a factual error. No changes made.

UK lose due to their condition.' Section 2.2. (page 24)			

Issue 3 Data extraction and quality assessment methodology

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
1. ERG comment: It is normally recommended that two reviewers are involved in data extraction to avoid bias and error. The lack of information regarding how data extraction was addressed presents issues regarding transparency. Section 4.1.3. (Page 42)	Remove comments regarding recommendation of two reviewers for data extraction.	As stated in the provided clinical SLR report, data extraction was performed by two reviewers: 'Data were independently captured by a single investigator, with validation performed by a second, senior investigator.' (P5)	Not a factual error. It is unclear what is meant by validation. No changes made.

2. "According to Appendix D of the CS, the quality assessment of the RCTs was completed using the using the University of York, Centre for Reviews and Dissemination guidelines. The non-randomised LTE and RWE studies were assessed using the Downs and Black checklist. ERG comment: It is normally recommended that two reviewers are involved in the assessment of study quality to avoid bias and error." Section 4.1.4 page 42	We propose adding the <u>underlined</u> sentence: "According to Appendix D of the CS, the quality assessment of the RCTs was completed using the using the University of York, Centre for Reviews and Dissemination guidelines. The non-randomised LTE and RWE studies were assessed using the Downs and Black checklist. <u>Quality assessment of belimumab studies was performed by a single reviewer and verified by a second, senior reviewer."</u>	Providing additional information noted by the ERG as missing	Not a factual error at the time of writing the report. Thank you for the additional information.
3. "No information was provided on the number of reviewers who assessed the quality of included studies." Section 4.2.4 page 54	We propose amending the sentence as above, i.e., to "Quality assessment of belimumab studies was performed by a single reviewer and verified by a second, senior reviewer."		Not a factual error at the time of writing the report. Thank you for the additional information.

Issue 4 Availability of data in the paediatric population

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
1. 'Evidence is missing for specific populations, such as children and patients with severe active CNS lupus.' Table 1.1 (page 11), Table 1.2 (page 12) and Section 3.1 (page 34)	Remove the reference to missing evidence in children throughout the ERG report.	Comparative evidence in the paediatric population for belimumab vs standard care alone is provided from the paediatric PLUTO trial. Results from its double-blind phase can be found in Appendix O to the CS. At Week 52, compared with placebo, numerically higher proportions of patients receiving belimumab met the primary efficacy endpoint (SRI4) and results of this small paediatric study were consistent with the phase 3 programme of belimumab in adults. However, due to the severity and relatively low prevalence of childhood onset SLE, a large-scale trial powered for statistical significance testing was considered unfeasible, and the analysis of PLUTO was descriptive. The results available from PLUTO suggest that outcomes in the paediatric population can be expected to be similar to those observed in adults. Because the majority of the evidence in belimumab pertains to the adult population, the paediatric population was not included in the economic modelling or the systematic literature review (which was	As stated by the company (CS, Table 1): "GSK presents the results of PLUTO, the paediatric trial of IV belimumab compared with placebo within an appendix of the submission. The paediatric population recruited in PLUTO is limited (due to the rarity of childhood SLE) and the study was not statistically powered to show a difference between treatment groups." Therefore, the ERG still believes reliable evidence for children is missing. No changes made.

		an update of the searches performed for TA397).	

Issue 5 PSM analysis- based calibration

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
1. "This means that the calibration factor is biased towards belimumab preventing organ damage, as most patients withdrew from the BLISS non-US LTE before 5 years, regardless of discontinuation of belimumab over that period, and those remaining on treatment at 5 years are likely to either have responded unusually well to belimumab or had slower progressing SLE than those that withdrew." (Table 1.5 pages 14 and 15)	We suggest removing the relevant fragments.	The BLISS non-US LTE does not support the notion that patients continuing on trial "responded unusually well to belimumab". In the BLISS non-US LTE, only 6 patients (0.8%) withdrew due to lack of efficacy and 36 (4.9%) due to physician decision. The most common reason for withdrawal was withdrawal by subject (151 patients, 20.5%). The withdrawal rate by the end of Year 5 (40.8%) does not support the statement that "most patients withdrew from the BLISS non-US LTE before 5 years". Of the 735 patients entering this study, 345 patients (46.9%) continued beyond 5 years. However, by the end of Year 5, 90 patients (12.2%) completed the study while 300 (40.8%) withdrew from it. Patients completed the study as belimumab became commercially available in their country, so study completion cannot be considered equivalent to cessation of treatment with belimumab. Overall, data from the BLISS non-US LTE do not suggest a loss of belimumab	Not a factual error. The ERG report states that the withdrawal rate was concerning because of the potential for bias. It cannot be assumed that the BLISS trial patients who did not sign up for the LTE did not sign up because they were able to get belimumab as part of regular practice, just like it cannot be assumed all patients not signing up for the LTE did so because they received no benefit from belimumab. The ERG still believes there is a high probability of bias from a high withdrawal rate. The implicit assumption the CS makes is that anyone who withdrew prior to 5 years had the same outcome trajectory as those that did not, which is a problem.

		efficacy over time and do not support the concerns around using the 5-year data point for calculating the calibration factor. It is also possible that the ERG meant the BLISS US LTE, which was used in the PSM analysis. In this LTE, 14 patients (5.2%) withdrew due to lack of efficacy and 17 (6.3%) withdrew due to physician decision. Again, withdrawal by subject was the most common reason for withdrawal (n=31, 11.6%). At the end of year 5, 192 of 268 patients (71.6%) continued in the study, so again, it is incorrect to state that most patients withdrew before 5 years. Long-term data from this study also do not suggest a loss of belimumab efficacy over time, neither do they spark concern around using the 5-year data point for calibration in the model. See Issue 12 for responses specific to use of the calibration factor within the economic model.	The ERG does not question the first bit of this point – the 5 year outcome rate is a reasonable estimate of the outcome for people who stay on belimumab for 5 years. What we question is applying the efficacy of belimumab over 5 years taken solely from patients on belimumab for 5 years and applying that to everyone who is eligible to take belimumab. This is a factual error: we did mean the US LTE. This has been corrected. Both withdrawal because of a physician decision and withdrawal by the subject could be withdrawal due to lack of efficacy – there is no way to know if lack of efficacy played a role in any/all of the decisions to stop belimumab. Unless there were major AEs (which would be problematic by themselves), it is difficult to see why a physician would stop belimumab if it was working well. The ERG point is that the withdrawal rate is very high, when considering patients starting in any

			BLISS study that progressed to an LTE.
2. "Patients with five years of follow-up of belimumab are therefore more likely to either have slowly progressing SLE, had a favourable response to belimumab with no or limited adverse effects, or a combination, compared with patients who either withdrew before five years had elapsed on belimumab, or chose not to enter the BLISS LTE studies. The CS states in table 64 that "patients who had not demonstrated a sufficient response with belimumab during the Phase 3 studies would unlikely have continued into the extension study", emphasising this point." (Section 4.4.1 page 74)	We suggest removing the relevant fragment and modifying the resulting conclusions.	The SmPC for belimumab states that "The patient's condition should be evaluated continuously. Discontinuation of treatment with Benlysta should be considered if there is no improvement in disease control after 6 months of treatment." However, data from the BLISS LTE studies suggest that withdrawals due to loss of efficacy were infrequent (see point 1 in this Issue), which does not support the notion of belimumab efficacy being overestimated by using 5-year LTE data for calibrating the model. Discontinuation of belimumab by patients who responded at the 6-month landmark was included in the economic model and can be expected to occur in clinical practice for a variety of reasons. The most common reason for withdrawal from the BLISS LTEs was patient decision. It is reasonable to envisage that regular monthly clinic appointments for infusions taking place over many years can become a burden, particularly in patients whose disease is well controlled, enabling them to lead active lives. This could lead to treatment cessation or irregular/missed dosing in patients who do respond well to belimumab. Wider availability of the SC formulation of	Not a factual error. The ERG does not question the efficacy of belimumab after 5 years for those who remained on belimumab for those 5 years. We do question applying the 5 year efficacy to all 5 years, including for people who discontinue it for a variety of reasons.

		belimumab is likely to improve long-term adherence to treatment.	
3. "The largest issue is that the withdrawal rate of patients in the BLISS LTE studies was high. Patients with five years of follow-up of belimumab are more likely to either have slowly progressing SLE, had a favourable response to belimumab with no or limited adverse effects, or a combination, compared with patients who either withdrew before five years had elapsed on belimumab, or chose not to enter the BLISS LTE studies. Because a maximum of 34% of the original BLISS trial patients were included in the PSM, the size of this bias from using the calibration factor could be high, and almost certainly biases the model to make belimumab seem more cost-effective than it is." Section 4.6 page 77	We suggest revising this fragment in line with the explanations provided for the two points above	Please see points 1 and 2 in this issue.	See above – none of these are factual errors.
4. "One way of producing a less biased estimate of	We suggest including a statement that it is uncertain whether such an	While a UK cohort would be preferentially used in the UK HTA setting, it is	Not a factual error.

	long-term SDI at least for those on belimumab treatment would be as follows: instead of matching BLISS LTE patients with TLC, create propensity scores using UK SLE cohort data (e.g., BILAG BR data), which could be tailored so only HDA-1 and HDA-2 subgroups are included, then weight the BLISS LTE data to make it generalisable to a UK cohort using the propensity scores. This does not require follow-up data for the UK cohort, just enough information to weight the BLISS LTE data so it is generalisable to a UK cohort." Table 1.5 pages 14 and 15	analysis would be feasible within the time frame for this appraisal.	uncertain how many of the 203 patients initiating non-biologic treatment in the BILAG-BR could be matched to the BLISS non-US LTE. These patients are likely to be a less severe SLE group in general, as they are being managed on non-biologic therapies. In addition, due to data availability issues, such an analysis would need to be completed by the University of Manchester and this may not be possible within the appraisal timelines. Furthermore, it should be noted that the Toronto Lupus Cohort (TLC) was identified in a systematic literature review as the preferred comparator cohort based on its size, the extent of organ damage, and the severity of SLE disease activity. The TLC collected data at relatively short (3–4-month) intervals and used disease severity and organ damage progression scales similar to those used within the BLISS LTE study.[5]	
5.	"the PSM analysis itself is probably biased towards belimumab being effective" Table 1.5 pages 14 and 15	We suggest removing the relevant fragment.	The PSM analysis was based on robust, well-established, and reproducible methodology and held up to peer review, resulting in a publication in a high-profile rheumatology journal.[5] We do not believe this is a fair representation of the study results.	Not a factual error. See also comments above.

Issue 6 Relevance of pooled data from BLISS-52 and BLISS-76 to the UK setting

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
1. "There was a relative lack of evidence for clinical effectiveness of belimumab seen in the BLISS-76 trial. The results favourable for belimumab submitted for the pooled population across trials were largely driven by BLISS-52 results. The SLE population in BLISS-76 is more likely to resemble that in the UK than that in the BLISS-52. Therefore, the BLISS-76 results are probably more relevant to the decision problem than those from BLISS-52, and results from the pooled population may overestimate the effectiveness of belimumab in the UK population."	We propose removing the relevant fragment.	The BLISS-76 trial was positive, meeting its primary endpoint, so that this trial did provide evidence for clinical effectiveness of belimumab. With regards to relevance to the UK setting, BLISS-76 subgroup analysis of SRI-4 response by region actually demonstrated that numerically greatest benefit of belimumab 10 mg/kg vs placebo was observed in patients enrolled in Western Europe and Israel (see Figure 7-1 in the BLISS-76 CSR). In the Western Europe and Israel subgroup, 38 out of 75 patients (50.7%) receiving 10 mg/kg belimumab achieved SRI-4 at 52 weeks compared with 15 out of 64 (23.4%) patients receiving placebo. Treatment by region interaction was, however, not statistically significant. This issue was extensively discussed during TA397and the conclusion from the Committee (see Section 4.7of the guidance) was as follows: "….The clinical experts explained that, because the UK is a multi-ethnic country and systemic lupus erythematosus affects many ethnic groups more severely than white	Not a factual error. This was the same critique as presented by the ERG for TA397.

		populations, data from different populations would still be relevant to the UK. Furthermore, the committee understood from the clinical experts that clinical practice varies between countries, for example, in the US higher doses of corticosteroids are used than in the UK. Therefore, there may also be issues about the relevance of the data from BLISS-76. On balance, the committee concluded that BLISS-76 was more representative of the population of England and Wales than BLISS-52. However, data from BLISS-52, and therefore from the pooled analysis would be relevant."	

Issue 7 Incorrect reporting of BILAG-BR data

<table>
<tr><th>Description of problem</th><th>Description of proposed amendment</th><th>Justification for amendment</th><th>ERG Response</th></tr>
<tr><td>1. "The analysis included [redacted] patients receiving belimumab and [redacted]</td><td>"The analysis included [redacted] patients receiving belimumab and [redacted]</td><td>As stated on page 26 of Appendix P to the submission, the primary analysis investigated the treatment assigned at baseline of each treatment round. Therefore, the correct numbers should be [redacted] and [redacted] for rituximab and belimumab, respectively, in line with Table 3 page 27 of Appendix P. Please also note incorrect confidentiality marking in this fragment of the ERG report, as outlined in "Issue 9: Discount rate incorrectly stated.
<table>
<tr><th>Description of problem</th><th>Description of proposed amendment</th><th>Justification for amendment</th><th>ERG Response</th></tr>
</table>
</td><td>Appendix P states:</td></tr>
</table>

ving belimumab and [redacted] receiving rituximab." Section 4.3.2 page 72, and Section 4.6 page 77 of the ERG Report	receiving rituximab."	1. The discount rate reported in section 3.5 on page 36 for the vials of belimumab is [redacted]. This is incorrect. The proposed discount rate for the IV vials is [redacted]	Amend to correct discount rate.	The discount rate seems to be a calculation error.	This has been amended to [redacted]%.	[redacted] Therefore, the ERG is correct. We have changed the confidentiality marking.

Issue 8 Annual costs of belimumab incorrectly stated.

Description of problem	Description of proposed amendment	Justification for amendment	
1. In section 5.2.9 Resources and costs (p100), it states: *"This results in drug-related costs of [redacted] for the first year and [redacted] for subsequent years for the SC formulation and [redacted] and [redacted] for the IV formulation*	Amend to: "This results in drug-related costs of [redacted] for the first year and [redacted] for subsequent years for the IV formulation and [redacted] for each year of the SC formulation (with the patient access	The correct values are stated in Table 69, Page 168 of the CS.	By drug-related cost, the ERG meant cost of the drug plus administration cost. This description has now been added to the ERG report.

		(with the patient access scheme included).”	scheme included).”		

Issue 9 Dates of economic searches are stated.

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
The ERG report states on page 79 that *‘The date the searches were conducted was not supplied, and no numbers of records retrieved were provided for either the NICE website or PubMed searches. No additional databases or other resources were searched.’*	We propose removing the relevant fragment.	The dates of the economic searches are stated – see titles of tables in *Appendix G. Published cost-effectiveness studies*.	Not a factual error. The ERG comments refer to the search methods in Appendix H/I. In Appendix H/I the company do not give the dates of the searches conducted.

Issue 10 Use of calibration factor in the model.

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response

		1. "Since the company only used the calibration factor on the belimumab arm of the model, it was not necessary to use the PSM to obtain the calibration factor. The PSM reduces the number of patients included with the aim to obtain a comparable set of patients to the TLC. As no comparative estimates for treatment effectiveness or long-term outcomes are derived, the purpose of using this reduced (matched) set of patients is	We suggest revising this fragment in line with the explanation provided.	Comparative estimates for the calibration factors were derived from the PSM study. The corresponding comparative ST estimate was higher than 1, implying that the model underestimated organ damage progression for the ST arm. The approach taken not to use the calibration factor for ST represents a conservative stance. If the relevant calibration factor was used in the ST arm, this would result in an increased rate of organ damage for	Not a factual error. The company are correct in pointing out that given that they had used the PSM, not applying the calibration factor to the ST arm is conservative. The ERG did not question this. The ERG just questions whether the calibration factor is reliable, since it is derived from the PSM the generalisability of which to the UK setting is very questionable.	

		questionable, in particular because the resulting set of patients is not necessarily more generalisable to the UK setting than the original sample" (Page 92)		this arm, reducing the ICER. Not applying the calibration factor for ST in the model, however, does not imply that it was therefore not necessary to base the belimumab calibration factor on the PSM. Without the PSM it would not be possible to know if the comparative ST calibration factor would be lower or higher than 1, and whether the ST arm organ damage progression was over- or underestimated in the model.		
		2. "It is further questionable	We suggest revising this	The HDA-2 population has	Not a factual error. There is no	

			whether results of the PSM are applicable to the HDA-2 population." (page 92)	fragment in line with the explanation provided.	more active disease than the PSM-based population. The relative, comparative benefit of belimumab compared to ST has been shown to be greater in the HDA populations for many outcomes in the Phase 3 studies. It is reasonable to assume this would translate to increased benefit in terms of rate of organ damage progression and this would imply that the relative difference in calibration factors between belimumab and ST could indeed be higher in the HDA-2	evidence for the generalisability of the PSM to the HDA-2 population.	

				population. Therefore, adopting the PSM based calibration factor to the HDA-2 population could in fact be conservative for belimumab. In addition, many conservative steps were taken in applying the calibration to the model, i.e. not also adjusting the ST arm and restricting the organ damage benefit to 6 years in the base case.		
		3. "Calibrating the existing model to one observed time-point (five years) is problematic, especially given the attrition of patients treated with belimumab in the BLISS LTE	We suggest revising this fragment in line with the explanation provided to the points above .		Not a factual error. See our responses to Issue 5.	

		studies mentioned above and in Section 4.2.6." (Page 93)				
		4. The pre-calibration model already assumes a long-term treatment effect based largely on clinical expert opinion. In TA397, it is stated that "In the simulation model, an assumption was made that the additional absolute effect of belimumab on disease activity reduction remains constant after one year. This is a key model assumption and was discussed with Professor	We suggest revising this fragment in line with the explanation provided.	The long-term treatment effect assumption already used in the model and validated by expert opinion concerns the treatment effect on sustained disease activity reduction. That is a different disease outcome than organ damage development. The indirect positive effect of sustained disease activity reduction on reduced organ damage progression	Not a factual error. The statement is correct. The company seems to be questioning the context, if anything. However, the ERG would argue that treatment effectiveness is related to both disease outcomes (disease activity and organ damage) and it is relevant, given the uncertainty surrounding long-term treatment effectiveness on both outcomes, to assess assumptions for their potential	

		Petri who has observed patients on belimumab in her clinic for a number of years as part of the Phase 2 open-label extension study".93 The ERG clinical expert considers this a reasonable assumption. (Page 93)		(modelled through the natural history models) was shown in the model to still underestimate the impact of belimumab on organ damage reduction as per the calibration exercise based on the LTE data. Hence, there is no double adjustment of a long-term treatment effect.	bias and how this adds up.	
		5. "The pre-calibration model already assumes that the absolute SS value is adjusted in comparison to that predicted by JH. In TA397,44 it is stated that "The adjusted JH model was used to predict the	We suggest revising this fragment in line with the explanation provided.	The pre-calibration model adjusted the SELENA-SLEDAI score, which is a measure of disease activity. This is a different outcome to that of irreversible organ damage progression, to	Not a factual error. See issue above for explanation.	

		SLEDAI score of a patient treated with SoC after one year. The model allows for the selection of the original and adjusted JH model. For the base case analysis, the adjusted JH model has been used". In the final appraisal document of TA397, the committee also noted that "that there was uncertainty about whether the equations derived from the Johns Hopkins data could be reliably applied to the target population because of differences in study populations".93		which the calibration factor was applied. As explained in the justification to the previous point, there is no double adjustment of a long-term treatment effect.		

Given that the adjustment favours the effectiveness of belimumab, it is likely that this adjustment causes the ICER to be underestimated." (page 93)					

Issue 11 Responder rule implementation in the model

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
1. In the model, the probability of being a responder	Discard this issue as it is not a model error (see explanation).	Response is determined at 24 weeks and is based on the 24-week SS score data from the BLISS trials for the HDA subgroups. Belimumab patients who do not respond at 24 weeks (i.e. those that do not have at least a 4 or more decrease in SS score from baseline to week 24) are discontinued from belimumab in the model and continue as a ST patient.	Not a factual error. The company's explanation is not addressing the point, which is

			is based on baseline SS score, which is linked to the responder criteria applied to patients in the BLISS trials (i.e., only for patients with a reduction of ≥ 4 points in SS		Corresponding costs in the model for belimumab treatment are only applied for the first 24 weeks (sub function costs in clsPatient copied below, relevant section highlighted) and this represents what would happen in clinical practice. Hence, there is no underestimation of costs in the model. Non-responders at 24 weeks could indeed have been responders at 52 weeks, but that is not an option in the model as they have already been taken off belimumab treatment at week 24 as per the responder rule. So indeed, non-responders on belimumab (at week24) could still have an SS point improvement of more than 4 points at 52 weeks, following their switch to ST at week 24.
					
					'short term costs
					Call mCosts.AddShort(curTime - 1, cur
					
					'medication costs
					If mTx = Belimumab Then
					Call mCosts.AddMed(curTime - 1, 1
					ElseIf mTx_pr = Belimumab And NatDisc
					Call mCosts.AddMed(curTime - 1, C
					ElseIf mTx_pr = Belimumab And respond
					Call mCosts.AddMed(curTime - 1, 2
					End If
						that some patients are designated non-responders at 24 weeks despite achieving response at 52 weeks (and the concern is that if it were modelled it could likely be shown that some of those fulfilled the response criteria at 24 weeks as well). This issue is a result of how the company modelled whether a patient is	

		score). Hence, it is estimated at baseline in the model and not directly linked to the actual improvement in SS score in the model. In turn, actual SS scores are estim			a responder irrespective of their characteristics. If non-responders in the model who discontinue treatment actually have a response, then this is likely not in line with clinical practice and may under-estimate cost.	

			ated based on a regression model where response is an independent variable, given that a 24-week time point does not exist in the model. As a result, a large proportion of				

		patients is classed as non-responders but experiences >4 points reduction in SS at 52 weeks. This could lead to under-estimation of costs in the model compared to clinica				

			l practice as patients with no response do not continue belimumab. A structural adjustment to the company's model is necessary to resolve this issue. (page 94)				

		2. Due to a modelling error, belimumab non-responders have the same reduction in disease activity at 52 weeks as patients in the standard therapy arm. (page 95)	We suggest revising this fragment in line with the explanation provided and discarding changes to the CS model to overcome the perceived error in the modelling of non-responder disease activity.	Response is assessed at 24 weeks, not at 52 weeks. In the first year of the model, belimumab non-responders at 24 weeks are discontinued from belimumab immediately and changed to ST. This means that they are on ST at 52 weeks (and no longer on belimumab) and have been for 28 weeks, which means that the use of the ST disease reduction is appropriate for belimumab non-responders at 52 weeks. The ERG change to the model (as per the below) needs to be undone, as no error needs correcting, and in turn the implemented change introduces an error to the programming routine. Non-responder patients need to be dropped to SoC in the original routine called UpdSSFirstHalfYear, as that is the timepoint when non-responders are dropped off from belimumab, not at 52 weeks.	Not a factual error. The company's modelling is not in line with the CS and not in line with the evidence presented in the CS (patients on ST can have better response than belimumab non-responders). It is therefore an error. If the ERG introduced another error by correcting the company's error, we request	

```
Private Sub UpdSS(firstyear As Boolean)

    Dim tmp As Double, mSS_Bel As Double, fact_eff As Double
    tmp = curSS

    If firstyear Then
        mSS_SoC = mSS_SoC + mSS_B * (mInput.SSRed(1))

        If curPat.mTx = Belimumab Then
            mSS_Bel = mSS_SoC + diffwk52
            curSS = mSS_Bel
            ' ERG_1: moved the next 3 lines here from first h
            If mInput.respAnalysis = True And curPat.responde
                curPat.mTx = SoC
            End If

        Else
            curSS = mSS_SoC
        End If
```

that the company fix their error in a way that does not introduce another error and provide clear explanation of what the error would be and how it was avoided as well as detail on the implementation.

Issue 12 Sampling of organ damage and death occurs after allocation to treatment.

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
1. In the VBA, first, a	This is not an issue, see	Organ damage and mortality	Not a factual error. The

			simulated patient is allocated to a treatment and organ damage and death are only sampled within the treatment arm. This leads to the same simulated patient (same age, gender, SS score) experiencing differential organ damage and times of death only dependent on allocation to treatment arm but not caused by this allocation (so just because of sampling). This induces noise in the model and makes validation	explanation provided.	depends, amongst other things, on AMS (covariate included in the time to event equations), which depends on the SELENA-SLEDAI score. In turn, the progression of SELENA-SLEDAI score depends on treatment allocation. So, contrary to what is stated, the occurrence and timing of organ damage and mortality does depend on treatment allocation. Therefore, organ damage items involved, and death cannot be sampled before treatment allocation. It would forego the benefits of belimumab on	same simulated patient has differential organ damage scores. Ours was an accurate description of the problems this can cause.	

			difficult. (page 108, in the table, and the entire more detailed text under e))		reduced organ damage development and mortality. It is true that this introduces additional sampling error. It could also mean you can sometimes get extreme differences between the same patient in both arms. This is all caused by sampling error. , but no solution to that has been identified. Organ damage and death events reported as in year 0 of the patients traces occur in the first year of the model, so after model entry. Baseline organ damage (before model entry) and all other baseline characteristics for		

				that matter, are sampled before treatment allocation and patients are cloned for both arms. Noise, or sampling error has been checked and mitigated by simulating enough patients. Convergence plots were generated and checked to make sure model results are stable and sampling error reduced to a minimum. To ensure stable results, free from sampling errors, convergence checks had been performed for the original model. In checking different seeds and checking different amounts of patient samples, it was		

				found that simulating 50,000 patients in each arm led to stable results, free from noise. This also holds true for the updated model.		
		Incorrect confidentiality marking" below.				

Issue 13 Incorrect reporting of reasons for withdrawal from the BLISS trials and the BLISS LTE studies

<table>
<tr><th>Description of problem</th><th>Description of proposed amendment</th><th>Justification for amendment</th><th>ERG Response</th></tr>
<tr>
<td>1. “Reasons for withdrawal from the BLISS-76 and BLISS-52 trials88, 89 for belimumab patients included: 47 patient requests (4% of all belimumab patients); 75 adverse events (6.7%); 57 withdrawals due to lack of efficacy (5.1%); 6 withdrawals due to lack of compliance (0.5%); 19 patients lost to follow-up (1.7%); 22 protocol violations (2.0%); 13 investigator decisions</td>
<td>“Reasons for withdrawal from the BLISS-76 and BLISS-52 trials, for 10mg/kg belimumab patients included: 16 patient requests (2.8% of all belimumab 10 mg/kg patients); 34 for adverse events (6.0%); 26 withdrawals due to lack of efficacy (4.6%); 3 withdrawals due to lack of compliance (0.5%); 9 patients lost to follow-up (1.5%); 8 protocol violations (1.4%); 6 investigator decisions (1.1%); and 11 other reasons (1.9%).”</td>
<td>The numbers quoted do not align with the CSRs, which state as follows:
BLISS-52 (Table 6-1 page 69)
<table>
<tr><th></th><th>Placebo N = 287</th><th>1 mg/kg N = 288</th><th>10 mg/kg N = 290</th><th>All Groups N = 865</th><th>P-value[1]</th></tr>
<tr><td>Completed</td><td>226 (78.7%)</td><td>240 (83.3%)</td><td>241 (83.1%)</td><td>707 (81.7%)</td><td>0.2827</td></tr>
<tr><td>Withdrawn</td><td>61 (21.3%)</td><td>48 (16.7%)</td><td>49 (16.9%)</td><td>158 (18.3%)</td><td>0.2827</td></tr>
<tr><td>Subject Request</td><td>7 (2.4%)</td><td>6 (2.1%)</td><td>3 (1.0%)</td><td>16 (1.8%)</td><td>0.3978</td></tr>
<tr><td>Adverse Event</td><td>19 (6.6%)</td><td>16 (5.6%)</td><td>15 (5.2%)</td><td>50 (5.8%)</td><td>0.7458</td></tr>
<tr><td>Lack of Efficacy</td><td>16 (5.6%)</td><td>12 (4.2%)</td><td>12 (4.1%)</td><td>40 (4.6%)</td><td>0.6520</td></tr>
<tr><td>Lack of Compliance</td><td>1 (0.3%)</td><td>1 (0.3%)</td><td>1 (0.3%)</td><td>3 (0.3%)</td><td>1.0000</td></tr>
<tr><td>Lost to Follow-up</td><td>4 (1.4%)</td><td>6 (2.1%)</td><td>3 (1.0%)</td><td>13 (1.5%)</td><td>0.5761</td></tr>
<tr><td>Protocol Violation</td><td>7 (2.4%)</td><td>2 (0.7%)</td><td>3 (1.0%)</td><td>12 (1.4%)</td><td>0.2027</td></tr>
<tr><td>Investigator Decision</td><td>3 (1.0%)</td><td>2 (0.7%)</td><td>3 (1.0%)</td><td>8 (0.9%)</td><td>0.9136</td></tr>
<tr><td>Other</td><td>4 (1.4%)</td><td>3 (1.0%)</td><td>9 (3.1%)</td><td>16 (1.8%)</td><td>0.1591</td></tr>
<tr><td>Pregnancy[2]</td><td>4 (1.4%)</td><td>3 (1.0%)</td><td>8 (2.8%)</td><td>15 (1.8%)</td><td>-</td></tr>
</table>
[1] P-value for comparison across 3 treatment groups obtained from likelihood ratio or Fisher's exact test.
[2] In addition, Subject BR001-005 in the placebo group discontinued treatment due to an AE of spontaneous abortion and Subject RO005-001 in the 10 mg/kg group discontinued treatment due to investigator decision (pregnancy).
</td>
<td>The ERG report presented results for both belimumab arms combined and using the withdrawal rates at the end of each study. This is not a factual error.
However, the number of withdrawals were slightly wrong, and have been amended.</td>
</tr>
</table>

<table>
<tr>
<td>(1.2%); and 25 other reasons (2.3%).”

Section 4.4.1 page 73</td>
<td></td>
<td>BLISS-76 (Table 6-1 page 75). Note that discontinuations to week 52 were used when correcting the statement from the ERG to combine reasons for discontinuation across BLISS-52 and BLISS-76.

Table 6-1 Subject completion status through Week 52 and Week 76

<table>
<tr><th></th><th colspan="5">Through Week 52</th><th colspan="5">Through Week 76</th></tr>
<tr><th></th><th>Placebo N = 275</th><th>1 mg/kg N = 271</th><th>10 mg/kg N = 273</th><th>All Groups N = 819</th><th>P-value[1]</th><th>Placebo N = 275</th><th>1 mg/kg N = 271</th><th>10 mg/kg N = 273</th><th>All Groups N = 819</th><th>P-va</th></tr>
<tr><td>Completed</td><td>205 (74.5%)</td><td>216 (79.7%)</td><td>209 (76.6%)</td><td>630 (76.9%)</td><td>0.3511</td><td>186 (67.6%)</td><td>199 (73.4%)</td><td>191 (70.0%)</td><td>576 (70.3%)</td><td>0.3</td></tr>
<tr><td>Withdrawn</td><td>70 (25.5%)</td><td>55 (20.3%)</td><td>64 (23.4%)</td><td>189 (23.1%)</td><td>0.3511</td><td>89 (32.4%)</td><td>72 (26.6%)</td><td>82 (30.0%)</td><td>243 (29.7%)</td><td>0.3</td></tr>
<tr><td>Subject Request</td><td>24 (8.7%)</td><td>14 (5.2%)</td><td>13 (4.8%)</td><td>51 (6.2%)</td><td>0.1179</td><td>28 (10.2%)</td><td>17 (6.3%)</td><td>20 (7.3%)</td><td>65 (7.9%)</td><td>0.2</td></tr>
<tr><td>Adverse Event</td><td>16 (5.8%)</td><td>13 (4.8%)</td><td>19 (7.0%)</td><td>48 (5.9%)</td><td>0.5608</td><td>23 (8.4%)</td><td>18 (6.6%)</td><td>23 (8.4%)</td><td>64 (7.8%)</td><td>0.6</td></tr>
<tr><td>Lack of Efficacy</td><td>15 (5.5%)</td><td>12 (4.4%)</td><td>14 (5.1%)</td><td>41 (5.0%)</td><td>0.8524</td><td>20 (7.3%)</td><td>12 (4.4%)</td><td>17 (6.2%)</td><td>49 (6.0%)</td><td>0.3</td></tr>
<tr><td>Lack of Compliance</td><td>2 (0.7%)</td><td>1 (0.4%)</td><td>2 (0.7%)</td><td>5 (0.6%)</td><td>1.0000</td><td>2 (0.7%)</td><td>2 (0.7%)</td><td>2 (0.7%)</td><td>6 (0.7%)</td><td>1.0</td></tr>
<tr><td>Lost to Follow-up</td><td>3 (1.1%)</td><td>4 (1.5%)</td><td>6 (2.2%)</td><td>13 (1.6%)</td><td>0.5754</td><td>4 (1.5%)</td><td>6 (2.2%)</td><td>6 (2.2%)</td><td>16 (2.0%)</td><td>0.7</td></tr>
<tr><td>Protocol Violation</td><td>5 (1.8%)</td><td>2 (0.7%)</td><td>5 (1.8%)</td><td>12 (1.5%)</td><td>0.5128</td><td>6 (2.2%)</td><td>6 (2.2%)</td><td>6 (2.2%)</td><td>18 (2.2%)</td><td>0.9</td></tr>
<tr><td>Investigator Decision</td><td>2 (0.7%)</td><td>3 (1.1%)</td><td>3 (1.1%)</td><td>8 (1.0%)</td><td>0.8272</td><td>3 (1.1%)</td><td>3 (1.1%)</td><td>4 (1.5%)</td><td>10 (1.2%)</td><td>0.9</td></tr>
<tr><td>Other</td><td>3 (1.1%)</td><td>6 (2.2%)</td><td>2 (0.7%)</td><td>11 (1.3%)</td><td>0.3030</td><td>3 (1.1%)</td><td>8 (3.0%)</td><td>4 (1.5%)</td><td>15 (1.8%)</td><td>0.2</td></tr>
<tr><td>Pregnancy[2]</td><td>-</td><td>2 (0.7%)</td><td>1 (0.4%)</td><td>3 (0.4%)</td><td>-</td><td>-</td><td>3 (1.1%)</td><td>1 (0.4%)</td><td>4 (0.5%)</td><td></td></tr>
</table>
[1] P-value for comparison across 3 treatment groups obtained from likelihood ratio or Fisher's exact test.
[2] Includes Subjects MX003-003, MX007-001, and MX008-009 in the 1 mg/kg group and Subject US041-017 in the 10 mg/kg group who were withdrawn because becoming pregnant. In addition, Subject US061-002 in the 10 mg/kg group became pregnant and was lost to follow-up (see Section 10.3.5).
</td>
<td></td>
</tr>
<tr>
<td>2. “Reasons for withdrawal from the LTE studies (at any time), included: 151 pregnancies (20.5% of all BLISS LTE patients); 69 adverse events (9.3%); 69 other reasons, usually withdrawal of consent (9.3%); 36 physician</td>
<td>“Reasons for withdrawal from the LTE studies (at any time), included: 182 patient withdrawals (18.1% of all BLISS LTE patients); 94 adverse events (9.3%); 91 other reasons, usually withdrawal of consent (9.1%); 54 physician decisions (5.3%); 16 withdrawals due to lack of compliance</td>
<td>Having reviewed the CSRs of the LTE studies and verified the submission for this appraisal, we were unable to locate the sources of the information quoted by the ERG. The proposed amendment in based on the CSRs for the LTE studies, which state as follows:

BLISS non-US LTE (Table 5 page 36 of the CSR)</td>
<td>This is a factual error: we only included the BLISS non-US LTE withdrawal rates, and have amended the report with the suggested combined withdrawal rates. Note that there were 53 physician decisions (36+17), not 54 as suggested.</td>
</tr>
</table>

decisions (439%); 13 withdrawals due to lack of compliance (1.8%); 22 patients lost to follow-up (3.0%); six withdrawals due to lack of efficacy (0.8%); and four protocol deviations (0.5%).” Section 4.4.1 page 73	(1.6%); 34 patients lost to follow-up (3.4%); 20 withdrawals due to lack of efficacy (2.0%); and 5 protocol deviations (0.5%).”		

	Overall
Subjects Starting Interval	735
Withdrawn, n (% of subjects starting interval)	367 (49.9)
Adverse event	69 (9.4)
Lack of efficacy	6 (0.8)
Lost to follow-up	22 (3.0)
Non-compliance with study drug	10 (1.4)
Other	69 (9.4)
Physician decision	36 (4.9)
Protocol deviation	4 (0.5)
Withdrawal by subject	151 (20.5)
Completed, n (% of subjects starting interval)	368 (50.1)

BLISS US LTE (Table 4 page 42 of the CSR):

	Overall
Subjects Starting Interval	268
Withdrawn, n (% of subjects starting interval)	128 (47.8)
Adverse event	25 (9.3)
Lack of efficacy	14 (5.2)
Lost to follow-up	12 (4.5)
Non-compliance with study drug	6 (2.2)
Other	22 (8.2)
Physician decision	17 (6.3)
Protocol deviation	1 (0.4)
Withdrawal by subject	31 (11.6)
Completed, n (% of subjects starting interval)	0
Ongoing, n (% of subjects starting interval)	0

Issue 14 Discount rate incorrectly stated.

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
2. The discount rate reported in section 3.5 on page 36 for the vials of belimumab is [redacted]. This is incorrect. The proposed discount rate for the IV vials is [redacted]	Amend to correct discount rate.	The discount rate seems to be a calculation error.	This has been amended to [redacted]%.

Issue 15 Annual costs of belimumab incorrectly stated.

Description of problem	Description of proposed amendment	Justification for amendment	
2. In section 5.2.9 Resources and costs (p100), it states: *"This results in drug-related costs of [redacted] for the first year and [redacted] for subsequent years for the SC formulation and [redacted] and [redacted] for the IV formulation (with the patient access scheme included)."*	Amend to: "This results in drug-related costs of [redacted] for the first year and [redacted] for subsequent years for the IV formulation and [redacted] for each year of the SC formulation (with the patient access scheme included)."	The correct values are stated in Table 69, Page 168 of the CS.	By drug-related cost, the ERG meant cost of the drug plus administration cost. This description has now been added to the ERG report.

Issue 16 Dates of economic searches are stated.

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
The ERG report states on page 79 that *'The date the searches were conducted was not supplied, and no numbers of records retrieved were provided for either the NICE website or PubMed searches. No additional databases or other resources were searched.'*	We propose removing the relevant fragment.	The dates of the economic searches are stated – see titles of tables in *Appendix G. Published cost-effectiveness studies*.	Not a factual error. The ERG comments refer to the search methods in Appendix H/I. In Appendix H/I the company do not give the dates of the searches conducted.

Issue 17 Use of calibration factor in the model.

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
6. "Since the company only used the calibration factor on the belimumab arm of the model, it was not necessary to use the PSM to obtain the calibration factor. The PSM reduces the number of patients included with the aim to obtain a comparable set of patients to the TLC. As no comparative estimates for treatment	We suggest revising this fragment in line with the explanation provided.	Comparative estimates for the calibration factors were derived from the PSM study. The corresponding comparative ST estimate was higher than 1, implying that the model underestimated organ damage progression for the ST arm. The approach taken not to use the calibration factor for ST represents a conservative stance. If the relevant calibration factor was used in the ST arm, this	Not a factual error. The company are correct in pointing out that given that they had used the PSM, not applying the calibration factor to the ST arm is conservative. The ERG did not question this. The ERG just questions whether the calibration factor is reliable, since it is derived from the PSM the generalisability of which to the UK setting is very questionable.

effectiveness or long-term outcomes are derived, the purpose of using this reduced (matched) set of patients is questionable, in particular because the resulting set of patients is not necessarily more generalisable to the UK setting than the original sample" (Page 92)		would result in an increased rate of organ damage for this arm, reducing the ICER. Not applying the calibration factor for ST in the model, however, does not imply that it was therefore not necessary to base the belimumab calibration factor on the PSM. Without the PSM it would not be possible to know if the comparative ST calibration factor would be lower or higher than 1, and whether the ST arm organ damage progression was over- or underestimated in the model.	
7. "It is further questionable whether results of the PSM are applicable to the HDA-2 population." (page 92)	We suggest revising this fragment in line with the explanation provided.	The HDA-2 population has more active disease than the PSM-based population. The relative, comparative benefit of belimumab compared to ST has been shown to be greater in the HDA populations for many outcomes in the Phase 3 studies. It is reasonable to assume this would translate to increased benefit in terms of rate of organ damage progression and this would imply that the relative difference in calibration factors between belimumab and ST could indeed	Not a factual error. There is no evidence for the generalisability of the PSM to the HDA-2 population.

		be higher in the HDA-2 population. Therefore, adopting the PSM based calibration factor to the HDA-2 population could in fact be conservative for belimumab. In addition, many conservative steps were taken in applying the calibration to the model, i.e. not also adjusting the ST arm and restricting the organ damage benefit to 6 years in the base case.	
8. "Calibrating the existing model to one observed time-point (five years) is problematic, especially given the attrition of patients treated with belimumab in the BLISS LTE studies mentioned above and in Section 4.2.6." (Page 93)	We suggest revising this fragment in line with the explanation provided to the points above .		Not a factual error. See our responses to Issue 5.
9. The pre-calibration model already assumes a long-term treatment effect based largely on clinical expert opinion. In TA397, it is stated that "In the simulation	We suggest revising this fragment in line with the explanation provided.	The long-term treatment effect assumption already used in the model and validated by expert opinion concerns the treatment effect on sustained disease activity reduction. That is a	Not a factual error. The statement is correct. The company seems to be questioning the context, if anything. However, the ERG would argue that treatment effectiveness is related to both disease outcomes (disease activity and organ damage) and it is

model, an assumption was made that the additional absolute effect of belimumab on disease activity reduction remains constant after one year. This is a key model assumption and was discussed with Professor Petri who has observed patients on belimumab in her clinic for a number of years as part of the Phase 2 open-label extension study".93 The ERG clinical expert considers this a reasonable assumption. (Page 93)		different disease outcome than organ damage development. The indirect positive effect of sustained disease activity reduction on reduced organ damage progression (modelled through the natural history models) was shown in the model to still underestimate the impact of belimumab on organ damage reduction as per the calibration exercise based on the LTE data. Hence, there is no double adjustment of a long-term treatment effect.	relevant, given the uncertainty surrounding long-term treatment effectiveness on both outcomes, to assess assumptions for their potential bias and how this adds up.
10. "The pre-calibration model already assumes that the absolute SS value is adjusted in comparison to that predicted by JH. In TA397,44 it is stated that "The adjusted JH model was used to predict the SLEDAI	We suggest revising this fragment in line with the explanation provided.	The pre-calibration model adjusted the SELENA-SLEDAI score, which is a measure of disease activity. This is a different outcome to that of irreversible organ damage progression, to which the calibration factor was applied. As explained in the justification to the previous point,	Not a factual error. See issue above for explanation.

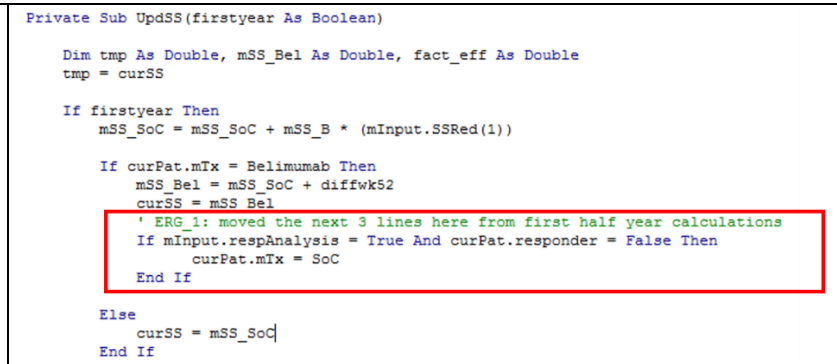
score of a patient treated with SoC after one year. The model allows for the selection of the original and adjusted JH model. For the base case analysis, the adjusted JH model has been used". In the final appraisal document of TA397, the committee also noted that "that there was uncertainty about whether the equations derived from the Johns Hopkins data could be reliably applied to the target population because of differences in study populations".93 Given that the adjustment favours the effectiveness of belimumab, it is likely that this adjustment causes the ICER to be underestimated." (page 93)		there is no double adjustment of a long-term treatment effect.	

Issue 18 Responder rule implementation in the model

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
3. In the model, the probability of being a responder is based on baseline SS score, which is linked to the responder criteria applied to patients in the BLISS trials (i.e., only for patients with a reduction of ≥ 4 points in SS score). Hence, it is estimated at baseline in the model and not directly linked to the actual improvement in SS score in the model. In turn, actual SS scores are	Discard this issue as it is not a model error (see explanation).	Response is determined at 24 weeks and is based on the 24-week SS score data from the BLISS trials for the HDA subgroups. Belimumab patients who do not respond at 24 weeks (i.e. those that do not have at least a 4 or more decrease in SS score from baseline to week 24) are discontinued from belimumab in the model and continue as a ST patient. Corresponding costs in the model for belimumab treatment are only applied for the first 24 weeks (sub function costs in clsPatient copied below, relevant section highlighted) and this represents what would happen in clinical practice. Hence, there is no underestimation of costs in the model. Non-responders at 24 weeks could indeed have been responders at 52 weeks, but that is not an option in the model as they have already been taken off belimumab treatment at week 24 as per the responder rule. So indeed, non-responders on belimumab (at week24) could still have an SS point improvement of more than 4 points at 52 weeks, following their switch to ST at week 24.	Not a factual error. The company's explanation is not addressing the point, which is that some patients are designated non-responders at 24 weeks despite achieving response at 52 weeks (and the concern is that if it were modelled it could likely be shown that some of those fulfilled the response criteria at 24 weeks as well). This issue is a result of how the company modelled whether a patient is a responder irrespective of their characteristics. If non-responders in the model who discontinue treatment actually have a response, then this is likely not in line with clinical practice and may under-estimate cost.

estimated based on a regression model where response is an independent variable, given that a 24-week time point does not exist in the model. As a result, a large proportion of patients is classed as non-responders but experiences >4 points reduction in SS at 52 weeks. This could lead to under-estimation of costs in the model compared to clinical practice as patients with no response do not continue		'short term costs
		Call mCosts.AddShort(curTime - 1, curcost)
		
		'medication costs
		If mTx = Belimumab Then
		Call mCosts.AddMed(curTime - 1, 1)
		ElseIf mTx_pr = Belimumab And NatDisc = True Then
		Call mCosts.AddMed(curTime - 1, 0.5)
		ElseIf mTx_pr = Belimumab And responder = False Then
		Call mCosts.AddMed(curTime - 1, 24 / 52)
		End If	

belimumab. A structural adjustment to the company's model is necessary to resolve this issue. (page 94)			
4. Due to a modelling error, belimumab non-responders have the same reduction in disease activity at 52 weeks as patients in the standard therapy arm. (page 95)	We suggest revising this fragment in line with the explanation provided and discarding changes to the CS model to overcome the perceived error in the modelling of non-responder disease activity.	Response is assessed at 24 weeks, not at 52 weeks. In the first year of the model, belimumab non-responders at 24 weeks are discontinued from belimumab immediately and changed to ST. This means that they are on ST at 52 weeks (and no longer on belimumab) and have been for 28 weeks, which means that the use of the ST disease reduction is appropriate for belimumab non-responders at 52 weeks. The ERG change to the model (as per the below) needs to be undone, as no error needs correcting, and in turn the implemented change introduces an error to the programming routine. Non-responder patients need to be dropped to SoC in the original routine called UpdSSFirstHalfYear, as that is the timepoint when non-responders are dropped off from belimumab, not at 52 weeks.	Not a factual error. The company's modelling is not in line with the CS and not in line with the evidence presented in the CS (patients on ST can have better response than belimumab non-responders). It is therefore an error. If the ERG introduced another error by correcting the company's error, we request that the company fix their error in a way that does not introduce another error and provide clear explanation of what the error would be and how it was avoided as well as detail on the implementation.

Issue 19 Sampling of organ damage and death occurs after allocation to treatment.

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
2. In the VBA, first, a simulated patient is allocated to a treatment and organ damage and death are only sampled within the treatment arm. This leads to the same simulated patient (same age, gender, SS score) experiencing differential organ damage and times of death only dependent on allocation to treatment arm but not caused by this	This is not an issue, see explanation provided.	Organ damage and mortality depends, amongst other things, on AMS (covariate included in the time to event equations), which depends on the SELENA-SLEDAI score. In turn, the progression of SELENA-SLEDAI score depends on treatment allocation. So, contrary to what is stated, the occurrence and timing of organ damage and mortality does depend on treatment allocation. Therefore, organ damage items involved, and death cannot be sampled before treatment allocation. It would forego the benefits of belimumab on reduced organ	Not a factual error. The same simulated patient has differential organ damage scores. Ours was an accurate description of the problems this can cause.

allocation (so just because of sampling). This induces noise in the model and makes validation difficult. (page 108, in the table, and the entire more detailed text under e))		damage development and mortality. It is true that this introduces additional sampling error. It could also mean you can sometimes get extreme differences between the same patient in both arms. This is all caused by sampling error. , but no solution to that has been identified. Organ damage and death events reported as in year 0 of the patients traces occur in the first year of the model, so after model entry. Baseline organ damage (before model entry) and all other baseline characteristics for that matter, are sampled before treatment allocation and patients are cloned for both arms. Noise, or sampling error has been checked and mitigated by simulating enough patients. Convergence plots were generated and checked to make sure model results are stable and sampling error reduced to a minimum. To ensure stable results, free from sampling errors, convergence checks had been performed for the original model. In checking different seeds and checking different amounts of patient samples, it was found that simulating 50,000 patients in each arm led to stable results, free from	

		noise. This also holds true for the updated model.	

Issue 20 Incorrect confidentiality marking

Location of incorrect marking	Description of incorrect marking	Amended marking	ERG Response
Belimumab ERG report P54	Baseline characteristics from HDA-2 group from BLISS-SC and pooled BLISS-52/76 trials are not given academic in confidence.	In all BLISS trials, the vast majority of patients were female: over [redacted] in BLISS-52 and BLISS-76, and over [redacted] in BLISS-SC. While this reflects the gender distribution of SLE, it limits statistical power to determine whether belimumab is as effective in men as in women, especially for the SC formulation. The ERG asked for the results of subgroup analyses by gender (Response to clarification, Question A29).[6] The interaction between treatment and gender was included in the logistic regression model for the primary endpoint, SRI-4. For BLISS-SC there was no significant interaction between treatment and gender but there were only [redacted] male patients included. In men the response rates were [redacted] for placebo and [redacted] for belimumab (OR [redacted], 95% CI [redacted]) and the corresponding response rates in women were [redacted] for placebo and [redacted] for belimumab (OR[redacted], 95% CI [redacted]). This indicates that belimumab was beneficial in women but there was a lack of evidence for men. In the analysis of the pooled BLISS-52 and -76 data the treatment x gender interaction was also not significant in the model. SRI-4 response rates in men were [redacted] for placebo [redacted] for belimumab 10 mg (OR [redacted], 95% CI [redacted]). The corresponding results for women were [redacted] for placebo and [redacted] for belimumab 10 mg (OR [redacted], 95% CI [redacted]). Both sets of results indicate that belimumab was beneficial in women but there was a lack of evidence for men. Further subgroup analyses of belimumab-SC showed little difference in effect by body weight or body-mass index (these analyses had reasonable power), and a subgroup analysis for age compared patients above and below 65 years but had limited statistical power, as only [redacted] participants were 65 years or older. Further subgroups analyses of belimumab-IV (10 mg/kg) showed little difference in effect between people with more or less than 15 U/mL baseline	This has been amended.

		anti-Sm, more or less than 1.64 baseline ANA titre, baseline immunosuppressive use, or being younger or older than 45 years (all analyses had reasonable statistical power). Patients with White/Caucasian, Asian and Alaska Native/American Indian ethnicities had similar positive effects of belimumab-IV, though Black/African American patients did better on placebo, though the difference was not statistically significant.	
Section 4.3.2, page 72 of the ERG Report	BILAG-BR data not marked AIC	The analysis included [redacted] patients receiving belimumab and [redacted] receiving rituximab. The two primary endpoints for SLE were BILAG-2004 and SLEDAI-2K, a disease specific instrument (SLICC/ACR damage index) was also used and HRQoL was measured using both generic and disease-specific instruments. Patients were recruited from 39 centres across the UK and followed up at three, six and 12 months then every 12 months. At any point they could stop treatment and restart with a second round of the same treatment or restart with a different one. The analysis used multilevel repeated measures regression modelling of outcomes at three, six and 12 months. Potential confounding variables were identified before analysis and included as covariates in the models. The results showed that there was [redacted] between belimumab and rituximab in SLEDAI-2K at 12 months of follow-up in the model adjusted for all covariates (mean difference [MD] -[redacted]), [redacted] in BILAG total score ([redacted]).	This has been amended.
Section 4.6 page 77	BILAG-BR data not marked AIC	The analysis included [redacted] patients receiving belimumab and [redacted] receiving rituximab	This has been amended.
Section 3.5 page 37	The SC formulation has not been launched therefore list prices for the SC formulation (and for the single pen and 4-pack sizes) are not public and are	"However, the company have received confirmation from the DHSC that they have agreed to a list price for this formulation of "[redacted] for Benlysta (belimumab) 1 pack of 4 sub-cut pens"'."	This marking has been amended by NICE technical team.

	currently commercially sensitive.		
Section 5.2.9 page 100	The SC formulation has not been launched therefore list prices for the SC formulation (and for the single pen and 4-pack sizes) are not public and are currently commercially sensitive.	"A difference was made between the two belimumab formulations – IV and SC. The company proposes a patient access scheme for both formulations [redacted]. The IV formulation requires 14 infusions in year 1 and 13 in every subsequent year, with administration cost of £154 per infusion. Two different list prices for the SC formulation have been recorded. The SC list price model received by the ERG puts the list price at £[redacted] per pen (£[redacted] per four pen package). A recent update on the list price puts the list price at £[redacted] per pen (£[redacted] per four pen package). Further, the model assumed vial wastage and the cost for the IV population to be weight-dependent. In comparison, the company assumes only one hour of specialist nurse time to teach the patient how to self-administer the drug. As the formulation is not weight-dependent, no vial wastage is assumed. This results in drug-related costs of £[redacted] for the first year and £[redacted] for subsequent years for the IV formulation and £[redacted] and £[redacted] for the SC formulation (with the patient access scheme and administration cost included)."	This marking has been amended by NICE technical team.

References

1. Rees F, Doherty M, Grainge M, et al. The incidence and prevalence of systemic lupus erythematosus in the UK, 1999-2012. Ann Rheum Dis. 2016;75(1):136-41.
2. Merrill JT, Neuwelt CM, Wallace DJ, et al. Efficacy and safety of rituximab in moderately-to-severely active systemic lupus erythematosus: the randomized, double-blind, phase II/III systemic lupus erythematosus evaluation of rituximab trial. Arthritis and rheumatism. 2010;62(1):222-33.

3. Morand EF, Furie R, Tanaka Y, et al. Trial of Anifrolumab in Active Systemic Lupus Erythematosus. New England Journal of Medicine. 2019;382(3):211-21.
4. Teng YO, Parikh SV, Saxena A, et al. O11 AURORA phase 3 study demonstrates voclosporin statistical superiority over standard of care in lupus nephritis (LN). Archives of Disease in childhood; 2020.
5. Urowitz MB, Ohsfeldt RL, Wielage RC, et al. Organ damage in patients treated with belimumab versus standard of care: a propensity score-matched comparative analysis. Annals of the Rheumatic Diseases. 2019;78(3):372.
6. GlaxoSmithKline. Belimumab for treating active autoantibody-positive systemic lupus erythematosus [ID1591] – Response to request for clarification from the ERG. GlaxoSmithKline, 2020.

NICE National Institute for Health and Care Excellence

Technical engagement response form

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (Review of TA397) [ID1591]

As a stakeholder you have been invited to comment on the ERG report for this appraisal. 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.

We need your comments and feedback on the key issues below. You do not have to provide a response to every issue. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be included in the papers in full and may also be summarised and presented in slides at the appraisal committee meeting.

Deadline for comments **5:00pm on 11 March 2021**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the ERG report which summarises the background and submitted evidence, and presents the ERG's summary of key issues, critique of the evidence and exploratory analyses. This will provide context and describe the questions below in greater detail.
- Please ensure your response clearly identifies the issue numbers that have been used in the executive summary of the ERG report. 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.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, 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.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**GlaxoSmithKline**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**

Key issues for engagement

Please use the table below to respond to questions raised in the ERG report on key issues. You may also provide additional comments on the key issue that you would like to raise but which do not address the specific questions.

Key issue	Does this response contain new evidence, data or analyses?	Response
Key issue 1: Evidence is missing for specific populations, such as children and patients with severe active central nervous system (CNS) lupus.	**NO**	**<u>Paediatric patients</u>** Clinical trial evidence evaluating the safety, efficacy and pharmacokinetics on the use of intravenous (IV) belimumab in children with systemic lupus erythematosus (SLE), derived from Part A of the Phase 2 PLUTO (The Pediatric Lupus Trial of Belimumab Plus Background Standard Therapy) study, was presented in Appendix O of the company submission, comprising of the primary and secondary endpoints and trial safety data. Given the severity and relatively low prevalence of childhood SLE, a double-blinded, placebo-controlled study with a large sample size powered for statistical significance testing was deemed unfeasible. Hence, this study was designed to descriptively evaluate efficacy and safety of belimumab in childhood onset SLE and its study design and endpoints were similar to previous Phase 3 IV belimumab studies in adults, as was the 10 mg/kg dose. As outlined in the belimumab IV European Public Assessment Report (EPAR)[(1)], the European Medicines Agency (EMA) was satisfied that there was sufficient efficacy and safety data to support the use of belimumab in children with SLE. For completeness, since the submission to NICE, we have also provided additional analyses of the primary endpoint and its individual components from Part A of the PLUTO study, for the HDA-1 and HDA-2 subgroups to demonstrate a comparable efficacy to the adult population (see Tables 1 and 2 below).

Table 1: PLUTO Paediatric IV Study - SRI-4 Response Rate at Week 52 - HDA-1 Subgroup

	Placebo N=13	Belimumab 10 mg/kg N=16
Response (primary efficacy analysis)[a], n (%)	5 (38.5%)	8 (50.0%)
Observed difference vs. placebo	-	11.54%
Odds ratio (95% CI)[a] vs. placebo	-	1.38 (0.29, 6.44)
4-point reduction in SELENA SLEDAI, n/n (%)	5/13 (38.5%)	9/16 (56.3%)
Odds ratio (95% CI)[a] vs. placebo	-	1.90 (0.41, 8.83)
No worsening in PGA, n/n (%)	6/13 (46.2%)	10/16 (62.5%)
Odds ratio (95% CI)[a,b] vs. placebo	-	1.58 (0.34, 7.42)
No new 1A/2B BILAG domain scores, n/n (%)	6/13 (46.2%)	10/16 (62.5%)
Odds ratio (95% CI)[a,c] vs. placebo	-	2.46 (0.47, 12.95)

Table 2: PLUTO Paediatric IV Study - SRI-4 Response Rate at Week 52 - HDA-2 Subgroup

	Placebo N=21	Belimumab 10 mg/kg N=27
Response (primary efficacy analysis)[a], n (%)	10 (47.6%)	16 (59.3%)
Observed difference vs. placebo	-	11.64%
Odds ratio (95% CI)[a] vs. placebo	-	1.65 (0.51, 5.32)
4-point reduction in SELENA SLEDAI, n/n (%)	10 (47.6%)	17 (63.0%)
Odds ratio (95% CI)[a] vs. placebo	-	2.00 (0.61, 6.56)
No worsening in PGA, n/n (%)	12 (57.1%)	18 (66.7%)
Odds ratio (95% CI)[a,b] vs. placebo	-	1.60 (0.47, 5.49)
No new 1A/2B BILAG domain scores, n/n (%)	12 (57.1%)	18 (66.7%)
Odds ratio (95% CI)[a,c] vs. placebo	-	1.71 (0.50, 5.82)

Currently, there is no clinical data to support the use of belimumab subcutaneous (SC) formulation in patients <18 years old and therefore the marketing authorisation for belimumab in patients aged 5-17 years old is currently restricted to the IV formulation only.

Therefore, due to limited available clinical data (i.e. no long-term evidence or data on the use of belimumab SC in children), an economic evaluation was not conducted on paediatric patients due to uncertainty of modelling on key parameters; all inputted data pertains to an adult population.

CNS lupus

The current marketing authorisation, as per Summary of Product Characteristics for belimumab IV[(2)] and SC[(3)] formulations, does not include patients with severe active CNS lupus, as these patients were not included in BLISS pivotal IV and SC trials. Therefore, we acknowledge there is no evidence for the use belimumab in this patient population and do not anticipate NICE to issue a guidance relating to the use of belimumab in patients with severe active CNS lupus.

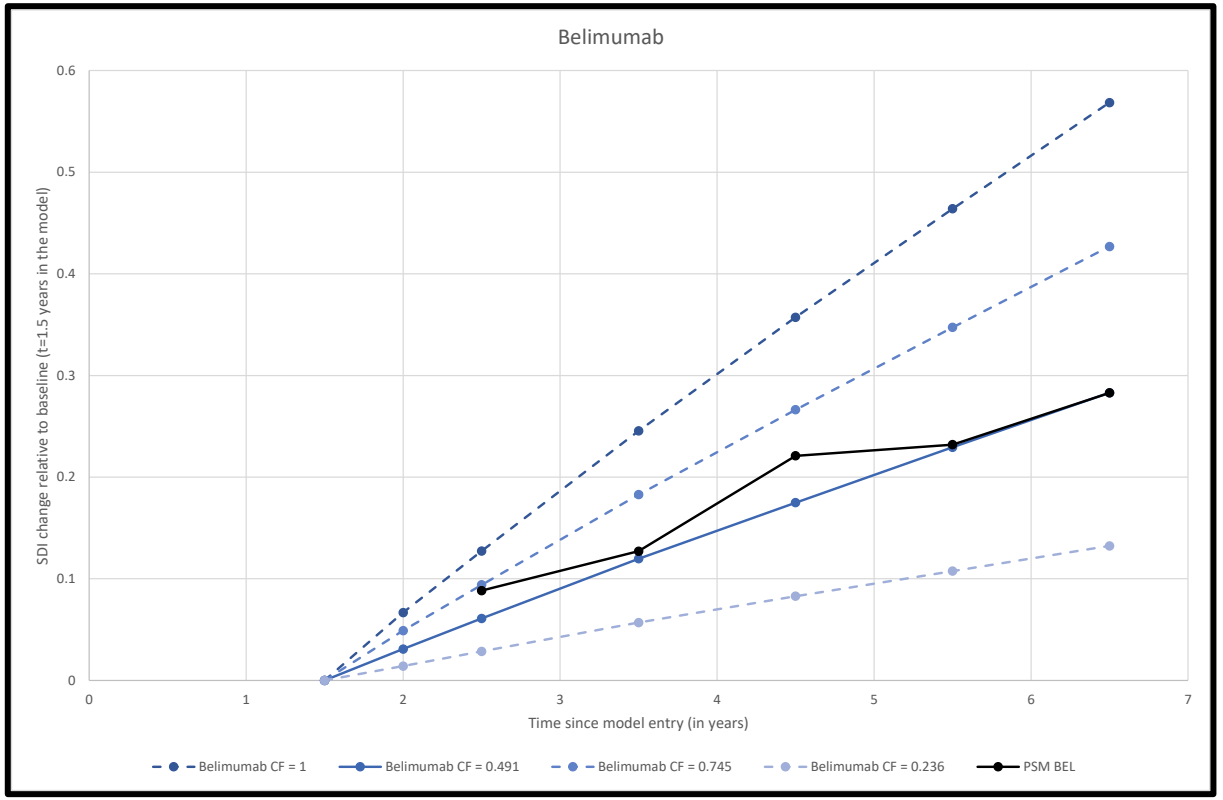
Key issue 2: Some comparators listed in the NICE scope were not included	**NO**	The ERG has acknowledged the data limitations which preclude a meaningful comparison between belimumab and rituximab due to the lack of a robust comparable dataset to the BLISS pivotal trials. Similarly, the ERG and clinical expert consulted acknowledge that cyclophosphamide is not an appropriate comparator for belimumab. This is aligned with clinical advice we received and to the best of our knowledge on the available evidence base.
Key issue 3: Short follow-up in the main comparative trials (BLISS-SC, BLISS-52 and BLISS-76)	**NO**	We acknowledge this point but would like to highlight that the 52 and 76-week follow up periods in the Phase 3 studies are in line with EMA guidance[4] and other comparative trials in SLE, such as AURORA,[5] EXPLORER,[6] and TULIP 1/2.[7] Furthermore, the company submission (Document B, section 2.6 and appendix M) we have presented 7 and 8 years of follow-up data from the BLISS-76 US and BLISS-52/76 non-US LTE studies respectively, and up to 13 years of follow-up data from the Phase 2 LTE study (LBSL-02) which demonstrate that the efficacy of belimumab is maintained and the drug is well tolerated.

Key issue 4: Using the propensity score-matching (PSM) analysis in calibrating the cost-effectiveness model can severely bias the results in favour of belimumab	**NO**	The ERG report raised the following key issues: 1. Patients are not necessarily representative of either BLISS-76 US LTE patients or Toronto Lupus Cohort (TLC) patients, but representative only of patients matching between these cohorts: this is unlikely to be representative of patients in the UK. 2. The application of the calibration factor derived at 5 years to the whole 5 years. 3. The calibration factor derived from the PS-matched comparative analysis of 0.491 effectively doubles the effectiveness of belimumab for preventing organ damage, compared with the Johns Hopkins Natural History Model (JH NHM). We will attempt to address these issues raised. ***1. Patients are not necessarily representative of either BLISS 76 US LTE patients or TLC patients, but representative only of patients matching between these cohorts: this is unlikely to be representative of patients in the UK.*** We acknowledge the comment made by the ERG that the population presented in the PS-matched comparative analysis could be seen to be representative of only the patients matched between the BLISS 76 US LTE study and the TLC and not a broader SLE population. However, we believe that the PS-matched comparative analysis was conducted robustly and that it is relevant to SLE patients in the UK for the reasons outlined below. The matching cohort, TLC, was identified systematically, based on pre-defined cohort characteristics (including captured outcomes, cohort size and severity). Despite availability of data, patients captured in the TLC preceding 1990 were excluded to maximise comparability across the groups. Furthermore, those with ≥ 15 years of follow-up were also excluded. A strength of using the TLC was that patients who were otherwise indicated for treatment did not receive belimumab simply because it was not available which enforces the notion that other than the availability of belimumab, all other clinical characteristics should or could be considered comparable. As is the case with PS-matched studies, it was not possible to match on all possible predictor variables for organ damage progression in both the BLISS-76 US LTE dataset and the TLC dataset. After discussion

		with clinicians internal to GSK and Professor Urowitz, it was agreed that the predictors for organ damage progression identified in the literature were not all independent. The study team believed that by matching on key predictors that could be identified, this may also indirectly include many of the other unobserved variables by proxy. Sensitivity analyses were conducted to test the sensitivity of the results to the source of the belimumab population and the methods for conducting the PS-matched comparative analysis. Whilst the primary analysis was conducted on the BLISS-76 US LTE cohort only, which allowed matching on most predictors, a secondary exploratory analysis was performed on the more geographically dispersed pooled BLISS 52 and 76 LTE (i.e. US and non-US cohorts) population. The results of the pooled analyses of US and non-US patients were also similar. The results were also robust to the weighting approach applied, differences in baseline corticosteroid dose and study entry date. Two clinical experts were consulted to understand the generalisability of the BLISS-76 US LTE cohort to the UK SLE population used in the PS-matched comparative analysis. They believed that the BLISS-76 US LTE extension study population clinically reflected UK SLE patients based on baseline characteristics. Furthermore, they also agreed that the BLISS-76 US PS-matched cohort was also clinically reflective of a UK SLE population. This was based on a review of baseline characteristics of the matched cohort and BILAG-BR. Additionally, the experts acknowledged that although there were differences geographically, the results could still be extrapolated to a UK SLE patient population. Further, in terms of differences in severity of disease, although UK patients eligible for belimumab may have more active disease (compared with the matched TLC cohort), the experts believed that as patients with higher disease activity demonstrated a greater response to belimumab compared with the overall study population, then the positive benefit seen from the PS-matched comparative analysis in slowing down organ damage progression could be extrapolated to the UK belimumab eligible SLE population in those who respond and continue on belimumab. ***2. The application of the calibration factor derived at 5 years to the whole 5 years.*** The PS-matched comparative analysis provided the opportunity to incorporate the long-term benefit of belimumab on reducing the progression of organ damage into the cost-utility models. We acknowledge

		there were limitations in incorporating this evidence (for example, without access to change in SDI, specifically for HDA-1 and HDA-2 populations), and therefore we were cautious in our approach. Firstly, the model was re-simulated to reflect the baseline characteristics of the matched cohort. This showed that the model underestimated the benefit that belimumab was having on organ damage progression (as measured by the SLICC/ACR damage index (SDI)). Since the single calibration factor in use in the cost-effectiveness model is based on the total SDI change from baseline at 5 years, we explored the possibility that the calibration factor could be different had a different time point been selected on which to apply to in the model. We compared SDI scores at Years 1-4 (obtained from the regression equations in Section 6.2.5 in the PS-matched analysis clinical study report (tables 105-108)) (solid black line) to those obtained when applying the single calibration factor (0.491) in the cost-effectiveness model (solid blue line) (**Error! Reference source not found.**). The applied calibration factor of 0.491 was determined by anchoring the model simulation estimates on the 5-year change from the PS-matched analysis, being the primary outcome of the study. As seen in **Error! Reference source not found.**, with a calibration factor of 0.491, the estimated model results are reflected well for years 2, 4 and 5, and slightly below at years 1 and 3 of the PS-matched results for belimumab (see solid black line). Since the calibrated values do not show a systematic under- or over-estimation of SDI scores over the years, we concluded that the use of a single calibration factor value applied for each of the years in the model was appropriate. The model is calibrated to predict the 5-year SDI increase in line with what is observed in the PSM analysis using real world data for belimumab and ST from the LTE belimumab studies and the Toronto Lupus Cohort, respectively. It is unclear how the SDI increase after 5 years should be extrapolated beyond this point. Hence, the most conservative assumption was made to apply the calibration factor for a maximum of 6 years.

Figure 1. Time since model entry versus SDI change relative to baseline for calibration factors associated with belimumab



PSM BEL: (solid black line) total SDI change from baseline at each year from the PS-matched comparative analysis; each data point represents one year of the PS-matched analysis.
Belimumab CF = 0.491: (solid blue line) modelled total SDI change from baseline at each year with calibration factor applied to belimumab responders

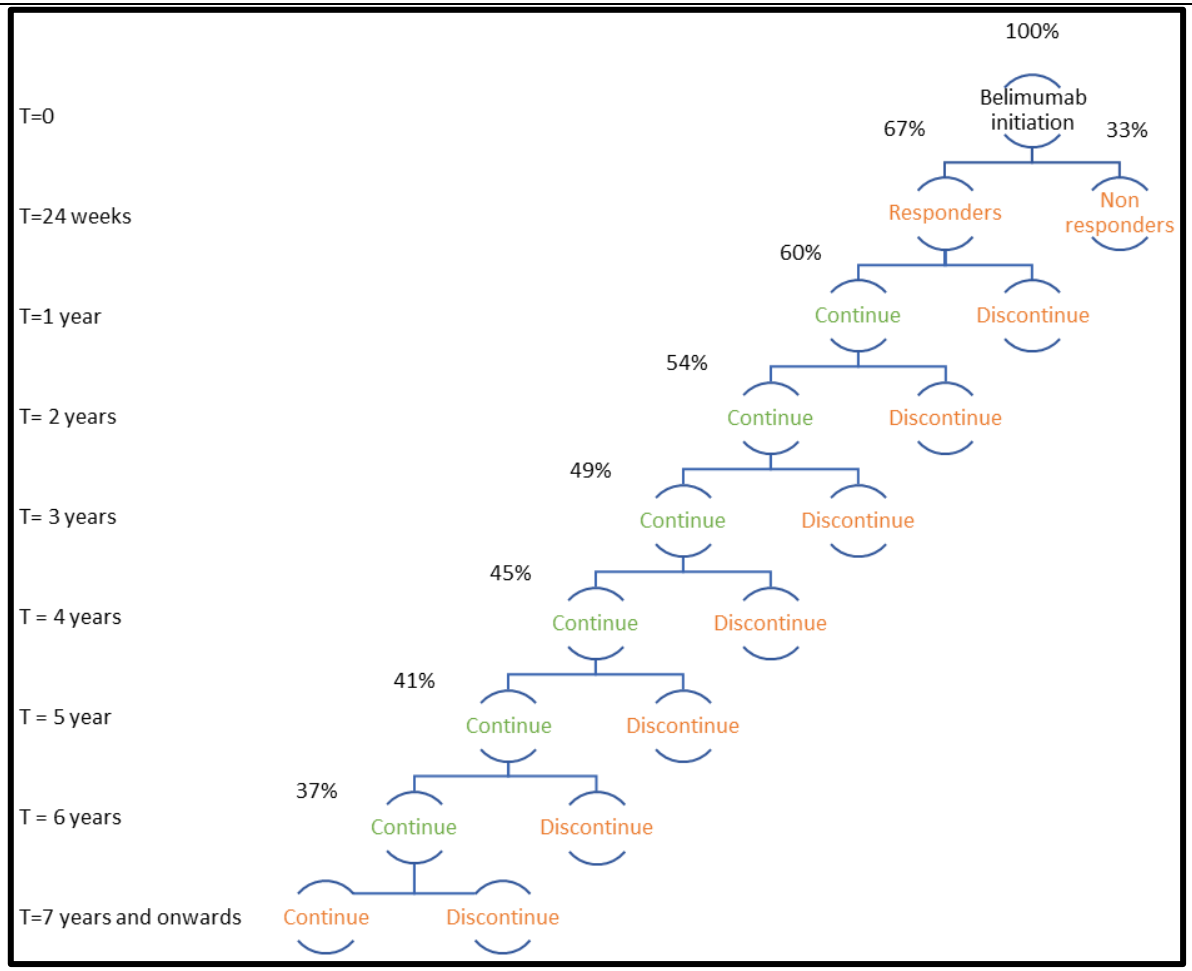
		3. The calibration factor derived from the PS-matched comparative analysis of 0.491 effectively doubles the effectiveness of belimumab for preventing organ damage, compared with the JH model. We acknowledge that the application of the calibration factor to the responders in the belimumab arm in the economic models does change the determinants of the cost-effectiveness of belimumab compared with ST, a key change from TA397. However, there is good reason why this is the case and why clinically it is a reasonable approach for implementation. Firstly, in TA397 Final Appraisal Document, it states *"…the company may have underestimated some of the benefits associated with delaying certain types of organ damage."* The primary measure of organ damage progression is a change from baseline in SDI. The SDI contains items that represent permanent, irreversible damage in a lupus patient. Items should be present for at least 6 months (with the exception that manifestations such as myocardial infarction and stroke are recorded once they occur). Damage is defined for 12 organ systems: ocular (range 0–2), neuropsychiatric (0–6), renal (0–3), pulmonary (0–5), cardiovascular (0–6), peripheral vascular (0–5), gastrointestinal (0–6), musculoskeletal (0–7), skin (0–3), endocrine (diabetes) (0–1), gonadal (0–1), and malignancies (0–2). Damage over time can only be stable or increase, to a maximum of 47 points. It is important to note the definition of SDI and in doing so, understanding that the impact of belimumab on organ damage progression was unlikely to have been realised in the pivotal short-term BLISS 52, BLISS 76 and BLISS SC studies. The former two studies underpinned the economic analysis included in TA397. Organ damage progression included in the previous analysis relied on a relationship between disease activity (SS) and organ damage (SDI) based on patients observed in the Johns Hopkins Cohort. This cohort would not be able to capture any improvement in organ damage progression; belimumab is the first ever licensed treatment for SLE with evidence that it directly impacts the progression of organ damage. The PS-matched comparative analysis therefore allows us to utilise the long-term change in SDI seen in patients on belimumab (compared with a matched cohort) and more appropriately model these benefits in a way that was not previously possible.

		Following the Technical Engagement call, our interpretation was that our commentary in the CS had led to some confusion as to how the calibration factor is applied in the model to ensure that the impact that belimumab has on organ damage progression is appropriately modelled. The BLISS 52 and 76 LTE (non-US and US cohort) data provides long-term data on the change in organ damage over time for patients who continue to receive benefit from belimumab beyond the randomised controlled trial period (up to 5 years of follow-up data was used in the PS-matched comparative analysis). The cost-effectiveness model distinguishes between responders and non-responders and only those who remain on therapy (responders at Week 24) receive the treatment effect of belimumab, including the calibrated results of the benefit on reducing organ damage progression. See **Error! Reference source not found.** from the VBA code routine "Update" in the clsOrgans class module of the code where this happens. The calibration factor depends on time (curTime) and the current treatment the patient is on (curpat.mTx) at that time, and is not to be mistaken with the treatment arm (curTx). Non-responders are taken off treatment, changing curpat.mTx to ST. So, the routine pulls from the ST calibration factor array for non-responders. The same holds true in case of natural discontinuation on belimumab, after which also the ST calibration factor will be applied moving forward. The definition of responders used in the cost-effectiveness model is reflective of the reimbursement criterion following the previous NICE guidance, and in line with the clinical guidance that patients will only continue on belimumab for as long as their clinician believes they are receiving benefit.

Figure 1. Routine "Update" in the clsOrgans class module code

```
Dim p_suc As Double


For i = 1 To mInput.nrOrgans
    If mAffected(i, 2) = False Then
        Xb = XBeta(i)
        SurvType = mInput.SurvType(i)

        Select Case SurvType
            Case dEXP
                Xb = Exp(Xb)
                hazard = HExponential(Xb, curTime, curTime - timeStep)
            Case dWEIB
                Xb = Exp(Xb)

                par2 = mInput.SurvPar2(i)
                hazard = HWeibull(Xb, par2, curTime, curTime - timeStep)
            Case dLOGLOG
                Xb = Exp(-Xb)

                par2 = mInput.SurvPar2(i)
                hazard = HLogLog(Xb, par2, curTime, curTime - timeStep)
            Case dGOMP
                Xb = Exp(Xb)

                par2 = mInput.SurvPar2(i)
                hazard = HGompertz(Xb, par2, curTime, curTime - timeStep)
            Case dLOGNORM
                'to be added or changed with other survival curve
        End Select

        p_suc = 1 - Exp(-hazard)

        'Adjustment with calibration factor
        p_suc = p_suc * mInput.calibFactor(curPat.mTx, curTime)

        If RandBern(p_suc) = True Then
            mAffected(i, 2) = True
            mDates(i) = curTime - 1

            'SLICC score
            mSLICC = mSLICC + mInput.SLICC(i)
        End If

    End If
Next

sliccOverTime(curTime) = mSLICC
```

		Therefore, in the belimumab arm in the model, patients stopping belimumab and returning to standard treatment only, due to being categorised as a belimumab **non-responder** at week 24 are not subject to the calibration factor at all, see **Error! Reference source not found.**. Furthermore, belimumab responder patients will only have the calibration factor applied from one year and until their time of withdrawal from belimumab or to a maximum of 6 years whichever occurs earliest. **Figure 3. Application of calibration factor to the belimumab arm in the economic model**



NOTE: The green 'Continue' denotes the application of the calibration factor is applied only to the proportion of belimumab patients who remain on treatment from 1.5 years to 6.5 years

So, working this through, the calibration factor for patients who continue on belimumab (responders) is 0.491 until year 6 inclusive. The calibration factor for patients who remain on belimumab year 7 onwards,

		and for patients who discontinue belimumab, for whatever reason, is 1. Percentages of patients still on belimumab and hence to which the calibration factor of 0.491 apply, do not account for death. The PS-matched comparative cohort was based on the population who continued on belimumab in BLISS 76 US LTE. Whilst not wanting to exclude 10 years of evidence evolution since the commencement of TA397, we believe we have been cautious with our approach in response to recognising a number of limitations. The model validation and calibration exercise were re-simulated with the baseline characteristics of the matched population, we have applied the resultant calibration to: o A subset of high disease activity patients (HDA-1 and HDA-2 populations) who demonstrated a greater response to belimumab in the Phase 3 studies compared with the ITT population. If this greater response to belimumab also translates to a greater benefit on delaying organ damage progression (as suggested by UK experts consulted), this could mean that the long-term effectiveness and cost-effectiveness of belimumab is underestimated in these high disease activity populations. o The calibration factor has also been applied to the SC model, with assumed comparability to the belimumab IV formulation. As a result of these limitations, we remain conservative in the way in which the calibration has been applied to the IV and SC models: • The resultant improvement in organ damage accrual is applied to the belimumab **responders** of the belimumab arm only. So, over time to 6 years, there is a 'dilution of effect' as patients discontinue belimumab. • The calibration was only applied for patients remaining on belimumab (i.e. responders) and was applied from entry to the model and continued until either patients discontinued belimumab due to natural discontinuation or until patients had received belimumab up to and including Year 6, whichever was earliest, despite there being longer-term data to support continued treatment effect out to 13 years [8].

		• The model validation suggested that organ damage accrual for those patients in the ST arm was underestimated (in which TLC patients had a higher increase in SDI compared with the John Hopkins Natural History Model). This has not been included which could further underestimate the benefit of belimumab. It is well understood that uncontrolled disease activity and cumulative corticosteroid dose are key contributors of irreversible organ damage and subsequent increased healthcare costs. Published long-term evidence has demonstrated that a large proportion of patients receiving belimumab long-term are able to reduce their steroid dose. Due to the known harmful side effects of steroids, clinical experts confirmed they would look to reduce the steroid dose as soon as their disease has been adequately controlled with belimumab. Furthermore, the available clinical evidence demonstrates the benefits of belimumab in reducing disease activity. This, in combination with the potential to reduce your steroid dose, has a positive effect on slowing down organ damage accrual. Therefore, we believe that the results of the application of the PS-matched analysis study to the economic model is appropriate and clinically plausible given that belimumab has potential steroid sparing effects.
Key issue 5: Data from the BILAG Biologic Register (BILAG BR) are not suitable for a comparison of belimumab with rituximab	**NO**	We acknowledge this point from the ERG which is in line with our approach in our company submission in which we explained that a formal indirect treatment comparison of rituximab and belimumab based on BILAG-BR data was not conducted due to the observational, exploratory nature of the data and the differences in cohort sizes, patient characteristics and duration of follow-up. This is also relevant to key issue 2, as BILAG-BR data are another (beyond the available RCT evidence) component of the evidence base for belimumab and rituximab that does not permit a reliable comparison between the two drugs.
Key issue 6: Rituximab + standard therapy was not included as comparator in the model	**NO**	We acknowledge this point from the ERG. Rituximab could not reliably be included as a comparator to belimumab within the economic model for several reasons. Relative to the BLISS trials, the primary rituximab trials had differing endpoints, patient populations and trial design. For example, the inclusion criteria of the published Phase 2/3 randomised, double-blind study of rituximab required SLE patients to have significantly active disease at screening [(6)] likely to correspond to a more severe patient population than the Phase 3 belimumab trials. Also, changes in SELENA-SLEDAI, an important short-term outcome which can be linked to longer term impact on organ damage, were not collected in the rituximab trial, making an indirect comparison difficult [(6)].

		Following on from TA397, the update to the systematic literature review did not identify any studies that directly compared belimumab with rituximab. Differences in the patient populations and measurement in end points from previous studies still precludes the conduct of any meaningful indirect and mixed treatment comparisons between belimumab and rituximab. It is also important to note that there is very limited published long-term follow-up data for rituximab, particularly in relation to the effect of this treatment on the progression of organ damage. As this disease manifestation is a key component of the health economic modelling it would not be possible to make any robust assessment on the comparative long-term effectiveness and cost-effectiveness between belimumab and rituximab. The data available from the BILAG-Registry shows that since 2016 (and before the NHS E policy of July 2020[9]), there is an overlap of patients i.e. some patients who receive rituximab would be eligible to receive belimumab. The patient characteristics are shared in Section B.2.3.4. As part of the analysis of the BILAG-BR, the University of Manchester did undertake a multilevel regression modelling exercise to explore the patient outcomes across the three cohorts: belimumab, rituximab and non-biologic. In these regression models, treatment effect estimates are compared with rituximab (reference) due to the availability of the largest sample size and results are reported as effect coefficients. The results suggest that for most health outcome measures patients in the belimumab cohort demonstrate a similar level of improvement to rituximab. However, the regression modelling could only be conducted out to 12 months because of the limited follow-up data available for belimumab patients in this study. It remains that reducing the risk of long-term organ damage is a key treatment goal for SLE patients and of most interest to clinicians. Whilst there is published data to support this for belimumab, there is limited equivalent evidence on impact for rituximab. This precludes conducting a robust indirect treatment comparison, and which could also not be resolved by the data captured through BILAG-BR due to the observational, exploratory nature of the data and the differences in cohort sizes, patient characteristics and duration of follow-up. As GSK does not feel that a robust comparison can be made with rituximab, we have concentrated on the comparator of ST only in this economic analysis.
Key issue 7: IV and SC formulations are not compared with each other, as two	**NO**	The models for IV and SC were presented separately to better reflect the evidence. This approach does not affect the information that may be derived from the models, it only precludes a direct comparison of model outcomes between the two formulations.

separate model files are provided.		Aside from the cost utility models presented, the indirect treatment comparison of the belimumab IV and SC formulations was included (company submission; Section B 2.6.5.2) [10]. Briefly, a network meta-analysis (NMA) was conducted, utilising patient level data in a fixed effects model which showed that belimumab IV and SC were found to have similar efficacy for the percentage of patients with an SRI-4 response, ≥4-point reduction in SELENA-SLEDAI and rate of severe SFI flares at Week 52 in patients with high disease activity. Please note the subgroups of patient included in the indirect treatment comparison had to meet one of the two following criteria:- • Criteria I: Low complement levels AND anti-dsDNA positive. • Criteria II: Low complement levels or a SELENA-SLEDAI score ≥10. Since submitting to NICE, the clinical outcomes for the HDA-1 and HDA-2 subgroups, comparing belimumab with placebo relevant to this appraisal have been published for both belimumab IV and SC formulations. Whilst a naïve unadjusted comparison it does provide additional evidence of similar efficacy between the two formulations in subgroups with higher disease activity [11].
Key issue 8: Use of calibration factor is likely biasing results	**NO**	Please see response to Key issue 4.
Key issue 9: Implementation of 24-week response and treatment continuation in the model is inconsistent	**NO**	GSK disagrees that there is an issue with the implementation of the belimumab 24-week response and treatment continuation in the model. A 24-week responder rule has been implemented in the model only for patients on belimumab to reflect the current NICE Guidance and clinical practice. At 24 weeks, a patient on belimumab is defined as either a responder or a non-responder to treatment. This is determined by drawing from a Bernoulli distribution with a success probability that depends on a patient's baseline SS score. These probabilities by baseline SS score have been derived from BLISS clinical trial data (BLISS-52 and BLISS-76 for IV and BLISS SC for SC) by dividing the number of belimumab patients that had an SS score reduction of at least 4 points at 24 weeks compared to baseline with the number of patients with that baseline SS score, for each relevant subgroup i.e. HDA-1 population and HDA-2 population. The number of patients and 24-week responders on IV and SC belimumab in the HDA-2 population are provide in Tables 3 and 4 respectively. As such, the modelling of the 24-week responder status is directly linked to the change from baseline in SS score at week 24.

Consequently every simulated patient with an SS score reduction of at least 4 points at 24 weeks compared to baseline is correctly defined as a 'responder', and every simulated patient with an SS score reduction of less than 4 points at 24 weeks is correctly defined as a 'non-responder'. No mismatch between the 24 weeks SS score and responder status was identified; this is visible in the 'Patient log' sheets of the accompanying models.

Table 3. Responder probabilities by baseline SS score, IV administration of belimumab (HDA-2 population)

Baseline SS score	Number of patients in belimumab arm for HDA-2	Number of belimumab patients responding at week 24 (SS score reduction of ≥ 4 points vs. baseline)	Probability of response on belimumab at 24 weeks
10	113	69	61.1%
11	9	5	55.6%
12	64	46	71.9%
13	8	7	87.5%
14	21	16	81.0%
15	3	3	100.0%
16	20	14	70.0%
17	1	0	0.0%

			18	7	3	42.9%
			19	3	3	100.0%
			20	6	4	66.7%
			21	1	1	100.0%
			22	4	3	75.0%
			23	1	1	100.0%
			24	0	0	N/A
			25	0	0	N/A
			26	0	0	N/A
			27	0	0	N/A
			28	0	0	N/A
			29	0	0	N/A
			30	1	0	0%

Table 4. Responder probabilities by baseline SS score, SC administration of belimumab (HDA-2 population)

Baseline SS score	Number of patients in belimumab arm for HDA-2	Number of belimumab patients responding at week 24 (SS score reduction of ≥ 4 points vs. baseline)	Probability of response on belimumab at 24 weeks
10	110	68	61.8%
11	11	7	63.6%
12	84	61	72.6%
13	8	7	87.5%
14	35	30	85.7%
15	2	1	50.0%
16	21	16	76.2%

17	1	1	100.0%		
18	12	8	66.7%		
19	2	2	100.0%		
20	8	6	75.0%		
21	0	0	0.0%		
22	1	1	100.0%		
23	0	0	0.0%		
24	1	1	100.0%		
25	0	0	0.0%		
26	0	0	0.0%		
27	0	0	0.0%		
28	0	0	0.0%		
29	0	0	0.0%		
30	0	0	0.0%		

Calculation of SS score at 52 weeks
In the model, a patient's SS score is determined at 52 weeks using the 52-week regression equation.

For belimumab IV **responders** at week 24 in the HDA-2 population (example here) the calculation at 52 weeks is as follows:

$$SS\_52 = SS\_0 + (-0.327 * SS\_0 - 0.318 * SS_R\_0)$$

For patients that were belimumab IV **non-responders** at week 24 and are therefore receiving ST at week 52, the SS score at week 52 is calculated as:

$$SS\_52 = SS\_0 + (-0.379 * SS\_0)$$

(NOTE: SS_0:= SS score at baseline for all belimumab patients; SS_R_0 = SS score at baseline for belimumab responders; SS_52:= SS score at Week 52)

		This implies (as recognised by the ERG) that for certain baseline SS scores, belimumab non-responders at week 24 (who are therefore on ST-alone from week 24 onwards) could have an SS reduction of more than 4 points at 52 weeks compared to baseline. This is possible for patients with a baseline SS score of 11 points or higher and applies to 47% of non-responders on IV belimumab treatment (HDA-2 population). It is critical to note that this observation does not mean that these patients were incorrectly classified in the model as non-responders due to the chronology of the measurements; at 24 weeks the SS score reduction was below the response threshold with an average reduction of 0.4 points in SS score. Following the subsequent discontinuation of belimumab and 28 weeks on ST, the SS score reduction at Week 52 compared with baseline accumulated to 4 points or more. This is not a modelling error but merely a result of the response criterion applied at Week 24 which aligns with the NICE Guidance and implementation in clinical practice. This result for belimumab "non-responders" (categorised at Week 24) who subsequently met the response criterion at Week 52 is clinically plausible based on the feedback we received from the two clinical experts with whom we consulted with regarding Issue 10; they explained that non-responders stopping belimumab at week 24 would receive alternative treatments in order to control their disease activity as soon as possible.
Key issue 10: Error in calculation of belimumab non-responder disease activity at 52 weeks	**NO**	This is not an error in the model, rather it is an assumption that belimumab non-responders take the average ST score (regression coefficient) rather than the belimumab non-responder regression coefficient from week 52 onwards. This assumption was made as patients who do not respond on belimumab at week 24 switch to ST and continue ST for the remainder of the modelled time horizon i.e., the remaining 28 weeks of the first year and any remaining cycles thereafter. This assumption was validated with experts for TA397 and has been re-validated with two UK clinical experts for this re-appraisal. The feedback from the clinical experts is that when patients discontinue belimumab at week 24 due to not meeting the NHSE responder criterion (a ≥4-point reduction in SELENA-SLEDAI score from baseline) their ongoing management will depend on the severity of their disease and level of disease activity. Typically, if a patient has on-going active disease, their steroid dose could be increased alongside a change of immunosuppressant depending on prior treatments. If disease is severe, patients could receive IV cyclophosphamide or rituximab preceded by a dose of IV methylprednisolone to control symptoms in the short term. The aim of treatment is to stabilise their disease as soon as possible

		which could be achieved anytime from 3-months after stopping belimumab if they respond to the alternative treatment regimen. It is worth noting that belimumab "non-responders" comprise a mix of patients; some may not show any benefit at all from treatment with belimumab and their disease activity would have stayed the same or even worsened compared with baseline; but others could have started to benefit from belimumab and have a drop in disease activity from 1 to 3 points, but had to stop due to the stringent Week 24 stopping criterion. Therefore, given the heterogeneity in level of benefit it seems reasonable to move non-responding (by Week 24) belimumab patients over to an average ST SS score value and after six-months of an alternative treatment management i.e. at Week 52. We acknowledge the comments from the ERG's clinical expert which suggested that assuming the full ST efficacy at Week 52 may be a little too early, although on the whole they would expect belimumab non-responders beyond 52 weeks to have similar disease activity to the average patient on ST (see ERG report page 96). However, we also consulted with two clinical experts who advised that when a patient stops belimumab at 24 weeks due to a lack of adequate response, they would most likely change their choice of immunosuppressant, consider an increase in oral steroid dose or give IV methylprednisolone +/- IV cyclophosphamide or rituximab. Therefore, they would expect and aim for an improvement in disease activity within 3- 6 months of changing treatments. We have therefore provided a scenario analysis which tries to balance these two views to explore the impact on the ICER if, for the belimumab Week 24 non-responders, we incorporate a **less than full** ST efficacy at Week 52 and a later timepoint for assuming **full** ST efficacy. The methodology applied in the model for this scenario analysis is detailed below: • Belimumab non-responders experience non-responder belimumab efficacy (regression coefficient) for the first 24 weeks in Year 1. • During the period Week 24 to Week 52, the average ST efficacy applies, • A full return to ST efficacy for non-responders was assumed to occur after one full year of ST treatment, which is after 1.5 years (Week 76) since model entry, instead of after 28 weeks of ST treatment (i.e. at Week 52) as in our base case. • After Week 76, efficacy is modelled with the natural history disease activity model (using ST efficacy in the regression equation). In summary, the first-year efficacy is modelled as a weighted efficacy reflecting 24 weeks of belimumab non-responder efficacy and 28 weeks of full ST efficacy. The second-year efficacy is modelled as the

average efficacy of the first 6 months with full ST efficacy and the second 6 months from the natural history model. The results of this scenario analysis are provided in Tables 5 and 6 below.

Table 5. Company base case and scenario analysis results, modelling of belimumab non-responders, IV model (HDA-2)

	Base case			Scenario analysis		
	SoC	Belimumab	Difference	SoC	Belimumab	Difference
Life Years	16.90	[redacted]	[redacted]	16.90	[redacted]	[redacted]
QALYs	9.81	[redacted]	[redacted]	9.81	[redacted]	[redacted]
Costs	£160,470	[redacted]	[redacted]	£160,470	[redacted]	[redacted]
ICER			£30,001			£31,927

Table 6. Company base case and scenario analysis results, modelling of belimumab non-responders, SC model (HDA-2)

	Base case			Scenario analysis		
	SoC	Belimumab	Difference	SoC	Belimumab	Difference
Life Years	17.12	[redacted]	[redacted]	17.12	[redacted]	[redacted]
QALYs	10.06	[redacted]	[redacted]	10.06	[redacted]	[redacted]
Costs	£151,999	[redacted]	[redacted]	£151,999	[redacted]	[redacted]
ICER			£30,566			£33,616

The approach taken in the base case remains a reasonable assumption in clinical practice. Acknowledging the heterogeneity of the high disease activity population, the additional scenario analysis shows that the ICER increases by approximately 6%-10% for the IV and SC models respectively.

<table>
<tr>
<td>Key issue 11:
Violation in utility estimation</td>
<td>YES</td>
<td>GSK agrees that there is an error in the utility regression. As we are unable to re-analyse the data needed for this utility regression equation in time for the response to the technical report, we have instead considered some scenario analyses in order to demonstrate the degree of impact on the ICER. In these analyses we have varied the coefficients in the regression equation used in the submission by one standard deviation in each direction (see Tables 7 and 8 below). The choice of an increase or decrease of the coefficients by one SD was arbitrary and it could be considered a substantial change, however this will hopefully address sufficiently the uncertainty arising from incorporating incorrect coefficients in the final utility regression model.

Table 7. Impact on utility regression coefficients by varying by 1 SD (in %) in the IV model

<table>
<tr><th>Utility coefficient</th><th>Coefficient decreased by one SD</th><th>Coefficient increased by one SD</th></tr>
<tr><td>log of age</td><td>10.1%</td><td>-10.5%</td></tr>
<tr><td>constant</td><td>9.1%</td><td>-9.4%</td></tr>
<tr><td>SLEDAI score</td><td>-0.8%</td><td>0.3%</td></tr>
<tr><td>Black ethnicity</td><td>0.3%</td><td>-0.2%</td></tr>
</table>

Table 8. Impact on utility regression coefficients by varying by 1 SD (in %) in the SC model

<table>
<tr><th>Utility coefficient</th><th>Coefficient decreased by one standard deviation</th><th>Coefficient increased by one standard deviation</th></tr>
<tr><td>log of age</td><td>10.5%</td><td>-10.4%</td></tr>
<tr><td>constant</td><td>9.5%</td><td>-9.3%</td></tr>
<tr><td>SLEDAI score</td><td>-0.5%</td><td>0.5%</td></tr>
<tr><td>Black ethnicity</td><td>0.4%</td><td>-0.4%</td></tr>
</table>

The results in Tables 9 Table and 10 below indicate the change in ICER compared to our submission base case with the adjustment made to the utility coefficients for the IV and SC models, respectively. Figures 4 and 5 visualise these compared to our submission base case ICERs for the IV and SC models, respectively.</td>
</tr>
</table>

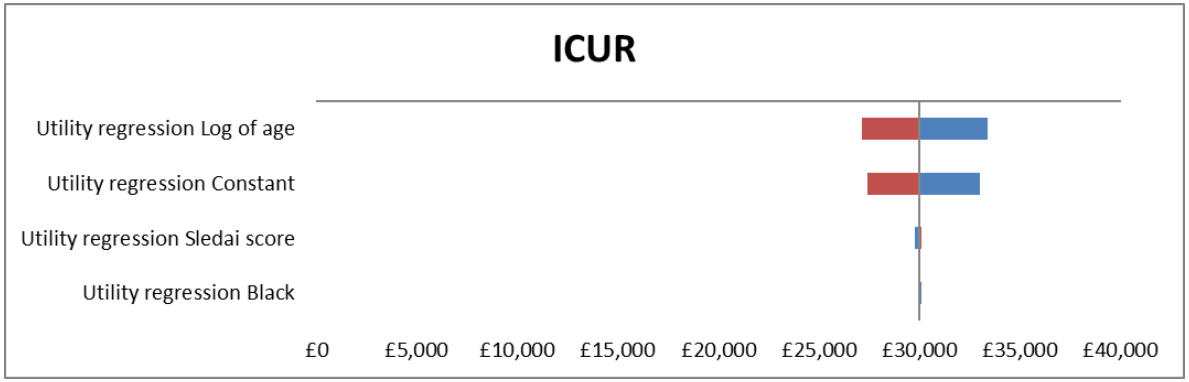
The results from these scenario analyses demonstrate that although there is an acknowledged error in the utility regression used in the models, this is likely to have a minimal impact on the ICER, providing reassurance in the validity of the cost-effectiveness results generated.

IV Model results:

Table 9. Utility coefficients used and ICERs obtained from sensitivity analysis compared with the base case analysis - IV model

Utility coefficient	Base case coefficient	Coefficient decreased by one SD	Coefficient increased by one SD	Base case ICER	ICER coefficient decreased	ICER coefficient increased
log of age	-0.1448	-0.1643	-0.1253	£30,001	£33,364	£27,142
constant	1.2970	1.2248	1.3674	£30,001	£33,017	£27,424
SLEDAI score	-0.0091	-0.0097	-0.0085	£30,001	£29,758	£30,101
Black ethnicity	-0.0538	-0.0752	-0.0336	£30,001	£30,076	£29,929

Figure 2. Tornado plot of ICERs in the IV model

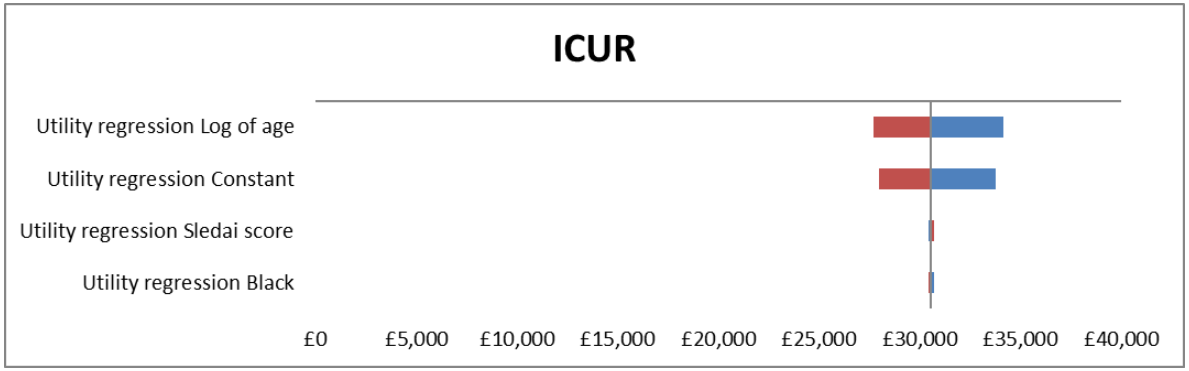


SC Model results:

Table 10. Utility coefficients used and ICERs obtained from sensitivity analysis compared with the base case analysis – SC Model

Utility coefficient	Base case coefficient	Coefficient decreased by one SD	Coefficient increased by one SD	Base case ICER	ICER coefficient decreased	ICER coefficient increased
log of age	-0.1448	-0.1643	-0.1253	£30,566	£34,137	£27,679
constant	1.2970	1.2248	1.3674	£30,566	£33,782	£27,967
SLEDAI score	-0.0091	-0.0097	-0.0085	£30,566	£30,415	£30,717
Black ethnicity	-0.0538	-0.0752	-0.0336	£30,566	£30,686	£30,454

Figure 3. Tornado plot ICERs of the SC model



Key issue 12: Uncertainty about organ damage utility multipliers	**YES**	The ERG highlighted uncertainty about organ damage utility multipliers and considered that they may over-estimate the impact of organ damage on HRQoL as the utility estimation function may capture this to a certain extent. The ERG suggested that GSK could investigate whether the weighting of organ damage

		items corresponds with the latest evidence and consult expert opinion on the magnitude of the organ damage utility multipliers. GSK has since re-consulted two clinical experts on the approach to deriving and applying the organ damage utility multipliers. Both clinical experts agreed that to their knowledge there are no definitive sources of utilities (weighting) by organ damage experienced by SLE patients and that GSK's approach to identifying such sources was considered reasonable. The overall findings from these consultations were as follows: • The incidence of avascular necrosis (AVN) presented in the submission appeared high compared to the other items within the musculoskeletal organ damage system, however the experts also commented that the impact of this item on a patient's quality of life when it occurs is significant. One clinical expert also noted that the frequency of a second episode of AVN is very uncommon and hence agreed with the low frequency presented within the submission. • The incidence of cardiovascular accident (CVA) would be expected to be higher than for cranial or peripheral neuropathy in the neuro-psychiatric organ damage system, however one clinician stated that when the latter occurred it could be equally debilitating. Other than these two areas, no other organ system items were highlighted by the clinicians as unreasonable in terms of the weighting applied to calculate the overall utility of that organ system or indeed the incidence assumed. Both clinicians also agreed that the hierarchy in the level of utility awarded to each organ system seemed reasonable. In line with the expert feedback, a scenario analysis was conducted where the following changes were made to the weightings provided in ***Table 67. Summary of quality-of-life values for the cost-effectiveness analysis from the previous submission, updated with values from the current literature search update*** in Document B of the company submission: • Neuropsychiatric organ damage system o The weighting of 'Cerebral vascular accident ever or resection (for causes other than malignancy)' was increased by 10% from 28% to 38%.

<table>
<tr>
<td></td>
<td></td>
<td>

The weighting of ‘Cranial or peripheral neuropathy’ was decreased by 10% from 31% to 21%.

Musculoskeletal organ damage system

The weighting of ‘Muscle atrophy / weakness’ was increased by 7% from 8% to 15%.
The weighting of ‘Avascular necrosis’ was decreased by 21% from 26% to 5%.
The weighting of ‘Ruptured tendon’ was increased by 7% from 8% to 15%.

Table 11 and Table 12 show the results of this scenario analysis in the IV and SC models for HDA-2 and show there is only a minor difference in each model for QALYS, resulting in a marginally improved ICER for belimumab.
Table 11. Company base case and scenario analysis results, modelling of updated utility multipliers, IV model (HDA-2)
<table>
<tr><th></th><th colspan="3">Base case</th><th colspan="3">Scenario analysis</th></tr>
<tr><th></th><th>SoC</th><th>Belimumab</th><th>Difference</th><th>SoC</th><th>Belimumab</th><th>Difference</th></tr>
<tr><td>Life Years</td><td>16.90</td><td></td><td></td><td>16.90</td><td></td><td></td></tr>
<tr><td>QALYs</td><td>9.81</td><td></td><td></td><td>9.65</td><td></td><td></td></tr>
<tr><td>Costs</td><td>£160,470</td><td></td><td></td><td>£160,470</td><td></td><td></td></tr>
<tr><td>ICER</td><td></td><td></td><td>£30,001</td><td></td><td></td><td>£29,704</td></tr>
</table>
Table 12. Company base case and scenario analysis results, modelling of updated utility multipliers, SC model (HDA-2)
<table>
<tr><th></th><th colspan="3">Base case</th><th colspan="3">Scenario analysis</th></tr>
<tr><th></th><th>SoC</th><th>Belimumab</th><th>Difference</th><th>SoC</th><th>Belimumab</th><th>Difference</th></tr>
<tr><td>Life Years</td><td>17.12</td><td></td><td></td><td>17.12</td><td></td><td></td></tr>
<tr><td>QALYs</td><td>10.06</td><td></td><td></td><td>10.11</td><td></td><td></td></tr>
<tr><td>Costs</td><td>£151,999</td><td></td><td></td><td>£151,999</td><td></td><td></td></tr>
<tr><td>ICER</td><td></td><td></td><td>£30,566</td><td></td><td></td><td>£30,442</td></tr>
</table>
</td>
</tr>
<tr>
<td>Key issue 13: Sampling of organ damage and death occurs after</td>
<td>NO</td>
<td>The issue described by the ERG relates to their difficulties in validating the model, specifically patient trajectories when all but treatment allocation is kept constant. The ERG comments that when comparing the outcomes of identical simulated patients undergoing ST and belimumab treatment, respectively, the clinical outcomes such as organ damage and time of death can vary substantially as a result of sampling of organ damage or death within each arm after sampling of treatment allocation. The ERG comments that</td>
</tr>
</table>

allocation to treatment		the impact of this is probably limited to the creation of noise which makes validation of the model performance more difficult. **Company position**: It is correct that treatment is sampled before clinical outcomes are modelled. Organ damage and mortality are modelled as a function of the SS score. The SS score in turn, is initially modelled as a function of baseline SS score, treatment allocation and treatment response. This dependency of clinical outcomes on treatment allocation entails that the structure of the model calculations cannot be changed to model organ damage outcomes before sampling of treatment allocation as it would ignore the impact of treatment on organ damage and mortality. Also, due to the stochastic nature of the model, outcomes such as organ damage of individual simulated patients are sampled from probability distributions whereas variation remains in the outcomes, no matter the order of calculations. For example, random intercept models were used for the modelling of change in SS scores. Consequently, even if treatment allocation could be modelled after organ damage outcomes were sampled, or even if we assumed belimumab and ST had equal treatment effectiveness for that matter, as the progression of disease is modelled including stochastic aspects will determine that the results even between two completely identical patients receiving identical treatment will be somewhat different. For this reason, a large number of patients is being simulated to ensure that sampling error is minimised, and model results are stable. This has been checked and confirmed by inspecting convergence plots, which indicate model results are stable and free from sampling error when a simulated cohort of 50,000 patients is analysed. In summary, we do not believe this is a key issue and the sampling order will not impact the validity of the model results.

References

1. Agency EM. Public Assessment Report (EPAR) - Benlysta, EMA/547408/2019 2019 [Available from: https://www.ema.europa.eu/en/documents/variation-report/benlysta-h-c-2015-ii-0062-epar-assessment-report-variation_en.pdf.
2. EMC. Benlysta 120 mg powder for concentrate for solution for infusion SmPC 2020 [Available from: https://www.medicines.org.uk/emc/product/4679/smpc.
3. EMC. Benlysta 200 mg solution for injection in pre-filled pen SmPC 2020 [Available from: https://www.medicines.org.uk/emc/product/11398/smpc

4. EMA. Guideline on clinical investigation of medicinal products for the treatment of systemic lupus erythematosus and lupus nephritis 2015 [Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-treatment-systemic-lupus-erythematosus-lupus_en.pdf.
5. Teng YO, Parikh SV, Saxena A, Solomons N, Huizinga RB. O11 AURORA phase 3 study demonstrates voclosporin statistical superiority over standard of care in lupus nephritis (LN). Archives of Disease in childhood; 2020.
6. Merrill JT, Neuwelt CM, Wallace DJ, Shanahan JC, Latinis KM, Oates JC, et al. Efficacy and safety of rituximab in moderately-to-severely active systemic lupus erythematosus: the randomized, double-blind, phase II/III systemic lupus erythematosus evaluation of rituximab trial. Arthritis Rheum. 2010;62(1):222-33.
7. Furie RA, Morand EF, Bruce IN, Manzi S, Kalunian KC, Vital EM, et al. Type I interferon inhibitor anifrolumab in active systemic lupus erythematosus (TULIP-1): a randomised, controlled, phase 3 trial. The Lancet Rheumatology. 2019;1(4):e208-e19.
8. Wallace DJ, Ginzler EM, Merrill JT, Furie RA, Stohl W, Chatham WW, et al. Safety and Efficacy of Belimumab Plus Standard Therapy for Up to Thirteen Years in Patients With Systemic Lupus Erythematosus. Arthritis & rheumatology (Hoboken, NJ). 2019;71(7):1125-34.
9. NHS England. Clinical Commissioning Policy Rituximab for refractory Systemic Lupus Erythematosus (SLE) in adults and post-pubescent children [200402P] 2020 [Available from: https://www.england.nhs.uk/wp-content/uploads/2020/07/Rituximab-for-refractory-Systemic-Lupus-Erythematosus-in-adults-and-post-pubescent-children-2.pdf.
10. Ramachandran S, Parks D, Kurtinecz M, Roth DA, Alfonso-Cristancho R. An indirect comparison of intravenous and subcutaneous belimumab efficacy in patients with SLE and high disease activity. Journal of Comparative Effectiveness Research. 2018;7(6):581-93.

11. Maslen T, Bruce IN, D'Cruz D, Ianosev M, Bass DL, Wilkinson C, et al. Efficacy of belimumab in two serologically distinct high disease activity subgroups of patients with systemic lupus erythematosus: post-hoc analysis of data from the phase III programme. Lupus science & medicine. 2021;8(1):e000459.

Updated cost-effectiveness results; revised PAS for belimumab IV: Belimumab + SoC vs SoC

GSK has proposed a new PAS to NHS England and Improvement (NHSE&I) which comprises a confidential common discount of [redacted] for the Benlysta intravenous (IV) and subcutaneous (SC) formulations. Although we have not yet had confirmation that this new PAS has been accepted by NHSE&I, as the Appraisal Committee meeting is only 5 weeks away, we are providing the new cost-effectiveness analyses for the IV formulation so that they can be considered by the committee if our new PAS is accepted. There are no changes to the cost-effectiveness analyses for the subcutaneous formulation provided in our submission template. Therefore, as agreed with the Project Manager for ID 1591 this document provides the updated cost effectiveness results inclusive of the revised PAS for belimumab IV.

In the Company submission (November 2020), the IV formulation PAS prices were [redacted] for the 120mg vial and [redacted] for the 400mg vial. To bring the IV formulation discount into line with that of the SC formulation relative to its list price, the updated PAS price for the IV formulations are now [redacted] for the 120mg vial and [redacted] for the 400mg vial.

The table below provides a summary of the updated IV model ICERs for HDA-1 and HDA-2 populations, with the previous ICERs provided for comparison.

Table 1. Summary of the updated IV model ICERs for HDA-1 and HDA-2 populations, with ICERs from the November 2020 company submission provided for comparison

	IV Model – HDA-2		IV Model – HDA-1	
Description of Scenario	As per November 2020 Company Submission	Updated PAS price	As per November 2020 Company Submission	Updated PAS price
Base case	£30,001	£29,162	£28,361	£27,522
1. Source of patient weight is BLISS trials	£28,095	£27,299	£25,983	£25,199
2. Belimumab treatment duration and effect restricted to 10 years	£20,485	£19,776	£18,538	£17,835
3. Calibration factors applied to both the belimumab and ST for 6 years	£23,419	£22,633	£21,245	£20,490
4. Calibration factors applied to belimumab only for patient lifetime	£24,187	£23,417	£22,301	£21,534
5. Discount rates 1.5% for both benefits and costs	£22,015	£21,384	£20,090	£19,464
6. Discount rates 1.5% for benefits and 3.5% for costs	£19,818	£19,264	£18,528	£17,980

Table 2. Base-case results for HDA-2 population –PAS price presented in Company Submission, November 2020 and the new PAS price

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
IV model – As per November 2020 Company Submission [redacted]							
ST	£160,470	16.90	9.81				
Belimumab IV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£30,001
IV model – Updated PAS price [redacted]							
ST	£160,470	16.90	9.81				
Belimumab IV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£29,162

All model outcomes presented are discounted.
Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years

Table 3. Results summary of the scenario analyses for HDA-2 – presented in Company Submission, November 2020 and the new PAS price

	IV model – As per November 2020 Company Submission [redacted]				**IV model –** Updated PAS prices [redacted]			
Description of Scenario	**Incremental Cost Belimumab**	**Incremental LYs Belimumab**	**Incremental QALYs Belimumab**	**ICER**	**Incremental Cost Belimumab**	**Incremental LYs Belimumab**	**Incremental QALYs Belimumab**	**ICER**
Base case	[redacted]	[redacted]	[redacted]	£30,001	[redacted]	[redacted]	[redacted]	£29,162
1. Source of patient weight is BLISS trials	[redacted]	[redacted]	[redacted]	£28,095	[redacted]	[redacted]	[redacted]	£27,299
2. Belimumab treatment duration and effect restricted to 10 years	[redacted]	[redacted]	[redacted]	£20,485	[redacted]	[redacted]	[redacted]	£19,776
3. Calibration factors applied to both the belimumab and ST for 6 years	[redacted]	[redacted]	[redacted]	£23,419	[redacted]	[redacted]	[redacted]	£22,633
4. Calibration factors applied to belimumab only for patient lifetime	[redacted]	[redacted]	[redacted]	£24,187	[redacted]	[redacted]	[redacted]	£23,417
5. Discount rates 1.5% for both benefits and costs	[redacted]	[redacted]	[redacted]	£22,015	[redacted]	[redacted]	[redacted]	£21,384
6. Discount rates 1.5% for benefits and 3.5% for costs	[redacted]	[redacted]	[redacted]	£19,818	[redacted]	[redacted]	[redacted]	£19,264

All model outcomes presented are discounted.
ICER, incremental cost-effectiveness ratio; LY, life years; QALYs, quality-adjusted life years

Table 4. Base-case results for HDA-1 population – presented in the Company Submission, November 2020 and with the new PAS price

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
IV model – As per November 2020 Company Submission [redacted]							
ST	£166,658	17.47	10.22				
Belimumab IV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£28,361
IV model – Updated PAS price [redacted]							
ST	£166,658	17.47	10.22				
Belimumab SC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£27,522

All model outcomes presented are discounted.
Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years

Table 5. Results summary of the scenario analyses for HDA-1 – presented in the Company Submission, November 2020 and the new PAS prices

	IV model – As per November 2020 Company Submission [redacted]				**IV model –** Updated PAS prices [redacted]			
Description of Scenario	**Incremental Cost Belimumab**	**Incremental LYs Belimumab**	**Incremental QALYs Belimumab**	**ICER**	**Incremental Cost Belimumab**	**Incremental LYs Belimumab**	**Incremental QALYs Belimumab**	**ICER**
Base case	[redacted]	[redacted]	[redacted]	£28,361	[redacted]	[redacted]	[redacted]	£27,522
1. Source of patient weight is BLISS trials	[redacted]	[redacted]	[redacted]	£25,983	[redacted]	[redacted]	[redacted]	£25,199
2. Belimumab treatment duration and effect restricted to 10 years	[redacted]	[redacted]	[redacted]	£18,538	[redacted]	[redacted]	[redacted]	£17,835
3. Calibration factors applied to both the belimumab and ST for 6 years	[redacted]	[redacted]	[redacted]	£21,245	[redacted]	[redacted]	[redacted]	£20,490
4. Calibration factors applied to belimumab only for patient lifetime	[redacted]	[redacted]	[redacted]	£22,301	[redacted]	[redacted]	[redacted]	£21,534
5. Discount rates 1.5% for both benefits and costs	[redacted]	[redacted]	[redacted]	£20,090	[redacted]	[redacted]	[redacted]	£19,464
6. Discount rates 1.5% for benefits and 3.5% for costs	[redacted]	[redacted]	[redacted]	£18,528	[redacted]	[redacted]	[redacted]	£17,980

All model outcomes presented are discounted.
ICER, incremental cost-effectiveness ratio; LY, life years; QALYs, quality-adjusted life years

Additional analyses provided in technical engagement response template

Below are updates for analyses provided as part of the technical engagement response, where the change in the PAS price for the IV model has affected ICERs. Please note that table numbers below correspond to those in the technical engagement response form.

Key issue 10: Error in calculation of belimumab non-responder disease activity at 52 weeks

Table 5. Company base case and scenario analysis results, modelling of belimumab non-responders, IV model (HDA-2)

	Base case			Scenario analysis		
	SoC	**Belimumab**	**Difference**	**SoC**	**Belimumab**	**Difference**
Life Years	16.90	[redacted]	[redacted]	16.90	[redacted]	[redacted]
QALYs	9.81	[redacted]	[redacted]	9.81	[redacted]	[redacted]
Costs	£160,470	[redacted]	[redacted]	£160,470	[redacted]	[redacted]
ICER			£29,162			£31,048

Key issue 11: Violation in utility estimation

Table 7. Impact on utility regression coefficients by varying by 1 SD (in %) in the IV model

Utility coefficient	Coefficient decreased by one SD	Coefficient increased by one SD
log of age	10.3%	-10.3%
constant	9.4%	-9.2%
SLEDAI score	-0.6%	0.6%
Black ethnicity	0.3%	-0.2%

IV Model results:

Table 9. Utility coefficients used and ICERs obtained from sensitivity analysis compared with the base case analysis - IV model

Utility coefficient	Base case coefficient	Coefficient decreased by one SD	Coefficient increased by one SD	Base case ICER	ICER coefficient decreased	ICER coefficient increased
log of age	-0.1448	-0.1643	-0.1253	£29,162	£32,518	£26,441
constant	1.2970	1.2248	1.3674	£29,162	£32,179	£26,716
SLEDAI score	-0.0091	-0.0097	-0.0085	£29,162	£28,994	£29,329
Black ethnicity	-0.0538	-0.0752	-0.0336	£29,162	£29,235	£29,093

Figure 1. Tornado plot of ICERs in the IV model

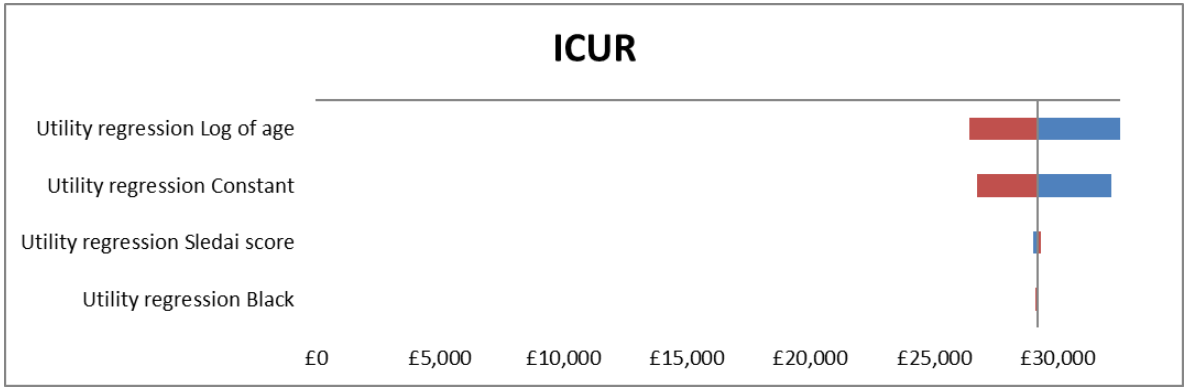


Table 11. Company base case and scenario analysis results, modelling of updated utility multipliers, IV model (HDA-2)

	Base case			Scenario analysis		
	SoC	Belimumab	Difference	SoC	Belimumab	Difference
Life Years	16.90	[redacted]	[redacted]	16.90	[redacted]	[redacted]
QALYs	9.81	[redacted]	[redacted]	9.65	[redacted]	[redacted]
Costs	£160,470	[redacted]	[redacted]	£160,470	[redacted]	[redacted]
ICER			£29,162			£29,172

Below are updates for the probabilistic sensitivity analyses provided as part of Document B of the submission, where the change in the PAS price for the IV model has affected ICERs.

Probabilistic results HDA-1 population

The probabilistic sensitivity analysis (1,000 iterations) resulted in an ICER of £28,058 per QALY for the IV formulation in the HDA-1 subgroup, with a probability of approximately 63.1% for the IV formulation to be cost effective for the commonly cited £30,000 willingness-to-pay-threshold per QALY (see Table 6). Figure 2 and Figure 3 present the incremental costs and QALYs, as well as the probability of cost-effectiveness at different willingness-to-pay thresholds.

Table 6. PSA results IV model, HDA-1 population

	Mean increment	Incremental cost-effectiveness ratio	Probability of cost-effectiveness
Costs	[redacted]		
QALYs	[redacted]		
ICER		£28,058	
Cost-effectiveness			63.1%

Figure 2. Scatter plot of incremental probabilistic outcomes, IV model, HDA-1 population

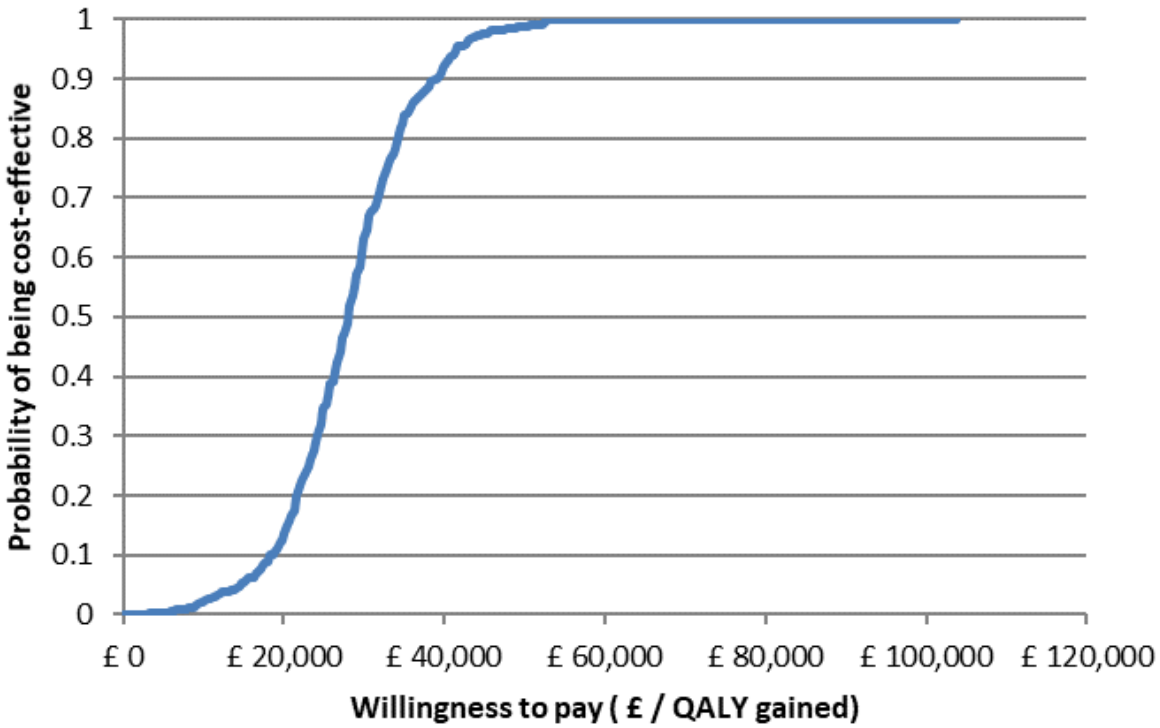


Figure 3. Cost-effectiveness acceptability curve IV model, HDA-1 population

Probabilistic results HDA-2 population

The probabilistic sensitivity analysis (1,000 iterations) resulted in an ICER of £30,808 per QALY for the IV formulation in the HDA-2 subgroup, with a probability of approximately 45% for the IV formulation to be cost effective for the commonly cited £30,000 willingness-to-pay-threshold per QALY (see Table 7). Figure 4 and Figure 5 present the incremental costs and QALYs, as well as the probability of cost-effectiveness at different willingness-to-pay thresholds.

Table 7. PSA results IV model, HDA-2 population

	Mean increment	Incremental cost-effectiveness ratio	Probability of cost-effectiveness
Costs	[redacted]		
QALYs	[redacted]		
ICER		£30,808	
Cost-effectiveness			45.3%

Figure 4. Scatter plot of incremental probabilistic outcomes, IV model, HDA-2 population

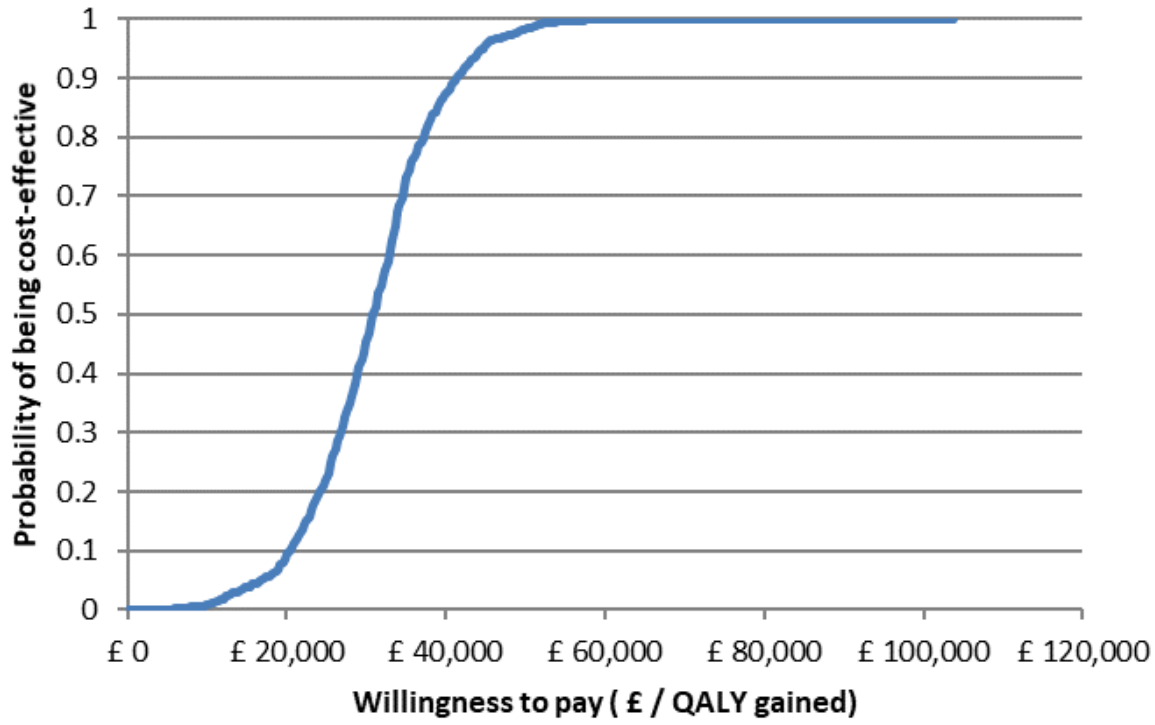


Figure 5. Cost-effectiveness acceptability curve IV model, HDA-2 population

NICE National Institute for Health and Care Excellence

Clinical expert statement & technical engagement response form

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (Review of TA397) [ID1591]

Thank you for agreeing to comment on the 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 you to complete questions where we ask for your views on this technology. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.
- In **part 2** we are asking you to give your views on key issues in the Evidence Review Group (ERG) report that are likely to be discussed by the committee. An overview of the key issues are summarised in the executive summary at the beginning of the ERG report.
- 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. In part 2 of this form we have included any of the issues raised by the ERG where we think having 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 return this form by **5:00pm** on **11 March 2021**

Completing this form

Part 1 can be completed anytime. We advise that the final draft of part 2 is completed after the expert engagement teleconference (if you are attending/have attended). This teleconference will briefly summarise the key issues, any specific questions we would like you to answer and the type of information the committee would find useful.

Important 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 want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Do not include medical information about yourself or another person that could identify you or the other person.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow.If confidential information is submitted, please also send a second 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.

PART 1 – Treating a patient with this condition and current treatment options	
About you	
1. Your name	**Professor Christopher Edwards**
2. Name of organisation	**University Hospital Southampton**
3. Job title or position	**Consultant Rheumatologist**
4. Are you (please tick all that apply):	☐ an employee or representative of a healthcare professional organisation that represents clinicians? **x** a specialist in the treatment of people with this condition? **x** a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	**x** yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.)

6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☐ yes
7. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**Nil funding from the tobacco industry**
The aim of treatment for this condition	
8. 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 slow, stop or reverse severe consequences of SLE including; Skin scarring, hair loss, joint pain and many other multisystem aspects of SLE As a result this will improve; Pain, mobility, reduce reliance on corticosteroids (and related steroid side-effects), mental health and disability
9. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm,	A reduction in disease activity of at least 2 in the SLEDAI score, reduced need for corticosteroids, reduction in BILAG A and B in all domains.

or a reduction in disease activity by a certain amount.)	
10. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes, SLE can be very difficult to treat and there are limited treatments for severe disease. This often results in unacceptably high doses of corticosteroids with side-effects or reliance on cyclophosphamide with potential permanent effects on fertility. The disease can be organ and life-threatening.
What is the expected place of the technology in current practice?	
11. How is the condition currently treated in the NHS?	A number of treatments – defined by BSR guidelines; Corticosteroids, immunosuppressives (methotrexate, azathioprine, mycophenolate mofetil), rituximab, cyclophosphamide.
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	BSR guidelines
• 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.)	Varies a lot – this is a complex and very heterogenous disease. It is very hard to do studies in this area an so it is impressive that belimumab has shown to be effective against a 'placebo arm' which includes a 'standard of care' with high dose steroids. General agreement between clinicians but often needs to be managed in specialist centres.

• What impact would the technology have on the current pathway of care?	Already available so there is experience that has been included in the BILAG Biologics Registry. It is very important for a group of SLE patients with severe disease.
12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Yes, but it would be good to expand access. The current requirements make it hard to use for many patients with need. The complexity of these requirements with additional need for specialist centre sign off anecdotally results in many patients not receiving belimumab therapy. As a result high doses of steroids are often used.
• How does healthcare resource use differ between the technology and current care?	Already in use.
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Secondary care – with access through a network organised by a specialist clinic.
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	None – all exists
13. Do you expect the technology to provide clinically meaningful	Greater access to a broader group of lupus patients would benefit a greater number of SLE patients. Hard to emphasise how hard it is to study and treat such a complex disease. It can be frustrating that it is easier to access high cost treatments for patients with inflammatory arthritis than for an organ and life-threatening disease like SLE.

benefits compared with current care?	
• Do you expect the technology to increase length of life more than current care?	Yes, if greater access the effects of chronic damage caused by uncontrolled systemic and tissue inflammation will be decreased. Also, less corticosteroid side-effects.
• Do you expect the technology to increase health-related quality of life more than current care?	Yes.
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Women are most effected by SLE. There is also more disease and more severe disease in black and minority ethnic populations.
The use of the technology	
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 (for example, any concomitant	Currently being used so no change.

treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	No additional testing but current TA defines both starting and stopping rules.
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?	Reduced corticosteroids side-effects
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	Already in use.

improve the way that current need is met?	
• Is the technology a 'step-change' in the management of the condition?	Already in use but was a step-change when introduced as only biological therapy currently licensed for SLE.
• Does the use of the technology address any particular unmet need of the patient population?	Uncontrolled SLE.
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?	Side-effects most often related to immunosuppression such as infections.
Sources of evidence	
20. Do the clinical trials on the technology reflect current UK clinical practice?	Mostly. I think there is less steroid use in the UK that would have increased the efficacy seen in the trials.
• If not, how could the results be extrapolated to the UK setting?	As above

• What, in your view, are the most important outcomes, and were they measured in the trials?	Yes, included but reduction in steroid use is very important.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	NA
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No
21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	Ongoing dataset held by the BILAG Biologics Registry
22. How do data on real-world experience compare with the trial data?	Encouraging – See BILAG Biologics Registry and other published international datasets.
Equality	

23a. Are there any potential equality issues that should be taken into account when considering this treatment?	SLE is most common and most severe in women and people from black and minority ethnic backgrounds.
23b. Consider whether these issues are different from issues with current care and why.	
Topic-specific questions	
24. Which high disease activity subgroup (HDA-1 or HDA-2) reflects the population who would benefit from treatment with belimumab in clinical practice?	Both

PART 2 – Technical engagement questions for clinical experts	
Issues arising from technical engagement	
We welcome your response to the questions below, but you do not have to answer every question. 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 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.	
Key issue 1: Evidence is missing for specific populations, such as children and patients with severe active central nervous system (CNS) lupus.	SLE is a very complex and heterogeneous disease and so that fact that efficacy is shown in adult populations without severe CNS disease is a major achievement. If you also take into account that lupus is present in around 1:1000 adults and far smaller numbers of children it is not surprising that a number of subsets have not been studied fully. Importance to emphasise how much more difficult it is to study SLE than for example rheumatoid arthritis.
Key issue 2: Some comparators listed in the NICE scope were not included.	

Key issue 3: Short follow-up in the main comparative trials (BLISS-SC, BLISS-52 and BLISS-76)	As above, these are very difficult area to study. Follow-up seemed reasonable clinically.
Key issue 4: Using the propensity score-matching (PSM) analysis in calibrating the cost-effectiveness model can severely bias the results in favour of belimumab	
Key issue 5: Data from the BILAG Biologic Register (BILAG BR) are not suitable for a comparison of belimumab with rituximab	Seems reasonable to use the BILAGBR. It is real-life UK data obtained and analysed by a very well respected group. Part of the reason to set up the Register was for this very reason.
Key issue 6: Rituximab + standard therapy was not	Rituximab is not licenced to treat lupus as it failed in Phase III studies. It is used as clinical experience suggests it is effective but it shows how complex this area is. How would you compare belimumab to data in phase III studies showing that rituximab is not effective.

included as comparator in the model	
Key issue 7: IV and SC formulations are not compared with each other, as two separate model files are provided.	No sure why we would expect them to be different. There are many examples of comparable therapies for other inflammatory rheumatic disease (infliximab, abatacept, tocilizumab) that are available in both sc and iv formulation. I am not aware of examples that show the efficacy or safety are different.
Key issue 8: Use of calibration factor is likely biasing results	
Key issue 9: Implementation of 24-week response and treatment continuation in the model is inconsistent	
Key issue 10: Error in calculation of belimumab non-responder disease activity at 52 weeks	

Key issue 11: Violation in utility estimation	
Key issue 12: Uncertainty about organ damage utility multipliers	
Key issue 13: Sampling of organ damage and death occurs after allocation to treatment	
Are there any important issues that have been missed in ERG report?	

PART 3 -Key messages

16. In up to 5 sentences, please summarise the key messages of your statement:

- SLE is a complex disease that causes tissue and organ damage along with reducing life-expectancy
- Belimumab provides an important additional treatment for those with severe disease

- A major benefit is reduction in corticosteroid use
- SLE is rare and complex to study so some populations and sub-groups will not be included in all trials
- Current TA access criteria are complicated and set in a way that excludes patients who might benefit

Thank you for your time.

Please log in to your NICE Docs account to upload your completed document, declaration of interest form and consent form.

.....

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

NICE National Institute for Health and Care Excellence

Clinical expert statement & technical engagement response form

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (Review of TA397) [ID1591]

Thank you for agreeing to comment on the 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 you to complete questions where we ask for your views on this technology. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.
- In **part 2** we are asking you to give your views on key issues in the Evidence Review Group (ERG) report that are likely to be discussed by the committee. An overview of the key issues are summarised in the executive summary at the beginning of the ERG report.
- 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. In part 2 of this form we have included any of the issues raised by the ERG where we think having 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 return this form by **5:00pm** on **11 March 2021**

Completing this form

Part 1 can be completed anytime. We advise that the final draft of part 2 is completed after the expert engagement teleconference (if you are attending/have attended). This teleconference will briefly summarise the key issues, any specific questions we would like you to answer and the type of information the committee would find useful.

Important 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 want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Do not include medical information about yourself or another person that could identify you or the other person.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow.If confidential information is submitted, please also send a second 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.

PART 1 – Treating a patient with this condition and current treatment options	
About you	
1. Your name	**Peter Lanyon**
2. Name of organisation	**NHS England Specialised Rheumatology Clinical Reference Group**
3. Job title or position	**Consultant Rheumatologist** **National Clinical Co-Lead for Rheumatology, Getting It Right First Time (GIRFT)**
4. Are you (please tick all that apply):	☐ an employee or representative of a healthcare professional organisation that represents clinicians? **X**☐ a specialist in the treatment of people with this condition? **X**☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☐ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it **X**☐ other (they didn't submit one, I don't know if they submitted one etc.)

6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☐ yes
7. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**
The aim of treatment for this condition	
8. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	The aim of drug treatment is to improve the health of people with Systemic Lupus Erythematosus (SLE). All treatments are intended to reduction the inflammatory disease activity, which in turn will 1. Reduce symptoms of currently active disease 2. Reduce the risk of new items of disease activity (flares) occurring 3. Reduce the risk of organ damage/scarring related to active disease 4. Improve quality of life
9. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity	Assessing response to treatments in SLE requires use of a global clinical scoring system - symptoms of active disease can occur in many different organ systems, and the severity of disease and the systems involved varies between patients. Two clinical scoring systems – BILAG and SLEDAI are used in both clinical trials and more recently have been embedded into clinical practice. Their introduction into routine clinical practice in the NHS has been catalysed by both BSR guidelines on the management of SLE and the NHS England Commissioning Policy for rituximab for refractory

by a certain amount.)	SLE. This Policy indicates a clinically meaningful response (for non renal disease) by 6 months as 1. A reduction in SLEDAI-2K score by ≥ 4 points from baseline 2. Loss of all A and B BILAG scores to < 1 B score with no new A or B scores in other organ domains at 6 months I think these are appropriate measures for defining clinical response
10. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes, definitely. And the evidence for this is now stronger than at the time of the initial TA. One of the best pieces of recent evidence for this is from an audit of the current NHS care of patients who are attending secondary care rheumatology clinics. A multi-centre audit of the 2018 BSR guideline on the management of adults with systemic lupus erythematosus (SLE) has audited 1021 episodes of care in 51 UK rheumatology units. The main findings from this audit are that • Almost a third of patients had on-going disease activity, notably independent of disease duration • Nearly half of those audited, and more than a quarter of those who had inactive SLE, were being prescribed maintenance prednisolone. • patients seen in a specialised SLE/CTD/vasculitis clinic were more likely to receive appropriate urine quantification and blood pressure measurement https://academic.oup.com/rheumatology/article-abstract/60/3/1480/6027899 In addition, in terms of unmet need, there is likely to be geographical variation in the uptake of Belimumab into routine NHS clinical practice, related to the policy implementation requirements.
What is the expected place of the technology in current practice?	
11. How is the condition currently treated in the NHS?	People who live this medical condition often need to have lifelong secondary healthcare, which is usually provided by either rheumatology or renal medicine specialists. Joint care coordinated across specialties e.g. renal, rheumatology and dermatology is increasingly common.

	In addition, for patients with severe or refractory disease, there is increasing involvement of specialised rheumatology regional networks in the care of patients, to support access to specialised expertise regardless of geography. Networked care has also been catalysed by the NHS England commissioning criteria for the use of rituximab in SLE. This requires treatment, at or involvement of a Specialised centre which is on the NHS England Provider Eligibility List. The medical drug treatments are tailored/escalated according to disease severity and include corticosteroids, antimalarials (hydroxychloroquine), immunosuppression (methotrexate, azathioprine, mycophenolate, cyclophosphamide) and biologic drugs (belimumab, rituximab). The place of these agents stratified according to disease activity is summarised very well in the BSR guideline.
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	Yes, guidelines produced by BSR in 2018 are the most common framework for care. Gordon C , Amissah-Arthur MB , Gayed M , Brown S, et al. ; British Society for Rheumatology Standards, Audit and Guidelines Working Group. The British Society for Rheumatology guideline for the management of systemic lupus erythematosus in adults. Rheumatology 2018;57:e1–45. Although not a clinical guideline, the NHS England commissioning policies are also used as a framework for treatment
• 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.)	Yes, the pathways of care are well defined in guidelines, but we need to be mindful that SLE is a rare disease. It is likely that variation in care will generally be greater in rare disease, one of the factors being the degree of clinical experience of the treating physician (hence the importance of networked specialised care). However fundamentally I do not think there is difference in opinion between medical professionals about treatment.
• What impact would the technology have on the current pathway of care?	It would create additional therapeutic “headroom” in a much needed space. Changing the eligibility for belimumab to include people with either anti-dsDNA antibodies or low C3/C4 complement makes very good clinical and biologic sense here and would improve the ability of the technology to improve patient care. Achievement of the general ambition of NHS England’s specialised commissioning function of ensuring access to specialised care and treatment regardless of geography would be enhanced if the requirement for treatment to be

	delivered at a specialised centre was removed. The pathway of care would also be improved further by access to SC as opposed to IV treatment. This is particularly important in terms of reducing COVID / infection related risks and supporting business continuity of NHS day case infusion units for patients for whom there is no alternative to an IV drug
12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Yes
• How does healthcare resource use differ between the technology and current care?	The main difference in resource use is that this is the only drug that currently requires long term intravenous administration, with the associated inconvenience for patients and associated administrative and financial costs of the day case procedure. Use of the SC formulation would significantly reduce resource use in day case units in terms of space, nursing time, administrative booking time etc. The real cost benefits of this are difficult to fully quantify because the availability of space and nursing time has been at a complete premium during this recent and any future pandemics. The ability to reduce even a small amount of this demand is very important
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Secondary care. It would seem appropriate for the current NHS England arrangements for commissioning high cost biologics for rare autoimmune diseases to continue e.g. treatment at or in conjunction with a specialised centre. The requirement for patients to travel to a specialised centre for treatment is likely to be a significant factor in any variation in current use, and I would recommend removing this requirement.
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	Very little for facilities, equipment or training. The small number of patients is not likely to place undue demands on existing rheumatology day case units, and particularly with SC use

13. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes. The trial evidence indicates that there are clinically meaningful benefits compared to Standard Care for the proposed cohorts of patients.
• Do you expect the technology to increase length of life more than current care?	Yes I would expect this to be the case. However, as far as I am aware there isn't yet any evidence of increased life expectancy as a direct result of the technology. But we know that the increased risk of mortality in SLE is likely to be due to a combination of both active disease, organ damage, and infection risk, for which steroids are an important risk factor. Hence strategies to control disease activity and reduce steroid burden (all prominent features in the BSR SLE audit) will have a beneficial impact on life expectancy as well as quality of life.
• Do you expect the technology to increase health-related quality of life more than current care?	Yes. These outcomes are included in the belimumab clinical trials and appear to show meaningful improvement.
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Yes. The trial data suggests the benefit is greater in people who have a greater disease activity. The suggestions that this technology is used in patients with a SLEDAI score of 10 or more and who have either positive dsDNA antibodies or low C3/C4 are therefore appropriate.
The use of the technology	
15. Will the technology be easier or more difficult to use for patients or healthcare professionals than	The only potential practical change is that SC formulation would potentially be much easier for patients than IV. There aren't any specific additional monitoring arrangements over and above that which would be needed for standard care for patients with similar disease activity.

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.)	
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Yes I would suggest the formal rules for starting treatment as proposed based on disease activity. There would need to be consensus broad of which drugs (and for how long) would need to be used prior to belimumab. Stopping criteria should be based on lack of response (or toxicity) No additional testing above and beyond routine clinical care is needed to apply these rules
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?	I don't have expertise in QALY calculation but anticipate that identifying the full benefits of reducing steroid related toxicity is difficult
18. Do you consider the	Yes there is no doubt that this is innovative. Much has been said already in this and previous appraisals about this being the first new drug for lupus in >50 years. But equally relevant is the fact that this drug (at the moment) is only

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?	used in lupus e.g. this is disease-personalised treatment. The importance of this fact for patients should not be underestimated. For the last 50 years drugs for patients with lupus have been borrowed from other conditions. For example we tell patients that drug X works for these other conditions and so we anticipate it will work in SLE. Being able to tell patients that this drug is specific for their condition is very important. In terms of innovation, its also relevant to say that there are no other new licensed drugs for SLE since this original first appraisal started. In terms of how this will meet "current need" – it has already been demonstrated that there is a major need for better disease control and steroid reduction – amongst patients with SLE who are currently receiving NHS care. This technology would appear to be a significant step in meeting these needs. Particularly as the longer term studies do not appear to show any new concerns about safety or continued efficacy.
• Is the technology a 'step-change' in the management of the condition?	Yes. For the reasons outlined above. And the SC formulation would also be the first SC drug available to patients with SLE.
• Does the use of the technology address any particular unmet need of the patient population?	Yes as per BSR audit there is a need for better disease control and reduced reliance on steroids
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?	In general the side effect profile does not seem different or concerning compared to standard of care.

Sources of evidence	
20. Do the clinical trials on the technology reflect current UK clinical practice?	Yes, as the clinical trials required patients to have active disease and also included standard treatment.
• If not, how could the results be extrapolated to the UK setting?	
• What, in your view, are the most important outcomes, and were they measured in the trials?	I think the trials were comprehensive in that they collected the main outcomes that can be formally measured in a trial including assessment of disease activity response, flares, change in disease related damage, steroid use and QOL including fatigue.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	Yes I think this is likely to be the case, as we know that active disease and predicts medium and long term damage which has impact on long term outcomes
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	Not that I am aware of
21. Are you aware of any relevant	I am not aware of any evidence other than any studies published since the review was undertaken and any

evidence that might not be found by a systematic review of the trial evidence?	unpublished BILAG BR data
22. How do data on real-world experience compare with the trial data?	In general my interpretation of these PMS extension studies is that they confirm ongoing tolerability and efficacy
Equality	
23a. Are there any potential equality issues that should be taken into account when considering this treatment?	Lupus is more likely to affect people from non-caucasian backgrounds and women of childbearing age. It is also important to consider reducing any inequity in access to treatment as a result of geography.
23b. Consider whether these issues are different from issues with current care and why.	This is related to the current requirement for treatment to be delivered at specialised centre
Topic-specific questions	
24. Which high disease activity subgroup (HDA-1 or HDA-2) reflects the population who would benefit from treatment with	HAD-2

belimumab in clinical practice?	

PART 2 – Technical engagement questions for clinical experts	
Issues arising from technical engagement	
We welcome your response to the questions below, but you do not have to answer every question. 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 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.	
Key issue 1: Evidence is missing for specific populations, such as children and patients with severe active central nervous system (CNS) lupus.	My understanding is that the PLUTO trial was not adequately powered. CNS lupus is outside of scope as excluded from BLISS trials
Key issue 2: Some comparators listed in the NICE scope were not included.	I agree with the ERG report. Cyclophosphamde is used for a different population including renal and CBS disease I agree with the ERG report that a head to head comparison with rituximab is very difficult - it has not met primary end points in key trials and also in real life comparisons - the eligibility criteria are different

Key issue 3: Short follow-up in the main comparative trials (BLISS-SC, BLISS-52 and BLISS-76)	In practical clinical terms – if a drug has met its end point compared to comparators over the duration of a clinical trial, unless there is a plausible proposition that there is a biologic reason for there to be a greater drop off in biological response in the treatment arm compared to comparator, I would usually expect the trial duration to predict the longer term outcome e.g. the lack of trial data beyond 76 weeks would not be a reason to doubt long term efficacy
Key issue 4: Using the propensity score-matching (PSM) analysis in calibrating the cost-effectiveness model can severely bias the results in favour of belimumab	
Key issue 5: Data from the BILAG Biologic Register (BILAG BR) are not suitable for a comparison of belimumab with rituximab	I agree that the eligibility criteria are different
Key issue 6: Rituximab + standard therapy was not included as comparator in the	Agree that the eligibility criteria are different

model	
Key issue 7: IV and SC formulations are not compared with each other, as two separate model files are provided.	
Key issue 8: Use of calibration factor is likely biasing results	
Key issue 9: Implementation of 24-week response and treatment continuation in the model is inconsistent	
Key issue 10: Error in calculation of belimumab non-responder disease activity at 52 weeks	
Key issue 11: Violation in	

utility estimation	
Key issue 12: Uncertainty about organ damage utility multipliers	
Key issue 13: Sampling of organ damage and death occurs after allocation to treatment	
Are there any important issues that have been missed in ERG report?	No I think this is comprehensive
PART 3 -Key messages	
16. In up to 5 sentences, please summarise the key messages of your statement: • There is significant unmet need for new treatments for SLE, as evidenced by recent NHS audit • Belimumab has a significant place here as an innovative treatment option to improve disease control, quality of life and improve long term outcomes. • I think that the proposed use in HDA2 patients is appropriate.	

- There are opportunities to look at how uptake of this innovation into the NHS, and reducing variation, could be improved.
- Reducing the requirement for IV treatment to be given at a specialised centre and enabling SC use would potentially address this

Thank you for your time.

Please log in to your NICE Docs account to upload your completed document, declaration of interest form and consent form.

.....

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

NICE National Institute for Health and Care Excellence

Patient expert statement and technical engagement response form

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (Review of TA397) [ID1591]

Thank you for agreeing to give us your views on this treatment 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.

About this Form

In **part 1** we are asking you to complete questions about living with or caring for a patient with the condition.

In **part 2** we are asking you to give your views on key issues in the Evidence Review Group (ERG) report that are likely to be discussed by the committee. An overview of the key issues are summarised in the executive summary at the beginning of the ERG report.

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. In part 2 of this form we have included any of the issues raised by the ERG where we think having a patient 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.

If you have any questions or need help with completing this form please email the public involvement team via pip@nice.org.uk (please include the ID number of your appraisal in any correspondence to the PIP team).

Please return this form by **5:00pm on 11 March 2021**

Completing this form

Part 1 can be completed anytime. We advise that the final draft of part 2 is completed after the expert engagement teleconference (if you are attending/have attended). This teleconference will briefly summarise the key issues, any specific questions we would like you to answer and the type of information the committee would find useful.

Please use this questionnaire with our hints and tips for patient experts. You can also refer to the Patient Organisation submission guide. **You do not have to answer every question** – they are prompts to guide you. There is also an opportunity to raise issues that are important to patients that you think have been missed and want to bring to the attention of the committee. The text boxes will expand as you type.

Important 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 want 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 15 pages.

PART 1 – Living with or caring for a patient with this condition and current treatment options	
About you	
1.Your name	**Jane Robinson**
2. Are you (please tick all that apply):	☐ **a patient with this condition?** ☐ a patient with experience of the treatment being evaluated? ☐ a carer of a patient with this condition? ☐ a patient organisation employee or volunteer? ☐ other (please specify):
3. Name of your nominating organisation.	Specialist Rheumatology CRG, NHS England
4. Has your nominating organisation provided a submission? Please tick all options that apply.	☐ **No, (please review all the questions below and provide answers where possible)** ☐ Yes, my nominating organisation has provided a submission ☐ I agree with it and **do not wish to** complete a patient expert statement ☐ Yes, I authored / was a contributor to my nominating organisations submission ☐ I agree with it and **do not wish to** complete this statement ☐ I agree with it and **will be** completing

5. How did you gather the information included in your statement? (please tick all that apply)	☐ **I am drawing from personal experience.** ☐ I have other relevant knowledge/experience (e.g. I am drawing on others' experiences). Please specify what other experience: ☐ I have completed part 2 of the statement **after attending** the expert engagement teleconference ☐ I have completed part 2 of the statement **but was not able to attend** the expert engagement teleconference ☐ I have not completed part 2 of the statement
Living with the condition	
6. What is your experience of living with this condition? If you are a carer (for someone with this condition) please share your experience of caring for them.	Daily issues with fatigue, mental acuity, joint pain and headaches. Frequent migraines. Daily issues with managing fatigue and cognitive issues especially and being able to do all the things you want, constantly having to plan how you use your limited energy, dealing with the consequences if you over do it. Any time you are ill the immediate reaction is lupus but then it might not be and a constant awareness that it might in fact be something else. Having to manage a long term health problem and medical appointments and drugs etc. is always hard and especially when you don't feel well. Being on drugs that can affect your immunity is concerning and has been so especially in the last year, along with shielding and the impact that has had on mental health as well as my physical health. Many people with lupus have to significantly change their life plans from career to lifestyle choices. I wasn't able to work at the level I wanted to and was under constant pressure with ill health reporting through the corporate system was particularly stressful and being referred to occupational health because of a "poor" sickness record. I chose to change how I worked and chose to work for myself to manage workload better and get out of the problems of having days where I was unable to work and this did

	allow me to have a life outside of work asl well but not an option for a lot of people. However, this was also challenging as often involved periods of working more which was ideal so reduced my available hours. Eventually I took a local a job supported by some free lance hours but consequently I earn less and am playing a lot less tax. I chose early on after being diagnosed with lupus and after a particularly challenging flare that there were times I wouldn't be able to look after myself so how could I look after a child. Whilst I don't regret that decision it was a particularly difficult one to make and at the same time seriously think that I may not have the career that I had dreamed of as well. Psychosocial issues are very challenging for people with lupus and there is very little support available. The stress of living with an illness is challenging to start with let alone when there are issues relating to not being able to work, how the disease makes you feel, the treatment makes you look, lack of energy to be social. Whilst no drug will help with these issues additional drug options will mean less people have these issues. People are treated differently for lupus depending if they are treated in a specialist centre with access to drug trials and those that are treated at general hospitals. It can be noticeable the different in treatments, especially those getting infusions when comparing treatment offered at different centres.
Current treatment of the condition in the NHS	
7a. What do you think of the current treatments and care available for this condition on the NHS?	7a) Treatment options seem to be limited. Flares are often treated with steroids and some people struggle to have a reasonable quality of life without a maintenance dose of steroids. It seems it can be a bit of an art finding the right combination of treatment to help and so the more options there are available the better. There is no one drug or set of drugs to use for people with lupus and often finding the right regime can take a considerable number of years. Treatment options are also variable depending what centres people attend for their lupus care are prescribed at those centres and this is difficult for patients to understand.

7b. How do your views on these current treatments compare to those of other people that you may be aware of?	Going forward it isn't possible for everyone to be treated at a specialist centre due to problems with access, time required to attend a clinic that isn't in your local town, issues getting referred, etc. 7b) I know many people who struggle to have any quality of life with the current treatment available and many have to rely on a partner or the State for income due to their lupus and the fact they haven't found the right combination of drugs to stay well long enough to have a good work/ life balance. More options will hopefully give more people with lupus the opportunity to find the right treatment for them and able to maintain their potential. People can spend years trying to find the right combination of treatment and in that time have a very poor quality of life, loss of confidence and function so that they are less able if the right treatment is found.
8. If there are disadvantages for patients of **current NHS treatments** for this condition (for example how the treatment is given or taken, side effects of treatment etc) please describe these	Many people need to supplement current treatment with corticosteroids for periods and some struggle to stop taking them completely. Long term use of steroids are a big concern and many people are dealing with risks of osteoporosis, diabetes, etc. Weight is a big concern for many as well especially as many find moving difficult so struggle to have a good exercise regime. People with lupus seem to have a lot of allergies and this can include some medications used for lupus and soon they find they've tried current options and there is nothing left. Often if someone has an allergy to one drug they will have an allergy to many. When options have been worked through patients are often just left to manage their symptoms struggling with day to day problems significantly impacting their work and home life.
Advantages of this treatment	
9a. If there are advantages of this treatment over current treatments on the NHS please describe these. For example, the impact on your Quality of Life your	9a) It's another option for people to have that they may be able to try to help their lupus symptoms improve. Any improvement in symptoms is an Improvement in your quality of life and increase people's chances to maintain and improve their lives, being able to be an effective family member, working and paying taxes,

ability to continue work, education, self-care, and care for others? 9b. If you have stated more than one advantage, which one(s) do you consider to be the most important, and why? 9c. Does this treatment help to overcome/address any of the listed disadvantages of current treatment that you have described in question 8? If so, please describe these.	caring, learning, having a social life. 9b) personally being an effective family member includes working, caring, learning and socialising 9c) I understand that this drug would be steroid sparing. I am not aware of issues relating to allergies and belimumab. I would hope that if this treatment is approved it would be easier for non specialist hospitals to use this treatment with their patients as well
Disadvantages of this treatment	
10. If there are disadvantages of this treatment over current treatments on the NHS please describe these? For example, are there any risks with this treatment? If you are concerned about any potential side affects you have heard about, please describe them and explain why.	Having to travel to attend specialist centres is a disadvantage. Some people will not want to receive an infusion but there is a sub cutaneous option where patients can do themselves. Equally some people may find this challenging. However, anyone I know who has infusions for their lupus are happy that they are getting help with their symptoms and generally feeling better that they are happy to have either infusion of sub-cut options for treatment if it makes them feel better.

Patient population	
11. Are there any groups of patients who might benefit more from this treatment or any who may benefit less? If so, please describe them and explain why. Consider, for example, if patients also have other health conditions (for example difficulties with mobility, dexterity or cognitive impairments) that affect the suitability of different treatments	If a patient meets the criteria for belimumab and it will fit in with their lifestyle I think everyone should be offered this treatment. I am unaware of issues that might make some less suitable for this treatment
Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged. Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race,	More women than men have lupus. Most people are diagnosed in their child-bearing years, more Afro-Caribbean people have lupus. Given people tend to live in communities of similar ethnicities and backgrounds there may be an impact of some people may not have access to some treatments if not in an area where there is no specialist centre or availability

religion or belief, sex, and sexual orientation or people with any other shared characteristics More information on how NICE deals with equalities issues can be found in the NICE equality scheme More general information about the Equality Act can and equalities issues can be found at https://www.gov.uk/government/publications/easy-read-the-equality-act-making-equality-real and https://www.gov.uk/discrimination-your-rights.	
Other issues	
13. Are there any other issues that you would like the committee to consider?	Whilst lupus does see people with periods of more significant disease (flares) it is a condition that has an impact on your daily life, it just gets more challenging when you are flaring. People regularly have to take major decisions about their choices in life which can have an ongoing impact, live with one or many symptom on a daily basis.

PART 2 – Technical engagement questions for patient experts	
Issues arising from technical engagement	
We welcome your response to the questions below, but you do not have to answer every question. If you think an issue that is important to 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 patient organisation that nominated you has 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.	
14a. What are the main benefits of this treatment for patients? If there are several benefits please list them in order of importance. Are there any benefits of this treatment that have not been captured?	Improvement in symptoms Use in children with SLE Available as a subcutaneous treatment as well as by infusion giving patients and clinicians options how to treat someone
14b. What are the benefits of this treatment for carers?	A person in your life who feels better will be a huge benefit but they will probably have less caring duties and a better quality of life themselves

15. Are there any important issues that have been missed in ERG report?	

PART 3 -Key messages

16. In up to 5 sentences, please summarise the key messages of your statement:

- Lupus does have periods of milder symptoms and acute flares BUT there isn't a day where I am not aware I have lupus and this is a common feature.
- Any drug that can reduce the impact of a long term illness is a benefit to the individual but also their family, their work colleagues and society as a whole reducing potential need for benefits and the ability to carry on paying taxes.
- Any drug that can reduce the use of corticosteroids in a patient will help the longer term health of the patient and reduce the risk of complications exacerbated by steroid use.
- There are two options for use of belimumab giving patients and clinicians choices for treatment to help improve health and social outcomes for patients.
- A child with lupus can have belimumab and this can seriously improve their outcomes and hopefully reduce their changes of ill health being a life long burden in their lives and the lives of their families.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

.....

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

.....

NICE National Institute for Health and Care Excellence

Technical engagement response form

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (Review of TA397) [ID1591]

As a stakeholder you have been invited to comment on the ERG report for this appraisal. 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.

We need your comments and feedback on the key issues below. You do not have to provide a response to every issue. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be included in the committee papers in full and may also be summarised and presented in slides at the appraisal committee meeting.

Deadline for comments **5:00pm on 11 March 2021**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the ERG report which summarises the background and submitted evidence, and presents the ERG's summary of key issues, critique of the evidence and exploratory analyses. This will provide context and describe the questions below in greater detail.
- Please ensure your response clearly identifies the issue numbers that have been used in the executive summary of the ERG report. 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.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, 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.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	[redacted]
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**LUPUS UK**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**

Key issues for engagement

Please use the table below to respond to questions raised in the ERG report on key issues. You may also provide additional comments on the key issue that you would like to raise but which do not address the specific questions.

Key issue	Does this response contain new evidence, data or analyses?	Response
Key issue 1: Evidence is missing for specific populations, such as children and patients with severe active central nervous system (CNS) lupus.	**YES**	**<u>Childhood Lupus (5-18)</u>** "Juvenile-onset systemic lupus erythematosus (jSLE) accounts for up to 20% of all SLE patients." "Key differences between juvenile- and adult-onset (aSLE) disease include higher disease activity, earlier development of damage, and increased use of immunosuppressive treatment in jSLE." "Lupus patients have higher mortality rates when compared to the general population and they are highest in young people. A study of 924 patients (413 with jSLE) demonstrated the standardized mortality ratio to be 18.3 in juvenile-onset disease compared to 3.1 in those with adult-onset disease." "Glucocorticoids remain a crucial part of jSLE management in the form of topical and/or low dose oral treatment for mild-to-moderate disease, or high dose oral or intravenous treatment for severe manifestations." *Juvenile-onset systemic lupus erythematosus: Update on clinical presentation, pathophysiology and treatment options.* *https://doi.org/10.1016/j.clim.2019.108274* Due to the earlier onset of disease in children and the higher disease activity, belimumab represents an important treatment option for this sub-group of patients. Belimumab has demonstrated a protective role in reducing the accumulation of damage to organs, as well as being a steroid-sparing treatment. This is of

		increased importance in young patients who will potentially face a greater disease and treatment burden throughout their life. “Patients with juvenile onset systemic lupus erythematosus (jSLE) have an accelerated risk of developing atherosclerosis, and cardiovascular disease (CVD) is a leading cause of mortality for patients.“ “The increased risk in jSLE is not explained by traditional CVD risk factors alone but is likely driven by interplay between disease associated chronic inflammation, traditional CVD risk factors (including dyslipidaemia) and risk factors associated with steroid treatment.” *Increased apolipoprotein-B:A1 ratio predicts cardiometabolic risk in patients with juvenile onset SLE.* *https://www.thelancet.com/journals/ebiom/article/PIIS2352-3964(21)00036-0/fulltext* Due to the small population of jSLE patients who may meet the criteria for treatment with belimumab, it is not possible to produce clinical effectiveness data similar to the adult population. Further prospective studies, ideally robust randomised controlled trials, are urgently needed to obtain more accurate date about belimumab in children. jSLE patients are typically treated based upon evidence from studies in adult-SLE cohorts. The limited evidence that is available for belimumab in jSLE does appear to be favourable: • A subset of the patients in this paper (39 patients) have jSLE. This is a ‘real world’ cohort description of belimumab use - https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5031077/#!po=0.877193 “In a subset of 39 patients with childhood-onset SLE, 65% responded favourably at 6 months, and 35% discontinued corticosteroids.” • Management of Paediatric Systemic Lupus Erythematosus: Focus on Belimumab - https://www.dovepress.com/front_end/cr_data/cache/pdf/download_1614943057_604213510d24e/dddt-216193-management-of-pediatric-systemic-lupus-erythematosus-focus-.pdf “Belimumab has been shown to be useful in the treatment of serologically active, refractory SLE and is FDA approved for use in SLE patients 5 years of age and older. Preliminary reports on both its efficacy and safety in the paediatric population are positive; however, its usefulness as a corticosteroid-sparing immunosuppressive agent in SLE patients with severe, active LN and/or active NPSLE remains unknown. This may limit its use in cSLE patients, as a large proportion of cSLE patients have renal

		and/or neuropsychiatric involvement. Belimumab may be most beneficial in cases of cSLE with seropositive, moderately active disease in which corticosteroids cannot be tapered or that are refractory to other immunosuppressive agents."
Key issue 2: Some comparators listed in the NICE scope were not included	**YES**	Whilst rituximab was mentioned in the initial scope, it is not a licensed treatment for SLE. In July 2020 NHS England published a clinical commissioning policy 'Rituximab for refractory Systemic Lupus Erythematosus (SLE) in adults and postpubescent children [200402P]' which specifies that rituximab should be considered **after** eligibility for belimumab has been assessed (based on the criteria of TA397) and ruled out. Due to this sequential approach, rituximab should not be considered a direct comparator. It is also worth noting the COVID-19 pandemic has introduced additional need for vaccinations and this may impact future treatment choices for SLE patients. As a B-cell depleter, rituximab can present challenges for important vaccinations. "It is recommended to wait for vaccination **at least 6 months after rituximab infusion**. However, if a vaccine, such as influenza, needs to be administered within a certain time interval, vaccination should be done, although lower vaccine effectiveness is expected." (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5042271/) If we are looking at a requirement for more annual vaccines and/or boosters to protect patients from circulating viruses (including COVID-19 and influenza), there is a risk that those being treated with rituximab may have to risk decreased vaccine efficacy if the immunisation timing does not adequately synchronise with their infusions. This would make rituximab a significantly less attractive treatment option as high-risk patients will face a potential increased risk of serious infection.
Key issue 3: Short follow-up in the main comparative trials (BLISS-SC, BLISS-52 and BLISS-76)	**NO**	N/A

Key issue 4: Using the propensity score-matching (PSM) analysis in calibrating the cost-effectiveness model can severely bias the results in favour of belimumab	**NO**	N/A
Key issue 5: Data from the BILAG Biologic Register (BILAG BR) are not suitable for a comparison of belimumab with rituximab	**NO**	N/A
Key issue 6: Rituximab + standard therapy was not included as comparator in the model	**YES**	Please see our response to Key issue 2.

Key issue 7: IV and SC formulations are not compared with each other, as two separate model files are provided.	**YES**	Real-world experience within the NHS demonstrated slow recruitment into the BENLYSTA sub study of the BILAG-BR. The overly restrictive HDA-1 population is an important factor, but there are also concerns about equity of access for people who may meet these criteria. Only a minority of patients are under the care of specialist centres where belimumab could be administered. A Rare Disease UK study (https://www.raredisease.org.uk/media/1601/centres-of-excellence.pdf) has previously shown that only 27% of patients with rare diseases are cared for in specialist centres. This presents a barrier to access for some patients who may live away from a specialist centre or have difficulty travelling due to their ill-health/finances/employment/childcare requirements. As such, for a range of reasons, some people may be less likely to benefit from this treatment if it continues to be administered only at specialist centres by IV infusion. Offering belimumab by sub-cutaneous injection would remove a significant barrier for many lupus patients who meet the eligibility criteria and are most likely to benefit from this treatment. Considering the proportion of lupus patients who are of working age, it would also reduce the treatment burden presented by 4-weekly, hour long infusions (with additional travelling time) which can incur a considerable amount of time away from the workplace and may disproportionately disadvantage patients from poorer backgrounds. In the clarification response GSK reports that in a follow-up study exploring patient experiences with the autoinjector, and those of switching from IV to SC belimumab, the majority of participants indicated they preferred the autoinjector to the IV, and were confident in the use of the autoinjector, rating it as convenient and easy to use. The burden of treatment needs to be appropriately considered, taking into account patient preference and barriers to access.
Key issue 8: Use of calibration factor is likely biasing results	**NO**	N/A
Key issue 9: Implementation	**NO**	N/A

of 24-week response and treatment continuation in the model is inconsistent		
Key issue 10: Error in calculation of belimumab non-responder disease activity at 52 weeks	**NO**	N/A
Key issue 11: Violation in utility estimation	**NO**	N/A
Key issue 12: Uncertainty about organ damage utility multipliers	**NO**	N/A
Key issue 13: Sampling of organ damage and death occurs after allocation to treatment	**NO**	N/A

Additional issues

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 (e.g. at the clarification stage).

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: Equality considerations.	Critique of Company's Definition of Decision Problem – Other relevant factors (p37)	**YES**	SLE is a disease predominantly affecting women and those from black, Asian and minority ethnic (BAME) backgrounds. There are important additional considerations because the burden of disease and treatment can be more significant within some ethnic groups. The administration of IV belimumab at specialist centres presents a potential barrier to access for some patients, especially those living further from specialist centres, of working age, with childcare requirements and on lower incomes. The 4-weekly, hour long infusions (with additional travelling time and cost) can incur a considerable burden. Many lupus patients of working age may choose not to access this treatment if they fear their employment may be put at risk by regular absences. This may also be a barrier for patients unable to arrange suitable childcare. Black and minority ethnic (BAME) households in the UK are over twice as likely to live in poverty as their white counterparts. This suggests that the barrier to accessing belimumab for cost-related reasons are likely to be more predominant amongst people from these communities. People from these BAME backgrounds are also more likely to experience severe disease and higher rates of premature mortality. People from these ethnic groups are already at a higher risk of developing diabetes and hypertension. It should be considered whether steroid-sparing treatments such as belimumab could present additional advantages over

			standard treatments by reducing some adverse events and risks of comorbidities. It also needs to be noted that double-stranded-DNA antibodies are less common in patients of African descent, so it could be perceived as discriminatory to stipulate dsDNA antibody positivity as a criterion and not consider other lupus-related antibodies.

NICE National Institute for Health and Care Excellence

Summary of changes to the company's cost-effectiveness estimate(s)

Company: If you have made changes to the company's preferred cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes.

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 ICER
Insert key issue number and title as described in the ERG report	Briefly describe the company's original preferred assumption or analysis	Briefly describe the change(s) made in response to the ERG report	Please provide the ICER resulting from the change described (on its own), and the change from the company's original base-case ICER
Company's preferred base case following technical engagement	Incremental QALYs: [QQQ]	Incremental costs: [£££]	Please provide the revised company base-case ICER resulting from combining the changes described, and the change from the company's original base-case ICER

NICE National Institute for Health and Care Excellence

Technical engagement response form

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (Review of TA397) [ID1591]

As a stakeholder you have been invited to comment on the ERG report for this appraisal. 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.

We need your comments and feedback on the key issues below. You do not have to provide a response to every issue. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be included in the committee papers in full and may also be summarised and presented in slides at the appraisal committee meeting.

Deadline for comments **5:00pm on 11 March 2021**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the ERG report which summarises the background and submitted evidence, and presents the ERG's summary of key issues, critique of the evidence and exploratory analyses. This will provide context and describe the questions below in greater detail.
- Please ensure your response clearly identifies the issue numbers that have been used in the executive summary of the ERG report. 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.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, 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.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	[redacted] **&** [redacted]
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**British Society for Rheumatology**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**N/A**

NICE National Institute for Health and Care Excellence

Key issues for engagement

Please use the table below to respond to questions raised in the ERG report on key issues. You may also provide additional comments on the key issue that you would like to raise but which do not address the specific questions.

Key issue	Does this response contain new evidence, data or analyses?	Response
Key issue 1: Evidence is missing for specific populations, such as children and patients with severe active central nervous system (CNS) lupus.	**No**	The results from the PLUTO study for children were included. Paediatric units located in England are now using IV belimumab in children if they are 5 years of age or older and meet the criteria. Numbers will be small because paediatric onset lupus is rarer than adult onset lupus. The pivotal studies excluded severe CNS lupus. IV cyclophosphamide and rituximab (sometimes in combination) are used to treat this rare and severe manifestation. This manifestation is often excluded from lupus RCTs.
Key issue 2: Some comparators listed in the NICE scope were not included	**No**	The appropriate comparators have been included here.
Key issue 3: Short follow-up in the main comparative trials (BLISS-SC, BLISS-52 and BLISS-76)	**No**	BLISS 52 was for 52 weeks and BLISS-76 was for 76 weeks. This (i.e. a year) is the standard duration for a lupus RCT. Subsequent observational data are available. Lupus studies will always face the issue that the progression of damage (from lupus or therapeutics) occurs at a rate that is slower than the reasonable timeframe over which a randomised trial could run. The complexity of extrapolating from disease activity measures over a 12-18 month period and long term disease damage has been discussed.

Key issue 4: Using the propensity score-matching (PSM) analysis in calibrating the cost-effectiveness model can severely bias the results in favour of belimumab	**No**	We do not have the expertise to add anything here only to comment that the propensity score-matching methodology has been subject to previous peer-review
Key issue 5: Data from the BILAG Biologic Register (BILAG BR) are not suitable for a comparison of belimumab with rituximab	**Yes**	There are over 900 patients in BILAG-BR who have received rituximab. Patients who meet the SLEDAI, anti dsDNA positivity and low complement criteria could be identified and the data compared. The number of patients who meet these criteria will be small as the belimumab criteria are very stringent. It will be around 90-100 patients for HAD-1. Such a search has previously been performed for the HAD-1 population and a letter was published by IN Bruce et al reporting this finding (Rheumatology (Oxford). 2017 Jun 1;56(6):1041-1043. doi: 10.1093/rheumatology/kex044.). The HAD-2 definition is more pragmatic and so there would be more patients who have received rituximab and meet these criteria and for whom the outcomes could be compared.
Key issue 6: Rituximab + standard therapy was not included as comparator in the model	**Yes**	The EXPLORER study did not meet its primary endpoint and this may have been related to trial design including the allowance of high doses of prednisolone and patients with very high disease activity were included in the study. Other available data would be through registry data. BILAG-BR published outcome data for rituximab (Rheumatology (Oxford) 2018 Mar 1;57(3):470-479. doi: 10.1093/rheumatology/kex395)**.**
Key issue 7: IV and SC formulations are not compared with each other, as two separate model files are provided.	**YES/NO**	We have been permitted to use S/C belimumab during the pandemic and it has been met very favourably from the patient perspective – fewer hospital visits/less time off work required. We can now start patients on s/c belimumab but we have chosen to give 3 pulses of belimumab using usual loading regime and then switch to s/c 3 weeks later. This may facilitate a faster response i.e. the patients get a better loading dose. A S/C

		loading dose has not been proposed as for some biologics. I am not aware of any published switching data.
Key issue 8: Use of calibration factor is likely biasing results	**YES/NO**	No comment
Key issue 9: Implementation of 24-week response and treatment continuation in the model is inconsistent	**YES/NO**	No comment
Key issue 10: Error in calculation of belimumab non-responder disease activity at 52 weeks	**YES/NO**	No comment
Key issue 11: Violation in utility estimation	**YES/NO**	No comment
Key issue 12: Uncertainty about organ damage utility multipliers	**YES/NO**	No comment
Key issue 13: Sampling of organ damage and death occurs after allocation to treatment	**YES/NO**	No comment

Additional issues

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 (e.g. at the clarification stage).

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**	No comment
Additional issue 2: Insert additional issue	Please indicate the section(s) of the ERG report that discuss this issue	**YES/NO**	No comment
Additional issue N: Insert additional issue			[INSERT / DELETE ROWS AS REQUIRED]

NICE National Institute for Health and Care Excellence

Summary of changes to the company's cost-effectiveness estimate(s)

Company: If you have made changes to the company's preferred cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes.

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 ICER**
Insert key issue number and title as described in the ERG report	Briefly describe the company's original preferred assumption or analysis	Briefly describe the change(s) made in response to the ERG report	Please provide the ICER resulting from the change described (on its own), and the change from the company's original base-case ICER
..	..	..	[INSERT / DELETE ROWS AS REQUIRED]
Company's preferred base case following technical engagement	**Incremental QALYs: [QQQ]**	**Incremental costs: [£££]**	**Please provide the revised company base-case ICER resulting from combining the changes described, and the change from the company's original base-case ICER**

NICE National Institute for Health and Care Excellence

Technical engagement response form

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (Review of TA397) [ID1591]

As a stakeholder you have been invited to comment on the ERG report for this appraisal. 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.

We need your comments and feedback on the key issues below. You do not have to provide a response to every issue. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be included in the committee papers in full and may also be summarised and presented in slides at the appraisal committee meeting.

Deadline for comments **5:00pm on 11 March 2021**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the ERG report which summarises the background and submitted evidence, and presents the ERG's summary of key issues, critique of the evidence and exploratory analyses. This will provide context and describe the questions below in greater detail.
- Please ensure your response clearly identifies the issue numbers that have been used in the executive summary of the ERG report. 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.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, 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.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	[redacted]
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**Renal Association**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None re tobacco. Updated disclosure re GSK – I am now Co-Chair International Advisory Panel, GSK Global GUIDE program – Virtual Scientific Forum that will meet 3-4 times per year. 1st meeting will be on 17/3/2021. This is a renumerated role. I am also on the speaker bureau for GSK sponsored symposia at EULAR and ERA meetings in June 2021.**

NICE National Institute for Health and Care Excellence

Key issues for engagement

Please use the table below to respond to questions raised in the ERG report on key issues. You may also provide additional comments on the key issue that you would like to raise but which do not address the specific questions.

Key issue	**Does this response contain new evidence, data or analyses?**	Response
Key issue 1: Evidence is missing for specific populations, such as children and patients with severe active central nervous system (CNS) lupus.	No	Agree minimal data presented for Children beyond the data from the PLUTO study but lupus in children is rare and hard to get trial data in this population. But it is licensed for those aged 5 and over and we use other drugs (eg rituximab) similarly in children to how we use in adults. The data from the PLUTO study are entirely consistent with data from the trials in adults – for instance see this review: https://pubmed.ncbi.nlm.nih.gov/32612353/ No one is suggesting using belimumab for those with active CNS disease and it's outside the scope of the current marketing authorisation. This does not detract from belimumab's usefulness in non CNS lupus.
Key issue 2: Some comparators listed in the NICE scope were not included	**No**	For the current marketing authorisation of belimumab, it is inappropriate to have cyclophosphamide as a comparator – cyclophosphamide is really only used for lupus nephritis or really severe systemic disease including active CNS disease; both of these are currently outside of licensing in the UK though the FDA has now approved belimumab for use in lupus nephritis. In the BLISS-LN trial, SoC could either be the cyclophosphamide-based Eurolupus regimen or an MMF based regimen – the latter was used in the majority. However, the purpose of the current

		submission, the use of belimumab being evaluated is in comparison to patients who would not routinely receive cyclophosphamide. Rituximab is used very differently to belimumab. It is essentially used to treat flare and not as a maintenance treatment in lupus. The current NHSE guidance suggests to use belimumab before rituximab so a different group of patients would be being compared – that in fact might bias towards belimumab as in theory rituximab should only be used for those with disease refractory to belimumab. There are no positive trials to compare belimumab with rituximab – and the argument about the different steroid usage in EXPLORER vs the BLISS studies is compelling.
Key issue 3: Short follow-up in the main comparative trials (BLISS-SC, BLISS-52 and BLISS-76)	**No**	The follow-up in the BLISS trials is no shorter than most lupus trials and have the advantage of the long term extension studies, the real world data and from a UK perspective, the BILAG-BR data.
Key issue 4: Using the propensity score-matching (PSM) analysis in calibrating the cost-effectiveness model can severely bias the results in favour of belimumab	**N**o	I am not an expert on the methodology but PSM is increasingly used to compare non directly compared groups; they have gone to great lengths to match. I am persuaded by their choice of cohort based on severity and standard of care and that the patients in the Toronto cohort did not have access to belimumab so there was no treatment choice bias. The different ethnic make up of the Toronto cohort to a more multi ethnic UK population remains a problem in the PSM analysis.
Key issue 5: Data from the BILAG Biologic Register (BILAG BR) are not suitable for a comparison of belimumab with rituximab	No	Agree – very different populations receive rituximab for lupus in the UK.
Key issue 6: Rituximab + standard therapy was not included as comparator in the model	**N**o	See my response to key issue 2. Rituximab has not been effective in any RCT to date despite a lot of RWE for effectiveness. However, it is used quite differently to belimumab and particularly in the UK, is mandated to be used AFTER belimumab

		and so will be being used in a more refractory patient population. This would in fact lead to a major bias in favour of belimumab
Key issue 7: IV and SC formulations are not compared with each other, as two separate model files are provided.	No	I don't think this is an issue… If the results with SC are more or less comparable with the IV trials, I see no reason for a head to head study.
Key issue 8: Use of calibration factor is likely biasing results	No	This is outside my scope of expertise but I am persuaded by their arguments on how the calibration factors were derived and used.
Key issue 9: Implementation of 24-week response and treatment continuation in the model is inconsistent	No	Not appropriate for me to answer. Answered by the company
Key issue 10: Error in calculation of belimumab non-responder disease activity at 52 weeks	No	Not appropriate for me to answer. Answered by the company
Key issue 11: Violation in utility estimation	No	I am not in a position to commen
Key issue 12: Uncertainty about organ damage utility multipliers	No	Always an imperfect science but weighting seems appropriate
Key issue 13: Sampling of organ damage and death occurs after allocation to treatment	No	Not appropriate for me to answer. Answered by the company

Additional issues

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 (e.g. at the clarification stage).

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:	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]

NICE National Institute for Health and Care Excellence

Summary of changes to the company's cost-effectiveness estimate(s)

Company: If you have made changes to the company's preferred cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes.

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 ICER
Insert key issue number and title as described in the ERG report	Briefly describe the company's original preferred assumption or analysis	Briefly describe the change(s) made in response to the ERG report	Please provide the ICER resulting from the change described (on its own), and the change from the company's original base-case ICER
..	..	..	[INSERT / DELETE ROWS AS REQUIRED]
Company's preferred base case following technical engagement	Incremental QALYs: [QQQ]	Incremental costs: [£££]	Please provide the revised company base-case ICER resulting from combining the changes described, and the change from the company's original base-case ICER

in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (review of TA397) [ID1591]

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
Sabine Grimm, Health Economist, Maastricht UMC, the Netherlands
Ben Wijnen, Health Economist, Maastricht UMC
Annette Chalker, Systematic Reviewer, KSR Ltd
Sean Harrison, Systematic Reviewer, KSR Ltd
Nigel Armstrong, Health Economist, KSR Ltd
Thomas Otten, Health Economist, Maastricht UMC
Mohammed Islam, Health Economist, KSR Ltd
Gill Worthy, Statistician, KSR Ltd
Lisa Stirk, 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, UK
YO19 6FD

Date completed 24/03/2021

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 report.[1]

In their response to technical engagement, the company submitted responses to the key issues raised in the Technical Report written by the NICE technical team, and some additional evidence relevant to these issues.[1]

1. Evidence is missing for specific populations, such as children and patients with severe active central nervous system (CNS) lupus

As stated by the company, the current marketing authorisation, as per Summary of Product Characteristics for belimumab IV[2] and SC[3] formulations, does not include patients with severe active CNS lupus, as these patients were not included in BLISS pivotal IV and SC trials. Therefore, the company acknowledges there is no evidence for the use belimumab in this patient population and do not anticipate NICE to issue a guidance relating to the use of belimumab in patients with severe active CNS lupus.

Regarding children, the company states that "Clinical trial evidence evaluating the safety, efficacy and pharmacokinetics on the use of intravenous (IV) belimumab in children with systemic lupus erythematosus (SLE), derived from Part A of the Phase 2 PLUTO (The Pediatric Lupus Trial of Belimumab Plus Background Standard Therapy) study, was presented in Appendix O of the company submission, comprising of the primary and secondary endpoints and trial safety data."[1]

In addition, the company have now also provided additional analyses of the primary endpoint and its individual components from Part A of the PLUTO study, for the HDA-1 and HDA-2 subgroups.

Currently, there is no clinical data to support the use of belimumab subcutaneous (SC) formulation in patients <18 years old and therefore the marketing authorisation for belimumab in patients aged 5-17 years old is currently restricted to the IV formulation only, according to the company.[1]

ERG comment:

As stated by the company, the PLUTO study was presented in Appendix O of the original submission. However, the data reported in Appendix O are for the total of 93 children (40 in the placebo group and 53 in the belimumab 10 mg/kg group) that participated in Part A of the PLUTO trial. No data were presented for the HDA-2 subgroup that is the population of interest for this appraisal.

The results presented here for the HDA-2 subgroup are similar to the results for the whole PLUTO trial population as presented in Appendix O of the original submission.[1] None of the results presented in Table 2 of the Response to technical engagement (TE) (PLUTO Paediatric IV Study - SRI-4 Response Rate at Week 52 - HDA-2 Subgroup) are statistically significant; which is also the same as the results for the whole trial population.

2. Some comparators listed in the NICE scope were not included

As stated in the ERG report:[4]

- Regarding cyclophosphamide, the ERG believes the company has a point because it is mainly used for different populations than belimumab and the adverse event profile of cyclophosphamide means that it is avoided if possible.
- The comparison with rituximab will be difficult because the evidence for rituximab is weaker as the phase 3 trials were negative due to very stringent end-points (and different to those used for belimumab) and is mostly from registries. BILAG BR data cannot be used to compare them easily

due to the different criteria for the use of rituximab and belimumab (See also Section 3.3 of the ERG report).

3. Short follow-up in the main comparative trials (BLISS-SC, BLISS-52 and BLISS-76)

As stated in the ERG report: BLISS-SC and BLISS-52 had a maximum follow-up of 52 weeks, while BLISS-76 had a maximum follow-up of 76 weeks.[4] All three trials did have long-term extension (LTE) phases. However, all patients received belimumab during the LTE. Therefore, these extension studies do not provide comparative evidence and there is no reliable comparative evidence available beyond 76 weeks.

4. Using the propensity score-matching (PSM) analysis in calibrating the cost-effectiveness model can severely bias the results in favour of belimumab

The company responded to key issue 4, stating the ERG raised three key issues and responded to each in turn. We will respond similarly.

4.1. Patients are not necessarily representative of either BLISS 76 US LTE patients or TLC patients, but representative only of patients matching between these cohorts: this is unlikely to be representative of patients in the UK.

The key issue here is that the participants in both the TLC and the BLISS LTEs have been highly selected into the propensity score matched (PSM) analysis, since all included participants in one cohort needed to have a similar propensity score to a patient in the other cohort. It is not relevant to this point whether the PSM analysis was performed robustly, nor particularly whether the TLC and BLISS LTEs are generalisable to a UK population, although this would help. By selecting subgroups of the two cohorts, the PSM analysis has defined a new cohort that may or may not be representative of either of the two parent cohorts, or indeed the UK population. It might be reasonable to assume little bias in the treatment effect of belimumab vs. SoC given an assumption of treatment effect being independent of the population characteristics. However, the approach taken by the company, at least in the base case, was to use only the outcome for belimumab treated patients to estimate the calibration factor in the model.

The only evidence the company has provided for the generalisability of the participants in the PSM analysis to a UK population was through stating that two clinical experts were consulted on this issue, and that the experts believed the matched participants were clinically reflective of a UK SLE population.

We have independently checked the baseline demographics in BILAG from Appendix P of the company submission with the baseline demographics of the matched participants in the PSM analysis of the BLISS-76 US LTE.[5] The populations seem similar in age, sex, ethnicity, SLE duration and current smoking status. However, rates of hypertension were different (54% in the PSM, [redacted] in BILAG), as were the baseline SLEDAI scores (8.5 in the PSM, [redacted] in BILAG). Other demographic variables presented in the PSM analysis could not be checked, and there are many pertinent variables that were not reported, such as measures of socio-economic outcomes and disease progression. Given the discrepancy in disease severity (as assessed by SLEDAI), we are not certain that the PSM analysis patients are generalisable to the UK SLE population. The clinical experts believed that patients with higher disease activity would have a greater response to belimumab, meaning the generalisation to the UK population would underestimate the treatment effect. This would make any generalisation an underestimate for the effectiveness of belimumab, which would not favour belimumab and thus be conservative.

Therefore, the ERG believes that while there is a potential for bias to affect the generalisability of results from the matched PSM cohort to UK SLE patients, there is not sufficient evidence to state that the bias would favour belimumab.

4.2. The application of the calibration factor derived at 5 years to the whole 5 years.

In Figure 1 of the response to technical engagement,[1] the company presented a plot based on estimated total SDI change from baseline to each year for years 1-5 using the PSM and compared this with plots of change from baseline given different calibration coefficients. However, it is unclear precisely how these calibration coefficients were estimated: the company stated that they "…explored the possibility that the calibration factor could be different had a different time point been selected on which to apply to in the model." However, the calibration coefficients were not labelled with any time points. Also, the plots are so diverse that it suggests the method used to estimate them probably clearly lacked face validity and therefore does not provide any plausible alternatives or validation for the base case calibration coefficient of 0.491.

There is also a question as to whether the SDI change PSM estimates were derived from the whole cohort of eligible patients in the PSM analysis (for example in the analysis up to year 2, including all participants with at least 2 years of follow-up data), or whether this is solely from patients with 5 years of follow-up data. If the latter, then there would be a bias inherent in selecting participants who have taken the drug for 5 years: these patients are much more likely to either have reduced disease activity regardless of their treatment, or belimumab has a larger than average effect in these participants, since disease progression despite treatment increases the probability of discontinuation. If the former, then this selection bias would not apply. However, we are not able to determine which of these situations applies, as we do not have access to the clinical study report referenced in the technical engagement: it was not provided to us as additional evidence. In fact, the company stated that their response to this issue contained no new evidence, data or analyses.

However, regardless of which patients were selected to estimate SDI at each of the years 1 to 4, there would remain the likely underestimation of progression at years 1 to 3 as can be seen by the straight line based on the calibration factor of 0.491 lying under the curve based on the PSM for the first 3 years, increasing the apparent effectiveness of belimumab. Specifically, the effectiveness of belimumab would be overestimated by the difference between the observed SDI change (solid black line in Figure 1 of the technical engagement response) and the 0.491 calibration factor (solid blue line in figure 1). Indeed, it is likely that if patients with 5 years of follow-up data had been used then using data from the whole cohort would probably increase the discrepancy between the resulting PSM-based curve and straight line based on the calibration factor. This is particularly problematic given patients withdraw from belimumab over time in the model, meaning more patients are affected by the calibration factor in years 1-3 than in years 4 and 5.

4.3. The calibration factor derived from the PS-matched comparative analysis of 0.491 effectively doubles the effectiveness of belimumab for preventing organ damage, compared with the JH model.

The ERG does not doubt that reducing organ damage may result in clinical benefit, and that this benefit is unlikely to become apparent within the relatively short time frames of the BLISS RCTs. The issue is in the estimation of this benefit. The ERG believes that the use of the calibration factor overestimates the benefit, due to the selection of patients previously mentioned. The calibration factor may also be biased through other means, including generalising to a different population, though this likely just increases the uncertainty and imprecision of the calibration factor. This uncertainty, including from the SDI change estimates themselves, is not accounted for by the company.

The ERG believes that the key issue has not been addressed, namely that patients who continue on belimumab for 5 years are likely to have progressed less than patients who took belimumab for 1, 2, 3 or 4 years before discontinuing, and therefore applying the calibration factor estimated based on 5 years to

all time points up to 5 years probably overestimates the effectiveness of belimumab for the preceding years.

5. Data from the BILAG Biologic Register (BILAG BR) are not suitable for a comparison of belimumab with rituximab

The company acknowledged this point which is why they did not perform a formal indirect treatment comparison of rituximab and belimumab based on BILAG-BR data.

6. Rituximab + standard therapy was not included as comparator in the model

The ERG acknowledges that including rituximab + standard therapy would have been difficult. The absence of this potentially relevant comparison remains a limitation.

7. IV and SC formulations are not compared with each other, as two separate model files are provided.

The company provided evidence supporting similar efficacy between IV and SC formulations, which is also in line with the expectations of the ERG's clinical expert. Ideally both formulations would be compared against standard therapy in one incremental analysis.

8. Use of calibration factor is likely biasing results

The ERG maintains that the use of the calibration factor likely introduces a bias in favour of belimumab. Please see the detailed response to Issue 4. In addition, although the company clarified that in the model the calibration factor was only applied to those patients remaining on belimumab treatment, it remained unclear whether the calibration factor was derived by calibrating the SDI scores in the overall belimumab treatment arm or the SDI scores of responders. The former method would likely over-estimate the effect of the calibration factor on organ damage.

9. Implementation of 24-week response and treatment continuation in the model is inconsistent

Unfortunately, despite the ERG's request at Technical Engagement, the company did not provide evidence to show that it is only the "chronology of events" that caused belimumab non-responders to have a >4 point reduction in SS scores at 52 weeks. This evidence could have been provided, as previously highlighted and discussed, by examining SS score reductions at 24 weeks of those patients that were classed as non-responders and had a >4 point reduction in SS scores at 52 weeks. It therefore remains unclear whether the implementation in the model is in line with the evidence.

10. Error in calculation of belimumab non-responder disease activity at 52 weeks

The ERG acknowledges the company's clinical experts' expectations of similar disease activity for belimumab non-responders and ST treated patients and also that improvement in disease may be possible within the first 52 weeks after discontinuing belimumab. However, currently the model does not capture any disadvantage belimumab non-responders may experience in the first 52 weeks because the first time the SS score is calculated after baseline is at 52 weeks. This is not in line with the evidence from the BLISS trials and the company do not provide an explanation for this discrepancy between evidence from BLISS and their assumption. Hence, the company's implementation introduces bias in favour of belimumab when compared to the evidence from BLISS trials. The company altered this implementation in a scenario (model file not provided to the ERG). This scenario reduces the impact of the difference in belimumab non-responders and ST treated patients by only applying it to 24 weeks instead of 52 weeks, but adds further impact by delaying the full return to ST efficacy by further 6 months. The result is an increase in the ICER slightly higher than that of the preferred ERG scenario. Given that the ERG did not

receive the company's model file, the ERG continues to use the original implementation of the company that can be found in the model code but that the company abandoned in their base-case, using the BLISS evidence for belimumab non-responders to incorporate this difference between belimumab non-responders and ST treated patients in the first 52 weeks (as detailed as model change 1 in ERG report).

11. Violation in utility estimation

The company did not fix the violation in utility estimation but instead provided analyses exploring the potential impact of this error on ICERs. The ERG agrees that the company's scenarios likely explore the full impact but also notes that the impact is not necessarily minor: The ICER increased / decreased by approximately £3,000 with only one of the coefficients varied and could increase / decrease further with combinations of coefficients varied. However, the ERG agrees with the company that the variation by 1 standard deviation is likely substantial. The ERG considers that this potential uncertainty about the ICERs should be borne in mind in decision-making.

12. Uncertainty about organ damage utility multipliers

The ERG appreciates that the company have addressed the highlighted uncertainty about organ damage multipliers by re-consulting with clinical experts. The impact of the company's scenarios on the ICERs was small. The ERG is therefore satisfied that, while uncertainty remains about organ damage impact on HRQoL, the company have addressed it as well as they could and the impact may be minor.

13. Sampling of organ damage and death occurs after

The ERG maintains its original position that good practice dictates that sampling error be minimized by first sampling all patient characteristics and then allocating to treatment arms. We appreciate that the company simulated a large number of patients, likely sufficient to mitigate any sampling error, but our critique stands – it is not in line with best practices and makes validation more difficult. That said, we think that this issue is less important to correct in this assessment, as this may not have an effect on model outcomes, as stated in the ERG report and also stated by the company.

Analyses

The ERG did not change the ERG base-case analyses. The analyses using the company's list price for the SC formulation are presented in Tables 1 and 2.

Table 1: ERG base-case using list price for SC formulation

Technologies	Total costs	Total LYs	Total QALYs	Incremental costs	Incremental QALYs	ICER (/QALY)
Replicated CS base-case						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£53,583
Standard therapy	£151,999	17.122	10.056			
Fixing errors 1: 1st year: SS reduction for beli non-responders						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£56,963
Standard therapy	£151,999	17.122	10.056			
Matter of judgement 2: Calibration factor removed = ERG base-case						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£94,139
Standard therapy	£151,999	17.122	10.056			

Table 2: ERG scenarios using list price for SC formulation

Technologies	Total costs	Total LYs	Total QALYs	Incremental costs	Incremental QALYs	ICER (/QALY)
ERG base-case						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£94,139
Standard therapy	£151,999	17.122	10.056			
Scenario 1: Use unadjusted JH model						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£106,867
Standard therapy	£151,873	18.369	11.036			
Scenario 2: Use calibration factor as per company's base-case						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£56,963
Standard therapy	£151,999	17.122	10.056			
Scenario 3: Use calibration factors on both arms						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£47,626
Standard therapy	£158,791	17.000	9.916			
Scenario 4: Remove impact of organ damage on HRQoL						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£87,730
Standard therapy	£151,999	17.122	12.082			
Scenario 6: HDA-1 subgroup						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£93,591
Standard therapy	£156,692	17.679	10.476			

REFERENCES

[1] National Institute for Health and Care Excellence. *Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (review of TA397) [ID1591]: Technical engagement response form*. London: NICE, 2021. 33p.

[2] EMC. *Benlysta 120 mg powder for concentrate for solution for infusion SmPC*, 2020 Available from: https://www.medicines.org.uk/emc/product/4679/smpc

[3] EMC. *Benlysta 200 mg solution for injection in pre-filled pen SmPC*, 2020 Available from: https://www.medicines.org.uk/emc/product/11398/smpc

[4] Riemsma R, Grimm S, Wijnen B, Chalker A, Harrison S, Armstrong N, et al. *Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (review of TA397) [ID1591]: a Single Technology Assessment*. York: Kleijnen Systematic Reviews Ltd, 2021. 128p.

[5] GlaxoSmithKline. *A multi-center, continuation trial of belimumab (HGS1006, LymphoStat-B), a fully human monoclonal anti-BLyS antibody, in subjects with systemic lupus erythematosus (SLE) who completed the phase 3 protocol HGS1006-C1056 in the United States. Data on file. 2 December 2015.*

in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (review of TA397) [ID1591]

ADDENDUM: Updated PAS for the IV formulation

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 Sabine Grimm, Health Economist, Maastricht UMC, the Netherlands Ben Wijnen, Health Economist, Maastricht UMC Annette Chalker, Systematic Reviewer, KSR Ltd Sean Harrison, Systematic Reviewer, KSR Ltd Nigel Armstrong, Health Economist, KSR Ltd Thomas Otten, Health Economist, Maastricht UMC Mohammed Islam, Health Economist, KSR Ltd Gill Worthy, Statistician, KSR Ltd Lisa Stirk, 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, UK YO19 6FD
Date completed	20/04/2021

Company's updated PAS for the intravenous formulation

The company have provided a new PAS for the IV formulation of belimumab. This PAS is a discount of comprises a confidential common discount of [redacted] for the Benlysta intravenous (IV) and subcutaneous (SC) formulations. This results in no changes to the cost-effectiveness analyses for the subcutaneous formulation. For the IV formulation, this new PAS results in small changes to the price per vial are now [redacted] for the 120mg vial and [redacted] for the 400mg vial.

The following are the revised ERG results using the company's new prices per vial for the IV population (ERG base-case in Table 1 and ERG scenarios in Table 2). It should be noted that using the discount of [redacted], the ERG calculated vial prices that are slightly different to the ones mentioned above, and these differences are in the order of magnitude of [redacted] for the 120mg vial and [redacted] for the 400mg vial. These differences are likely due to the discount of [redacted] being rounded up.

The ERG's critique of the company's analyses remains unchanged.

Table 3 presents the ERG base-case results for the SC formulation, showing the impact of the two changes separately first and then cumulatively in the ERG base-case.

Table 1: ERG analyses for IV formulation using updated IV PAS (deterministic unless indicated)

Technologies	Total costs	Total LYs	Total QALYs	Incremental costs	Incremental QALYs	ICER (/QALY)
CS base-case						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£29,162
Standard therapy	£160,470	16.900	9.809			
Fixing errors 1: 1st year: SS reduction for belimumab non-responders						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£30,839
Standard therapy	£160,470	16.900	9.809			
Matter of judgement 2: Calibration factor removed						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£49,202
Standard therapy	£160,470	16.900	9.809			
ERG base-case (changes 1 and 2)						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,817
Standard therapy	£160,470	16.900	9.809			
ERG base-case (probabilistic, changes 1 and 2)						
Belimumab				[redacted]	[redacted]	£53,910
Standard therapy						

Table 2: ERG scenarios for the IV formulation using updated IV PAS (deterministic)

Technologies	Total costs	Total LYs	Total QALYs	Incremental costs	Incremental QALYs	ICER (/QALY)
ERG base-case						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,817

Standard therapy	£160,470	16.900	9.809			
Scenario 1: Use unadjusted JH model						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£62,607
Standard therapy	£161,467	18.187	10.798			
Scenario 2: Use company's calibration factors						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£30,839
Standard therapy	£160,470	16.900	9.809			
Scenario 3: Use calibration factors on both arms						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£24,050
Standard therapy	£167,261	16.764	9.669			
Scenario 4: Remove impact of organ damage						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£47,366
Standard therapy	£160,470	16.900	11.941			
Scenario 5: Patient weight based on trial						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£49,433
Standard therapy	£160,470	16.900	9.809			
Scenario 6: HDA-1 subgroup						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£47,782
Standard therapy	£166,658	17.468	10.216			

Table 3: ERG analyses for SC formulation using original SC=updated IV PAS (deterministic)

Technologies	Total costs	Total LYs	Total QALYs	Incremental costs	Incremental QALYs	ICER (/QALY)
CS base-case						

Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£30,566
Standard therapy	£151,999	17.122	10.056			
Fixing errors 1: 1st year: SS reduction for belimumab non-responders						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£32,617
Standard therapy	£151,999	17.122	10.056			
Matter of judgement 2: Calibration factor removed						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£56,277
Standard therapy	£151,999	17.122	10.056			
ERG base-case (changes 1 and 2)						
Belimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£61,057
Standard therapy	£151,999	17.122	10.056			

1 Non-response and treatment continuation in the model (key issue 9 in ERG report).

a. Please provide the number of all modelled belimumab non-responders.

In the IV HDA-2 subgroup, 3,414 out of 10,000 simulated patients on belimumab treatment are non-responders (34.1%) at week 24.

b. Please provide the number of all modelled belimumab non-responders with a SS-reduction of >=4 points at 52 weeks.

In the IV HDA-2 subgroup, 1,588 non-responders (out of 3,414) had a SS-reduction of >=4 points at 52 weeks (46.5% of all belimumab non-responders).

c. For all these, please provide their SS-reduction at 24 weeks.

The mean SS reduction at 24 weeks for non-responders with an SS reduction of >=4 points at 52 weeks was 0.41. The SS reduction at week 24 ranged between 0.31 and 0.85, which means that there were no belimumab patients that were modelled as non-responders that had a SS-reduction of >=4 points at 24 week.

The SS score of belimumab non-responders at 24 weeks was modelled directly based on a regression model that was fitted to the BLISS trial data. This regression model was used for the modelling of all belimumab (responders and non-responders) and standard treatment SS scores at 24 weeks.

However, as the regression model provides SS estimates that reflect the mean score of the relevant sub-population, the range of the modelled SS reduction is considerably narrower than the range of SS reduction in the trial population.

d. Please also provide the number of patients that were modelled as non-responders that had a SS-reduction of >=4 points at 24 weeks. If this is positive, please explain this - and consider proposing an exploratory analysis to assess the impact of this on the ICER. If this is nought, please provide a clinical explanation for these belimumab non-responders' disease activity continuing to improve after 24 weeks.

As outlined in the answer to question 1c above, there were no patients that were modelled as non-responders that had a SS-reduction of >=4 points at 24 weeks.

GSK would like to confirm that the implementation of belimumab non-responders and treatment continuation in the IV model is in line with the evidence derived from the pooled Phase 3 IV trials (and the BLISS SC trial for the SC model). In addition, no belimumab patients that experienced a SELENA-SLEDAI reduction of ≥ 4 points at 24 weeks were modelled as non-responders. In the model, belimumab non-responders at week 24 discontinued belimumab and continued standard therapy in line with NICE TA397 guidance and the Summary of Product Characteristics. In the model, 46.5% of belimumab non-responders at week 24 had an SS score reduction of >=4 points at 52 weeks.

Based on our understanding from UK Rheumatologists, patients receiving belimumab who do not respond at week 24/6 months (defined by a < 4-point reduction in SELENA-SLEDAI in the health economic model), would cease treatment with belimumab and receive alternative treatments, which is in line with current NICE TA397 guidance and the Summary of Product Characteristics.

SLE patients have a high morbidity and mortality and therefore standard practice to seek alternative treatments for belimumab non-responder patients, which will result in a potential improvement in their disease activity (defined by the SELENA-SLEDAI) over time. However, as there are limited other treatments available, clinicians would usually aim to further optimise standard therapy by changing the patient's immunosuppressant (assuming 1 or 2 prior immunosuppressants), with an increase in steroid dose whilst response to treatment is assessed over the next 3-6 months. Other non-standard therapies may also be considered if the patient continues to have active disease. Therefore, it is not unreasonable to assume a further improvement in SELENA-SLEDAI score and in line with the average ST score by week 52 in a belimumab non-responder patient.

We note that the clinical expert opinion obtained by the ERG accepted as a reasonable assumption that belimumab non-responders (as assessed at week 24) have equal disease activity to the average patient on standard treatment beyond 52 weeks, see ERG report page 96. The company would like to highlight that a scenario analysis exploring an alternative modelling approach for the SS score development of belimumab non-responders has previously been submitted for the committee's consideration. In this scenario, return to standard therapy efficacy for belimumab non-responders was assumed to occur after one full year of standard therapy treatment alone (i.e. after week 76). This represents a delay of standard therapy efficacy compared with the base case where standard therapy efficacy was applied after 28 weeks of standard therapy (i.e. after week 52). This delay translates into a lower SS score reduction for non-responders at week 52. Results of the analysis for the IV formulation showed the approach had only a small impact on the ICER which was increased by less than £2,000 compared with the base case analysis, resulting in an ICER of £31,048.

2 Health-related quality of life

a. Could the company please provide utility values for the average modelled patient treated with belimumab and treated with standard therapy for baseline, year 1, year 6, year 10, year 20 for both the company's base-case and a scenario in which the calibration factor is not used.

The average utility values of patients treated with belimumab and treated with standard therapy for different time points are presented in Table 1 and Figure 1. Although utility scores are presented until 20 years, it is worth noting that in the base case model, the calibration factor is only applied for patients on belimumab for a maximum of 6 years, as long as they are a responder at week 24 and do not discontinue at week 52 for any reason. Furthermore, the calibration factor is only applied proportionally (i.e., for example, patients who discontinue due to non-response at the end of the third year of receiving belimumab, only receive 3 years' worth of benefit and not the full 6 years).

Table 1. Average utilities by treatment group over time

Model cycle (years)	Standard therapy (base case and scenario)	Belimumab (base case where calibration factor applied for first 6 years)*	Belimumab (no calibration factor applied scenario)
Baseline	0.586	0.599	0.595
1	0.615	0.642	0.635
6	0.601	0.626	0.616
10	0.578	0.597	0.590
20	0.536	0.547	0.543

* CF applied from 0-6 years only, irrespective of model length, to belimumab responders only.

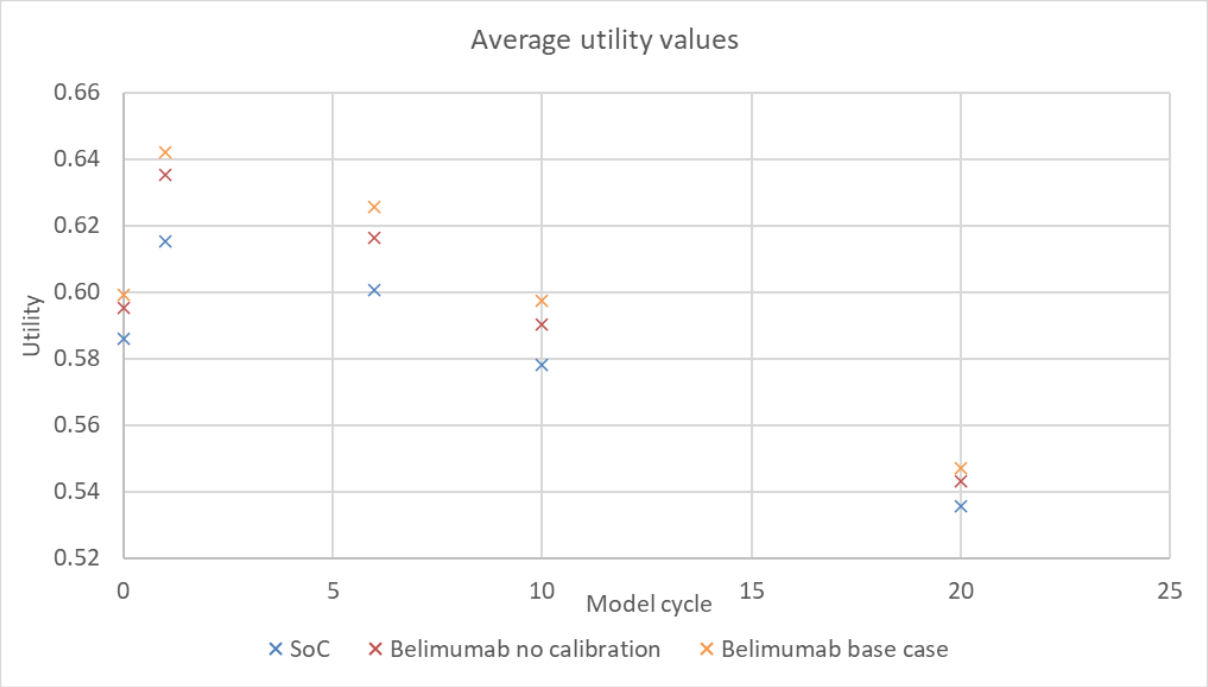


Figure 1. Average utilities by treatment groups over time

in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus (review of TA397) [ID1591]

ADDENDUM: Updated PAS for the IV formulation

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
Sabine Grimm, Health Economist, Maastricht UMC, the Netherlands
Ben Wijnen, Health Economist, Maastricht UMC
Annette Chalker, Systematic Reviewer, KSR Ltd
Sean Harrison, Systematic Reviewer, KSR Ltd
Nigel Armstrong, Health Economist, KSR Ltd
Thomas Otten, Health Economist, Maastricht UMC
Mohammed Islam, Health Economist, KSR Ltd
Gill Worthy, Statistician, KSR Ltd
Lisa Stirk, 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, UK
YO19 6FD

Date completed 10/05/2021

Further evidence regarding response at 24 weeks

We appreciate that the company provided further evidence at short notice.

1) The company state that no non-responders had a SS score reduction of >=4 points at 24 weeks. In fact the score reduction at 24 weeks among non-responders ranged between 0.31 and 0.85, according to the company. It was not explicitly stated whether this referred to the model output or the BLISS trials. The company refers to a regression model explaining this, which we assume is the equivalent of the model in Table 61 of the CS, but this was not provided to the ERG for the 24 weeks time point and hence in the model version available to the ERG it is not possible to replicate the 24 weeks SS scores.

The ERG questions whether a drop of more than 3 points in SS score (difference between >=4 and the range mentioned above) between 24 weeks and 52 weeks can be clinically plausible. In 1d, the company states that this could be explained by switching to alternative treatments – yet in the model (and presumably in NHS practice?), these patients will continue to receive SoC and the only difference is that belimumab treatment is discontinued. The company did not provide evidence to support this claim.

The company's scenario may not be informative in this context: first, it was unclear to the ERG which scenario the company referred to. Second, the impact may be not be sufficiently explored given that the company assumes that belimumab non-responders have identical SS scores to patients in the SoC arm.

Further information would be necessary to assess the plausibility of the company's modelling choices and impact of this issue, at least:

- SS improvement at 24 weeks in BLISS trials, stratified by response and average.
- SS improvement at 52 weeks in BLISS trials, stratified by response and average.
- Explanation as to what exactly in the BLISS trials caused the significant decrease in the SS score between 24 and 52 weeks and whether this is generalizable to the UK setting.
- The same regression results as presented for 52 weeks (Table 61 in CS) but for 24 weeks.
- Model file enabling calculation of 24 weeks SS scores, with output copied to validation sheet (next to 52 week SS scores per individual patient).

2) The utility values may be useful for committee. They do look relatively low (for example compared to those published by Wang et al 2015 in the Nature Scientific Reports https://www.nature.com/articles/srep13297?report=reader).