

# Belimumab: a technological advance for systemic lupus erythematosus patients? Report of a systematic review and meta-analysis

Ngianga-Bakwin Kandala,<sup>1,2</sup> Martin Connock,<sup>1</sup> Amy Grove,<sup>1</sup> Paul Sutcliffe,<sup>1</sup> Syed Mohiuddin,<sup>3</sup> Louise Hartley,<sup>1</sup> Rachel Court,<sup>1</sup> Ewen Cummins,<sup>4</sup> Caroline Gordon,<sup>5</sup> Aileen Clarke<sup>1</sup>

**To cite:** Kandala N-B, Connock M, Grove A, *et al.* Belimumab: a technological advance for systemic lupus erythematosus patients? Report of a systematic review and meta-analysis. *BMJ Open* 2013;**3**:e002852. doi:10.1136/bmjopen-2013-002852

► Prepublication history and additional material for this paper is available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2013-002852>).

Received 2 April 2013  
Revised 19 June 2013  
Accepted 20 June 2013

## ABSTRACT

**Objectives:** To undertake a systematic review and meta-analysis to investigate clinical effectiveness of belimumab for patients with systemic lupus erythematosus (SLE) and antinuclear and/or anti-double-stranded DNA (dsDNA) autoantibodies.

**Methods:** We searched eight electronic databases and reference lists for randomised controlled trials (RCTs) of belimumab against placebo or best supportive care. Quality assessment and random effects meta-analysis were undertaken.

**Design:** A meta-analysis of RCTs.

**Participants:** 2133 SLE patients.

**Primary and secondary outcome measures:** SLE Responder Index (SRI) at week 52.

**Results:** Three double-blind placebo-controlled RCTs (L02, BLISS-52 BLISS-76) investigated 2133 SLE patients. BLISS-52 and BLISS-76 trials recruited patients with antinuclear and/or anti-dsDNA autoantibodies and demonstrated belimumab effectiveness for the SRI at week 52. Ethnicity and geographical location of participants varied considerably between BLISS trials. Although tests for statistical heterogeneity were negative, BLISS-52 results were systematically more favourable for all measured outcomes. Meta-analysis of pooled 52-week SRI BLISS results showed benefit for belimumab (OR 1.63, 95% CI 1.27 to 2.09). By week 76, the primary SRI outcome in BLISS-76 was not statistically significant (OR 1.31, 95% CI 0.919 to 1.855).

## INTRODUCTION

Systemic lupus erythematosus (SLE) is an autoimmune disease subject to relapse and remission. Incidence is estimated between 1.0 and 10.0/100 000 person-years using different measures, and prevalence between 20 and 70/100 000.<sup>1 2</sup> SLE is a complex multiorgan disease with a number of different manifestations.<sup>3</sup> Patients almost always have fatigue,

## ARTICLE SUMMARY

### Article focus

- Systemic lupus erythematosus (SLE) is a complex multiorgan autoimmune disease subject to relapse and remission.
- Patients almost always have fatigue, skin rashes and arthritis and there is a wide variety of other problems which the disease can cause.
- Belimumab is a new treatment specifically targeted against SLE.

### Key messages

- Combining the results from two randomised controlled trials suggests that belimumab is clinically effective for SLE patients.
- However, all outcomes were systematically superior in one trial compared with the other.
- Different trial conduct and populations mean that it is too early to draw generalisable conclusions.

### Strengths and limitations of this study

- At the first sight, combined meta-analytic evidence suggests that belimumab is clinically effective for patients with severe SLE.
- We suggest that it is too early to draw strong conclusions because the two relevant trials cover different populations in different countries and there may be differences in trial conduct and outcome assessment.

often have skin rashes and arthritis and there is a wide variety of other problems which the disease can cause.

The American College of Rheumatology has defined 11 classification criteria, including rash, photosensitivity, oral ulcers, arthritis, serositis and renal and neurological disorder.<sup>4 5</sup> Assessment of the patient can be difficult, as flares of the disease have to be distinguished from its complications, from comorbidity especially infection and from adverse effects of medications.<sup>6</sup> SLE is more

For numbered affiliations see end of article.

### Correspondence to

Dr Ngianga-Bakwin Kandala; N-B.Kandala@warwick.ac.uk

common in women (in most studies 90% or more of cases are women<sup>2</sup>) and in those from black and other ethnic groups. Recently, age-adjusted incidence rates have been produced showing that the rates are highest in women aged 40 years and over.<sup>7</sup> Mortality rates show that 5-year survival is high, at over 90%<sup>8,9</sup> and an overall SMR has been calculated as 2.4.<sup>10</sup>

Antinuclear antibodies are present in virtually all patients with SLE.<sup>11</sup> Anti-dsDNA antibodies are present in 50–60% patients at some point in their disease, but often transiently with active disease.<sup>11</sup> Corticosteroids are the mainstay of treatment; they suppress disease, but they may cause organ damage. The aim of the treatment is to maintain normal function while suppressing disease activity and preventing organ damage,<sup>6</sup> achieving these conflicting aims can be difficult. Other drugs used include antimalarials such as hydroxychloroquine, and immunosuppressive drugs such as azathioprine and mycophenolate mofetil. More recently rituximab (a monoclonal antibody which reacts with the CD20 antigen expressed on B cells) has also been used, although the largest trial undertaken to date failed to reach its end point.<sup>12</sup>

Belimumab (Benlysta) is an IgG1 monoclonal antibody which inhibits the activity of the soluble cytokine BlyS (B lymphocyte stimulator; also known as BAFF).<sup>13</sup> In contrast to earlier SLE treatments, belimumab is targeted at the fundamental pathology of SLE and has been widely interpreted as representing a step change in treatment options.<sup>13</sup>

Belimumab was recently licensed in the USA and in Europe for the treatment of autoantibody-positive SLE and is the first drug to be so licensed for several decades. The European indication is for severely affected SLE patients with active, autoantibody-positive disease and a high degree of disease activity exemplified by positive anti-dsDNA and low complement despite standard therapy.<sup>13</sup> Belimumab is administered by intravenous infusion recommended at 10 mg belimumab/kg on days 0, 14 and 28, and at 28-day intervals thereafter. A course of belimumab treatment for a 64 kg patient using the US list price of \$1477 (£926.37) for a 400 mg vial<sup>14</sup> would be \$56 527 (£35 454)/year, and according to the USA average whole sale price of \$4.432 (£2780)/400 mg vial<sup>15</sup> would be \$42 545 (£26 684)/year.

A number of clinical measures have been developed for tracking the progression of SLE<sup>16</sup> and for estimating the effects of treatment.<sup>17</sup> These include the Physician's Global Assessment (PGA), the Safety of Estrogen in Lupus National Assessment-Systemic Lupus Erythematosus Disease Activity Index (SELENA-SLEDAI), the British Isles Lupus Assessment Group Index (BILAG Index) and the SLE Response Index (SRI). Their major features are summarised in figure 1. Their complexity means that outside specialised centres they may not be widely used in routine clinical practice. The multiplicity of SLE manifestations and of the systems developed to measure them has resulted in a proliferation of outcome measures that can be reported in trials of interventions for SLE. This in turn

means that by chance at least some outcome measures will generate favourable results for an intervention; hence, the US Food and Drug Agency (FDA) in conjunction with belimumab-trialists developed the SRI aimed at guarding against the possibility that worsening in overall disease might be masked by apparent improvement in a more narrowly defined manifestation.

Our objective was to synthesise findings from randomised controlled trials (RCTs) of belimumab for patients with SLE and antinuclear and/or anti-dsDNA autoantibodies to make an overall assessment of the performance of this drug in relation to comparator treatments using the SRI and other outcomes (as listed in figure 1) and to assess the findings of trials in the light of population samples and geographical factors.<sup>18</sup>

## METHODS

The study was undertaken as a part of the work for the National Institute for Health Research, Health Technology Assessment programme (grant funding reference 10/73/01). Further information is available from: <http://www.hta.ac.uk/>.

### Search scope

We searched for RCTs investigating belimumab administered intravenously for patients with SLE and antinuclear and/or anti-dsDNA autoantibodies. Comparators considered were belimumab versus placebo and belimumab versus best supportive care. Outcomes included all disease-related or health-status-related measures. There was no publication year restriction, but the search was restricted to English language references only.

### Search strategy

The following eight databases were searched: Cochrane Database of Systematic Reviews; the Cochrane Central Register of Controlled Trials (CENTRAL); DARE; EMBASE; HTA Database; MEDLINE; and Pre-Medline and Science Citation Index. Search strategies for these databases used a combination of terms related to the population and interventions listed above; the specific search strategies are provided in the online supplementary appendix S1. In MEDLINE and EMBASE, the subject strategies were combined with search strategies designed to identify RCTs (see online supplementary appendix S1).

Unpublished studies were identified using Clinical Trials; Current Controlled Trials; EU Clinical Trials Register; UK Clinical Research Network Study Portfolio; National Research Register; WHO Clinical Trials Search Portal; National Health Service Evidence; Conference Proceedings Citation Index—Science; and Google.

In addition, specific websites were searched: Medicines and Healthcare products Regulatory Agency, European Medicines Agency, US Food and Drug Administration (FDA) and the following specific conference proceedings: American College of Rheumatology, British Society of

**Figure 1** Summary of the major clinical measures used in systemic lupus erythematosus trials.

<p><b>SELENA-SLEDAI:</b> encompasses 24 weighted items scored dichotomously as present or absent in the previous 10 days, thus improvement or worsening of a manifestation is not captured. Overall disease activity is scored over a range of 0 to 105 points. A minimum clinically meaningful score change = a decrease of 6 points (overall improvement) or an increase of 8 points (overall worsening). A designated change in score (<math>\geq 4</math> points) between baseline and follow up can be used to dichotomise patients into responders or non-responders for overall disease.</p> <p><b>BILAG</b><sup>16</sup>: Includes 86 items grouped in 8 organ systems to assesses organ system involvement over the last 4 weeks compared to preceding 4 weeks based on physicians intention to treat using classifications ranging from A to E as follows: A = worsening usually requiring intensification of steroids or immunosuppressant treatments; B = worsening usually requiring antimalarials, nonsteroidal anti-inflammatory drugs (NSAIDs), or low dose steroids; C = stable disease (symptomatic therapy); D = improvement; E = system never involved. Unlike SELENA-SLEDAI it can detect worsening or improvement in individual organ system involvement.</p> <p><b>PGA:</b> employed to monitor change in patient overall disease activity; typically a visual analogue scale is used ranging between no disease = 0, mild disease = 1, moderate disease = 2, and severe disease = 3.</p> <p><b>SRI:</b> A composite instrument (combining elements of SELENA-SLEDAI, BILAG and PGA) developed by belimumab-trialists in conjunction with the US FDA. It allows patients to be dichotomised into responders or non-responders according to predefined assessment criteria in each of the component elements, such as: a SELENA-SLEDAI improvement of <math>\geq 4</math> points, plus no worsening in PGA score by <math>&gt; 0.3</math> points, plus no new BILAG organ system involvement scoring category A in one system or category B in two or more systems. An advantage of SRI, over any one of its components used alone, may be that it can detect SLE improvement in some initial manifestation(s) while guarding against the possibility that worsening in organ systems or overall disease activity might be masked.</p>
--

Rheumatology and the European League Against Rheumatism.

*Inclusion criteria:* Publications were included if they described results from RCTs of belimumab for SLE patients with positive autoantibodies. Two reviewers independently assessed retrieved publications for inclusion. There were no disagreements between reviewers.

*Data extraction:* Potentially relevant publications were obtained in full text and assessed by the same two reviewers. One reviewer extracted data for all specified primary and secondary outcome measures, for adverse events and deaths. A second reviewer checked the extracted data.

*Quality evaluation:* Quality assessment and risk of bias was guided by the Centre for Reviews and Dissemination checklist<sup>19</sup> based on all information in the included publications which specifies reporting of randomisation, concealment of allocation, group balance, blinding, drop-outs, outcome reporting bias and whether intention to treat analysis was used.

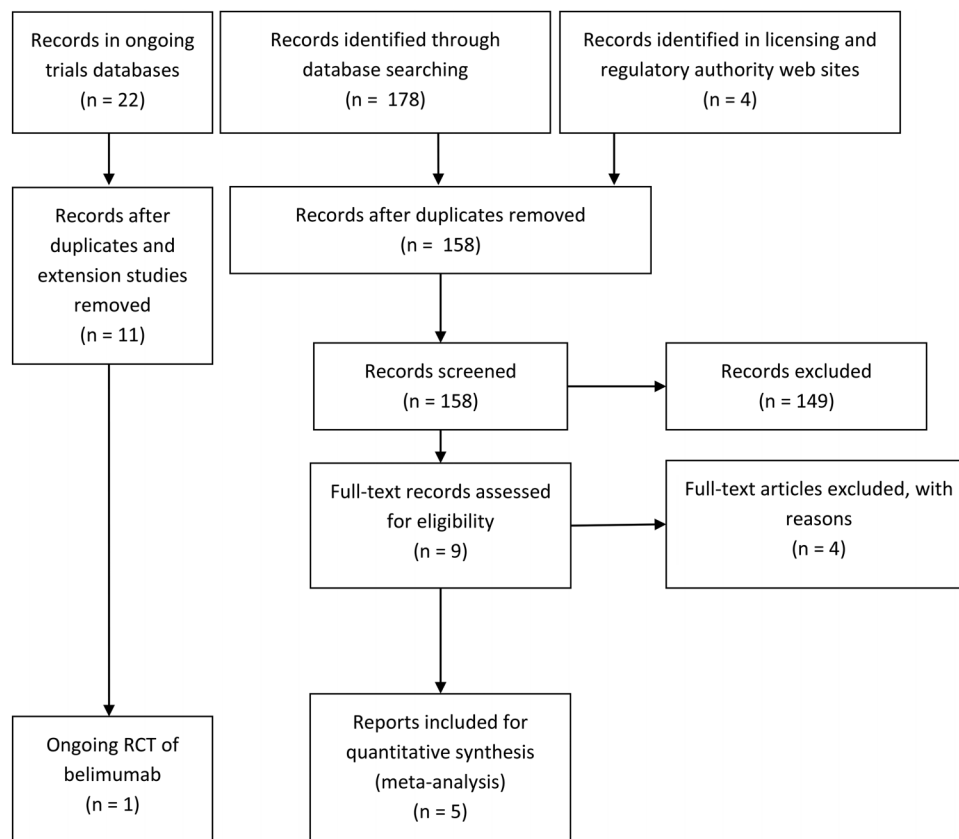
*Statistical analysis:* Unadjusted ORs and mean differences were calculated for binary and continuous outcomes, respectively. Statistical heterogeneity was calculated using

the  $I^2$  statistic.<sup>20 21</sup> There were very few studies for analysing publication bias.<sup>21</sup> Although our thorough search found no further studies, we cannot completely rule out that any method for combining the two trials may result in an overestimate or an underestimate of effect sizes due to publication bias. Adjusted outcome measures were tabulated where these were reported. A random effects meta-analysis<sup>22</sup> was undertaken using the DerSimonian Laird method in STATA V.11.<sup>23</sup> All graphs were prepared in Microsoft Excel 2010.

## RESULTS

### Characteristics of included studies

We identified three placebo-controlled RCTs of belimumab versus standard care: the phase III trials termed BLISS-52<sup>24</sup> and BLISS-76<sup>25</sup> and a phase II trial (study L02).<sup>26</sup> The PRISMA flow chart shows the process of identification of publications (see [figure 2](#)). We identified an on-going trial in Asia.<sup>27</sup> All three completed trials appeared to be of good quality; however, details of allocation concealment were meagre ([table 1](#)). In meta-analysis, we included the two phase III trials



**Figure 2** PRISMA 2009 flow diagram for belimumab in systemic lupus erythematosus randomised controlled trials and on-going trials.

(BLISS-52 and BLISS-76) since the population, trial design and primary outcome were different in the L02 trial.

BLISS-52,<sup>24</sup> BLISS-76<sup>25</sup> and study L02<sup>26</sup> have been published in peer-reviewed journals; however, the fullest accounts in the public domain are in the FDA licensing approval documents<sup>28 29</sup> and the manufacturer’s 2011 submission to the UK National Institute of Health and Clinical Excellence (NICE).<sup>30</sup> Each of these placebo-controlled randomised trials was designed with multiple randomised groups. In the L02 trial, patients received

1 or 4 or 10 mg/kg of belimumab or placebo, while in the BLISS trials the belimumab dose regimens were 1 or 10 mg/kg. Both the US and European licensing is for the 10 mg/kg dose regimen. In this article, we focus on efficacy results for the 10 mg/kg licensed regimen relative to placebo. We also consider the off license 1 and 4 mg/kg dose regimens for the investigation of adverse events.

Centralised, stratified randomisation was used in all three trials and the arms were generally well balanced. For phase III trials, stratification was undertaken

**Table 1** Quality assessment of the included trials

Quality items	L02	BLISS-52	BLISS-76
Does reporting suggest that randomisation was carried out appropriately?	Yes	Yes	Yes
Does reporting suggest that the concealment of treatment allocation adequate?	Unclear	Unclear	Unclear
Were the groups reported as similar at the outset of the study in terms of prognostic factors?	Yes	Yes	Yes
Were the care providers, participants and outcome assessors reported as blind to treatment allocation?	Yes	Yes	Yes
Were any unexpected imbalances in drop-outs reported 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 intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Yes	Yes

Quality assessment used information presented in the study journal articles and the manufacturer’s submission to the US FDA and was based on CRD guidance (2008)<sup>19</sup> for undertaking systematic reviews in healthcare. (CRD, Centre for Reviews and Dissemination; York: Centre for Reviews and Dissemination).



according to race, baseline proteinuria and disease activity score (SELENA-SLEDAI); for the phase I study, disease activity only was used as a stratification factor. All three trials recruited predominantly female patients (~90%) and were described as double blind. The two BLISS studies were conducted according to similar protocols.

There were differences in geographical distribution of the study centres and in the resulting ethnic racial make-up of the study populations (table 2 and figure 3). Thus in BLISS-76, 70% were Caucasian, 13% Native American and 3% Asian, whereas in BLISS-52, 27% were Caucasian, 32% native American and 38% were Asian. Table 3 lists the major protocol prespecified outcomes in the BLISS trials.

There were additional population differences between BLISS and L02 trials at recruitment. Reporting of results for patients with antinuclear and/or anti-dsDNA auto-antibodies in L02 was only included for a post hoc subgroup and primary outcomes measured in L02 were not comparable with those of the BLISS studies. For these reasons, L02 study results are included here only with regard to safety outcomes. For the BLISS trials, a composite novel primary outcome measure was developed a priori from discussions between the FDA and the manufacturer and termed the SRI (see figure 1 and table 3). The protocol prespecified primary end point was the proportion of SRI responders at week 52. This is taken as the primary outcome in this systematic review.

Efficacy results in the two BLISS trials for major binary effectiveness outcomes including the time to first SLE flare and to first severe flare are summarised in figure 4. ORs have been calculated using the proportions of patients with and without events reported in the journal articles for these trials.<sup>24 25</sup> Safety outcomes shown in figure 4 were calculated after combining the number of events across the three trials (L02, BLISS-52 and BLISS-76) and are taken from the FDA documents. The HRs for time to flares were poorly reported in journal articles and the data presented are taken from the manufacturer's submission to the FDA.<sup>28 29</sup> As shown in figure 4, both trials satisfied this primary end point with a better result for BLISS-52. The difference in percentage responders in the belimumab group relative to placebo group was larger in BLISS-52 (14%), than in BLISS-76 (9.4%).

For the other binary effectiveness outcomes, the BLISS-52 trial delivered results which were more favourable to belimumab than did BLISS-76, with the latter results failing to reach a conventional level of statistical significance except for the  $\geq 4$  point improvement in SLEDAI score at week 52. The journal articles and manufacturer's submissions to the FDA and to NICE used a logistic regression model and reported ORs adjusted according to the stratification factors employed at randomisation. Adjusted ORs for a response in BLISS-52 and in BLISS-76 were, respectively, 1.83 (95% CI 1.30 to 2.59;  $p=0.0006$ ) and 1.52 (95% CI 1.07 to 2.15;

$p=0.0207$ ). Again a superior response was found for the BLISS-52 trial. By week 76, the unadjusted OR for the SRI response in the BLISS-76 trial ceased to reach statistical significance (figure 4); this also held for the reported OR adjusted by logistic regression (OR 1.31, 95% CI 0.92 to 1.87,  $p=0.1323$ ).<sup>29</sup>

With regard to time for first flare or for first severe flare (each estimated over 52 weeks follow-up), the responses reported in the FDA submission are again superior for BLISS-52. Each outcome failed to reach conventional statistical significance for BLISS-76. The FDA submission additionally reported more mature results estimated over 76 weeks of follow-up for BLISS-76, and again these indicate a lack of statistical significance for both outcomes (HR for first flare: 1.05, 95% CI 0.88 to 1.27; HR for first severe flare 1.30, 95% CI 0.92 to 1.85).

Figure 4 shows the results for major safety outcomes. Although there were more serious adverse events, more serious infections and more deaths associated with belimumab than with placebo, none of the ORs for these outcomes reached statistical significance. There were 14 deaths during the controlled phase of the three trials; 3 in the placebo group ( $n=675$ ) and 11 in the belimumab groups ( $n=1458$ ) with 6 in the 10 mg/kg and 5 in the 1 mg/kg groups, respectively (OR 11.7; 95% CI 0.474 to 6.124). There were various causes of death: five were infection-related, three were strokes, three cardiovascular events, two suicides, one cancer, one from SLE-related complications and two were of unknown cause.

Results for continuous outcomes are summarised in figure 5. Mean changes from baseline reported in the BLISS journal articles and in the manufacturer's submissions to the FDA and NICE have been used to generate a mean difference statistic (sometimes termed 'weighted mean difference'<sup>31</sup>). These revealed superiority of response in BLISS-52 relative to BLISS-76 for all reported outcomes, a pattern similar to that for binary outcomes. Mean changes from baseline for FACIT-fatigue scores and for EQ-5D utility scores (not pictured) did not reach statistical significance and the improvement seen again in BLISS-52 for these was superior to that seen in BLISS-76.

In summary, BLISS-52 showed a systematic superiority over BLISS-76 in apparent benefit of belimumab across the full range of test responses (binary, time to event and continuous), which may reflect geographical differences between the trials (table 2 and figure 3). The primary outcome in BLISS-76 was achieved (adjusted OR 1.52, 95% CI 1.07 to 2.15), but large geographical differences within BLISS-76 were striking: rates of 32% (46 of 145) and 35% (47 of 136), for placebo and belimumab, respectively, were reported for patients from North America and Canada (a <3% greater response for belimumab), whereas for BLISS-76 patients outside these regions a 15% greater response for belimumab over placebo was reported, 71 of 137 (51.8%) for belimumab and 47 of 130 (36.1%) for placebo. In

**Table 2** Major characteristics of included studies

Study	Treatment (intravenous)	N	Mean age (SD) years	SELENA-SLEDAI at entry	Geographical distribution of patients	Ethnic make-up of trial participants	Number and location of Study Centres	
L02	Bel 1 mg/kg	114	42 (11)	>4 points	USA (98%), Canada (2%)	Caucasian	NR 69.9%	59 in N America
2006 Phase II	Bel 4 mg/kg	111				African American	NR 24.7%	
52 week	Bel 10 mg/kg	111				Latino	NR 18.5%	
	Placebo	113						
BLISS-52 2009	Bel 1 mg/kg	288	36 (11)	>6 points	Latin America (50%), Asia (38%), E Europe and Australia (13%)	Caucasian	229 27%	90 in Pacific Asia. 11 in S America and E Europe
Phase III 52 week	Bel 10 mg/kg	290				Asian	327 38%	
	Placebo	287				Black/African Am	30 4%	
						Alaskan Nat/Am	279 32%	
						Indian		
						Nat Hawaiian/Pacific Islander	0 0%	
						Multiracial	5 1%	
BLISS-76 2009	Bel 1 mg/kg	271	40 (12)	>6 points	USA and Canada (53%), W Europe (25%), E Europe (11%), Latin America (11%)	Caucasian	569 70%	136 in N America and Europe
Phase III 76 week	Bel 10 mg/kg	273				Asian	28 3%	
	Placebo	275				Black/African Am	118 14%	
						Alaskan Nat/Am	103 13%	
						Indian		
						Nat Hawaiian/Pacific Islander	1 0%	
						Multiracial	8 1%	

NR, not reported; SELENA-SLEDAI, Safety of Estrogen in Lupus National Assessment-Systemic Lupus Erythematosus Disease Activity Index.



Figure 3 Differing centre locations in the BLISS 52 and BLISS 76 multicentre trials.

comparison, the corresponding rates for patients from Latin America in BLISS-52 were 49% placebo (71 of 145), and 61% belimumab (85 of 140).

The manufacturer’s submissions to the FDA and to NICE combined results from the two BLISS trials by pooling the patients and applying the logistic regression

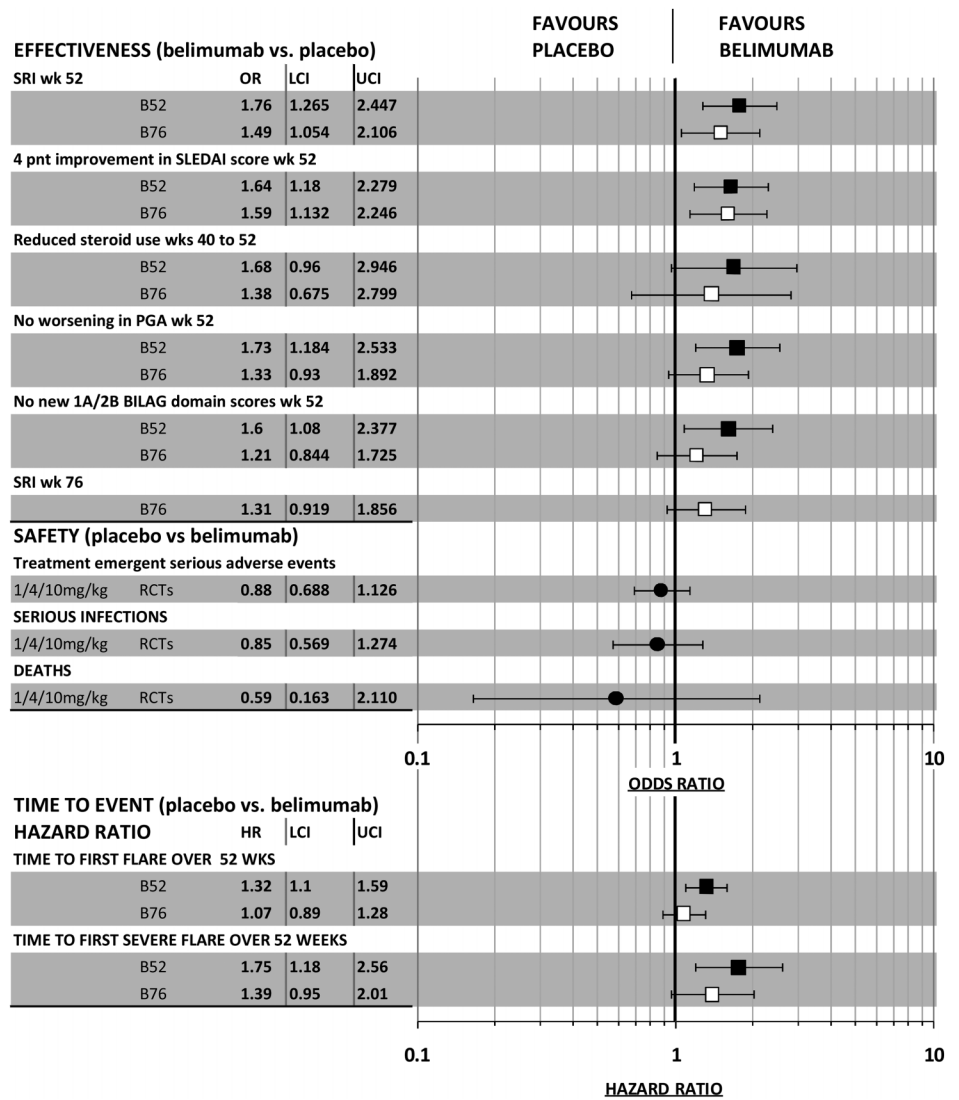
Table 3 Outcomes defined and prespecified in the BLISS 52 and BLISS 76 trials and their accompanying designated status

Outcome	Measure	Outcome specification
SLE Responder Index (SRI*)	Percentage responders at week 52	Primary outcome
Reduction in SLEDAI score by ≥4 points	Percentage responders at week 52	Major secondary outcome
<i>Change in PGA score from baseline</i>	<i>Mean change at week 24</i>	<i>Major secondary outcome</i>
Steroid reduction weeks 40–52	Percentage responders	Major secondary outcome
<i>SF-36 physical component summary score</i>	<i>Mean change at week 24</i>	<i>Major secondary outcome</i>
SLE Responder Index	Percentage responders at week 76	Major secondary outcome
<i>SLICC/ACR damage index</i>	<i>Mean change at week 52</i>	<i>Secondary outcome</i>
<i>FACIT-fatigue scale mean change from baseline</i>	<i>Mean change at clinic visits</i>	<i>Secondary outcome</i>
<i>EQ-5D score</i>	<i>Mean change at clinic visits</i>	<i>Secondary outcome</i>
<i>Change in PGA score from baseline</i>	<i>Mean change at week 52</i>	<i>Secondary outcome</i>
<i>SF-36 physical component summary score</i>	<i>Mean change at week 52</i>	<i>Secondary outcome</i>
SLEDAI SLE flare index over 52 weeks	Time to first flare	Secondary outcome
SLE Responder Index (SRI)	Percentage responders at timed clinic visits	Other outcome reported
Modified SLE responder index	Percentage responders at week 52	Other outcome reported
No worsening in PGA score by ≥0.3	Percentage responders at week 52	Other outcome reported
No new BILAG 1A/2B domain scores	Percentage responders at week 52	Other outcome reported
<i>Change in SLEDAI score from baseline</i>	<i>Mean change at week 52</i>	<i>Other outcome reported</i>
Death	Number during exposure	Safety outcome
Treatment emergent adverse events	Number during exposure	Safety outcome
Serious infections	Number during exposure	Safety outcome

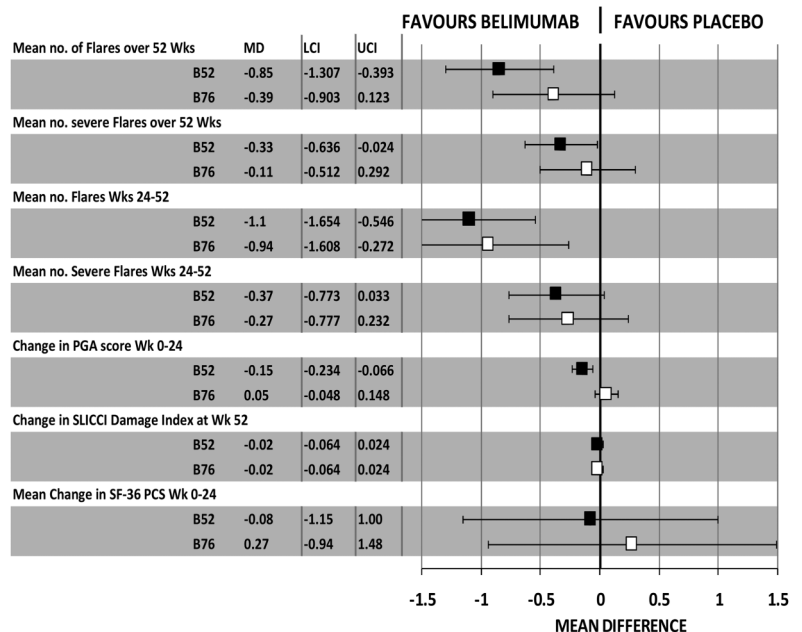
\*Composite outcome measure consisting of ≥4 points improvement in SLEDAI score, no worsening in PGA by ≥0.3 points and no new BILAG 1A or 2B domain scores. Continuous outcomes are in italics.

ACR, American College of Rheumatology; BILAG, British Isles Lupus Assessment Group; EQ-5D, EuroQoL 5 dimensions; FACIT, Functional Assessment of Chronic Illness Therapy; SF-36, Short Form 36-Item Health Survey; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index; SLICC, Systemic Lupus International Collaborating Clinics.

**Figure 4** Summary of results for major binary and time to event outcomes in belimumab randomised controlled trials.



**Figure 5** Summary of results for major continuous outcomes in BLISS 52 and BLISS 76 trials.





model described above; for the primary outcome (proportion of SRI responders at week 52), the difference between the belimumab and placebo groups was 11.8%.<sup>28</sup>

An alternate method of combining trials by the meta-analysis of study level results from the two BLISS trials showed a statistically significant benefit of belimumab for most main outcomes including SRI, SELENA-SLEDAI, worsening in PGA, steroid use, BILAG score, or time to first severe flare, and mean number of flares and severe flares over 52 weeks and weeks 24–52 (figure 6). Tests for statistical heterogeneity of ORs and HR outcomes were not significant. This meta-analysis offers an alternative to the manufacturer logistic regression and it is justified for two trials of substantial size (N=577 and N=548); however, these results, and those from pooling individual patient data from the two trials prior to logistic regression, mask the systematic difference between trials across all outcomes.

DISCUSSION

We undertook a systematic review of the clinical effectiveness of belimumab, a new treatment targeted at SLE patients with antinuclear and/or anti-dsDNA autoantibodies. We performed an extensive search and systematic review of both completed and on-going trials using a number of databases and by checking reference lists. Data were extracted independently and studies were quality assessed. Random effects meta-analysis was undertaken.

We identified three RCTs (L02, BLISS-52, BLISS-76) reporting data on over 2000 patients. In contrast to the BLISS trials, L02 recruited patients who were not necessarily current carriers of antinuclear or anti-dsDNA antibodies at the start of the study. L02 failed to demonstrate the clinical effectiveness for its primary end points.<sup>26</sup> Meta-analysis of the BLISS studies showed a benefit of belimumab with the main primary outcome

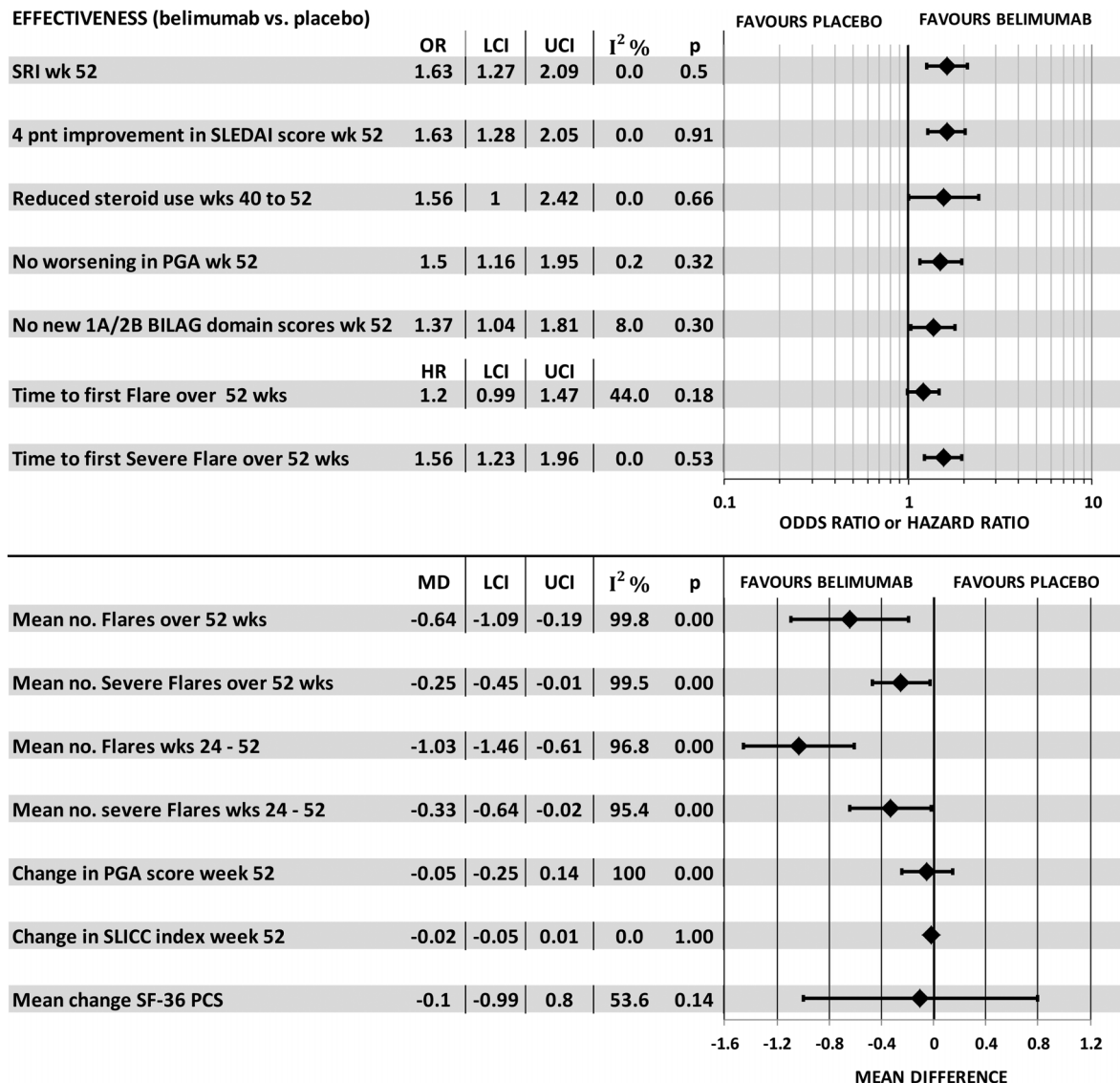


Figure 6 Meta-analysis of major outcomes in the two BLISS trials.

(SRI), showing improvement at 52 weeks (OR 1.63; 95% CI 1.27 to 2.09  $p < 0.001$ ), although by week 76 the proportion of SRI responders in the BLISS-76 trial ceased to reach statistical significance (OR 1.31; 95% CI 0.92 to 1.87  $p = 0.1323$ ). There were no significant differences between placebo and intervention groups in the quality of life or adverse events.

We found that the benefits of belimumab were systematically greater across the board (although not significantly so) in the BLISS-52 trial and although tests for statistical heterogeneity were negative, geographical location of study centres and the racial background and ethnicity of participants varied considerably. If the two BLISS trials were drawn from the same underlying populations, while one might expect outcomes to differ, we would anticipate that this would occur randomly between trials—some better and some worse than the other.

Very few studies have directly assessed the existence and the importance of geographical differences in trial outcomes.<sup>32–34</sup> Key factors contributing to such differences are variations in the underlying patient population characteristics and variation in study execution. Vickers *et al*<sup>33</sup> found that Eastern Asian and Eastern European studies had a higher proportion of positive trial results when compared with other countries. This is seen in the present case for the primary outcome where both the belimumab and placebo response rates in BLISS 52 study were greater than those in BLISS-76 and, remarkably, the placebo response rate in BLISS-52 (49%) was greater than that for the belimumab arm of BLISS-76 (43%). O’Shea and DeMets<sup>35</sup> also report that within the  $\beta$ -Blocker Heart Attack Trial, there was a difference not only in the direction but also in the size of the treatment effect between Canada and the USA, although it should be noted that the original aim of that trial was not the investigation of international differences in treatment effect. One study found that 96–99% of the total variance in the “Global utilisation of strategies to open occluded coronary arteries IV acute coronary syndromes” (GUSTO IV ACS) trial could be accounted for by patient-level factors.<sup>36</sup>

International trials need to harmonise the training of investigators, patient selection, treatment management, thresholds to centre admission, access to facilities, ascertainment of endpoints and, by implication, results of interest<sup>37–44</sup> since it is possible that in centres in different countries these factors may differ systematically.<sup>37</sup> Equally, the underlying differences in populations and countries (ethnicity, genetics, socioeconomic status and healthcare systems) and the nature and epidemiology of SLE according to ethnic background may result in differences in reporting the outcomes and pooled results.

The outcomes used in the BLISS trials would be unfamiliar to most of the investigators and it is possible that the criteria may have differed between countries. In particular, PGA is an important element of the outcomes measured (see figure 1). PGA was measured as an outcome in itself, and it is also incorporated in SRI.

PGA is of concern because as a global physician assessment of a patient’s SLE status it is subjective. At week 24 the advantage of belimumab over placebo in the percentage change in PGA score from baseline was considerably larger, by about 10%, for BLISS 52 than for BLISS 76; this larger result in one trial is likely to have had an important influence on the findings for relative effectiveness in the two trials.

The latest results of belimumab in patients with SLE (phase II study design, uncontrolled extension study) reported that of 449 patients with active SLE (USA/Canada) 177 (39.4%) patients remained on treatment after 7 years of therapy (ie, approximately 1746 cumulative patient-years) and that this subgroup exhibited durable sustained improvement in SLE disease activity (SRI and PGA).<sup>30</sup>

## CONCLUSIONS

In conclusion, systematic review and random effects meta-analysis of two RCTs of belimumab for patients with autoantibody-positive SLE demonstrated positive results in the main outcome at week 52. However, in view of the different populations studied at different locations in the BLISS trials and the consistently superior results from one trial compared with the other, we consider that population heterogeneity, geographical differences and variation in trial conduct and outcome assessment may have played a role in influencing the outcomes. However, the generalisability of results pooled meta-analytically or by logistic regression should be viewed with caution and we suggest that it is too early to draw strong conclusions in this case.

### Author affiliations

<sup>1</sup>Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry, UK

<sup>2</sup>KEMRI-University of Oxford-Wellcome Trust Collaborative Programme, Malaria Public Health and Epidemiology Group, Centre for Geographic Medicine, Nairobi, Kenya

<sup>3</sup>Health Sciences, The University of Manchester, Manchester, UK

<sup>4</sup>McMDC Ltd, McMaster Development Consultants, Glasgow, UK

<sup>5</sup>School of Immunity and Infection, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK

**Acknowledgements** The authors would like to thank the National Institute for Health Research, Health Technology Assessment programme for funding this work.

**Contributors** N-BK participated in conception and design, data analysis and interpretation, drafting the article, critical revisions for important intellectual content and the approval of the final article for submission. MC participated in conception and design, data analysis and interpretation, literature review, interpretation of results, drafting the article, critical revisions for important intellectual content and in the final approval of final article for submission. AG participated in the interpretation of results and critical revisions for important intellectual content. PS participated in literature review, interpretation of results and critical revisions for important intellectual content. SM participated in data analysis and interpretation, interpretation of results and critical revisions for important intellectual content. LH participated in literature review, interpretation of results and critical revisions for important intellectual content. RC participated in literature review and critical revisions for important

intellectual content. EC participated in interpretation of results and critical revisions for important intellectual content. CG participated in interpretation of results and critical revisions for important intellectual content. AC participated in conception and design, interpretation of results, drafting the article, critical revisions for important intellectual content and in the approval of the final article for submission. All authors read and approved the final manuscript.

**Funding** This work was supported by the National Institute for Health Research, Health Technology Assessment programme (project number 10/73/01).

**Competing interests** None.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** No additional data are available.

**Open Access** This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>

## REFERENCES

- Bernatsky S, Joseph L, Pineau CA, *et al.* A population-based assessment of systemic lupus erythematosus incidence and prevalence—results and implications of using administrative data for epidemiological studies. *Rheumatology (Oxford)* 2007;46:1814–18.
- Pons-Estel GJ, Alarcon GS, Scofield L, *et al.* Understanding the epidemiology and progression of systemic lupus erythematosus. *Semin Arthritis Rheum* 2010;39:257–68.
- Gordon C, Li CK, Isenberg DA. Systemic lupus erythematosus. *Medicine* 2010;38:73–80.
- Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1997;40:1725.
- Tan EM, Cohen AS, Fries JF, *et al.* The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1982;25:1271–7.
- Griffiths B, Mosca M, Gordon C. Assessment of patients with systemic lupus erythematosus and the use of lupus disease activity indices. *Best Pract Res Clin Rheumatol* 2005;19:685–708.
- Somers EC, Thomas SL, Smeeth L, *et al.* Incidence of systemic lupus erythematosus in the United Kingdom, 1990–1999. *Arthritis Rheum* 2007;57:612–18.
- Pons-Estel GJ, Alarcon GS, McGwin G Jr, *et al.* Protective effect of hydroxychloroquine on renal damage in patients with lupus nephritis: LXV, data from a multiethnic US cohort. *Arthritis Rheum* 2009;61:830–9.
- Pons-Estel GJ, Gonzalez LA, Zhang J, *et al.* Predictors of cardiovascular damage in patients with systemic lupus erythematosus: data from LUMINA (LXVIII), a multiethnic US cohort. *Rheumatology (Oxford)* 2009;48:817–22.
- Bernatsky S, Boivin JF, Joseph L, *et al.* Mortality in systemic lupus erythematosus. *Arthritis Rheum* 2006;54:2550–7.
- Kurien BT, Scofield RH. Autoantibody determination in the diagnosis of systemic lupus erythematosus. *Scand J Immunol* 2006;64:227–35.
- 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 Rheum* 2010;62:222–33.
- Summary of product characteristics. 2012. [ema.europa.eu/FindMedicine/HumanMedicines/EuropeanPublicAssessmentReports](http://ema.europa.eu/FindMedicine/HumanMedicines/EuropeanPublicAssessmentReports) (accessed Jan 2012).
- American College of Rheumatology. Belimumab for Systemic Lupus Erythematosus. 2011. [http://www.rheumatology.org/publications/hotline/2011\\_03\\_15\\_belimumab.asp](http://www.rheumatology.org/publications/hotline/2011_03_15_belimumab.asp) (accessed Jan 2012).
- New Drug Bulletin. Belimumab (Benlysta—Human Genome Sciences, Inc). 2012. [http://healthcare.utah.edu/pharmacy/bulletins/NDB\\_228.pdf](http://healthcare.utah.edu/pharmacy/bulletins/NDB_228.pdf) (accessed Jan 2012).
- Isenberg DA, Rahman A, Allen E, *et al.* BILAG 2004. Development and initial validation of an updated version of the British Isles Lupus Assessment Group's disease activity index for patients with systemic lupus erythematosus. *Rheumatology* 2005;44:902–6.
- Gordon C, Bertias G, Ioannidis JPA, *et al.* EULAR points to consider for conducting clinical trials in Systemic lupus erythematosus. *Ann Rheum Dis* 2009;68:470–6.
- Burgos PI, McGwin G Jr, Pons-Estel GJ, *et al.* US patients of Hispanic and African ancestry develop lupus nephritis early in the disease course: data from LUMINA, a multiethnic US cohort (LUMINA LXXIV). *Ann Rheum Dis* 2011;70:393–4.
- Khan KS, Ter Riet G, Glanville J, *et al.* Reviews of Research on Effectiveness. CRD's Guidance for Carrying Out or Commissioning Reviews. 2nd Edition. (ISBN 1900640201) CRD Report No. 4. York: NHS Centre for Reviews and Dissemination (CRD), University of York, 2001. <http://www.york.ac.uk/inst/crd/report4.htm> (accessed Jan 2012).
- Higgins JP, Thompson SG, Deeks JJ, *et al.* Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.
- Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. 2012. <http://www.cochrane-handbook.org> (accessed Jan 2012).
- DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177–88.
- STATA Statistical analysis and software. 2011. <http://www.stata.com/> (accessed Apr 2012).
- Navarra SV, Guzman RM, Gallacher AE, *et al.* Efficacy and safety of belimumab in patients with active systemic lupus erythematosus: a randomised, placebo-controlled, phase 3 trial. *Lancet* 2011;377:721–31.
- Furie R, Petri M, Zamani O, *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;63:3918–30.
- Wallace DJ, Stohl W, Furie RA, *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;61:1168–78.
- GlaxoSmithKline. GSK1550188 A 52 week study of belimumab versus placebo in the treatment of subjects with SLE located in North East Asia. 2011. <http://clinicaltrials.gov/ct2/show/NCT01345253> (accessed Jan 2012).
- U.S. Food and Drugs Administration. Arthritis Advisory Committee Meeting Briefing Document for the 16 November 2010 Meeting: Belimumab: BENLYSTA (belimumab): Treatment of Systemic Lupus Erythematosus: BLA 125370. 2010. <http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisDrugsAdvisoryCommittee/UCM233581.pdf> (accessed Jan 2012).
- U.S. Food and Drugs Administration. FDA briefing package. 2010. <http://www.fda.gov/downloads/advisorycommittees/committeesmeetingmaterials/drugs/arthritisdrugsadvisorycommittee/ucm233579.pdf> (accessed Jan 2012).
- GlaxoSmithKline. Single Technology Appraisal (STA): Belimumab for the treatment of active autoantibody-positive systemic lupus erythematosus. 2011. <http://www.guidance.nice.org.uk/TA/live/13307/56490/56490.pdf> (accessed Jan 2012).
- Egger M, Smith GD, Altman DG. *Systematic reviews in health care, meta-analysis in context*. 2nd edn. BMJ Books, 2001.
- Pan Z, Trikalinos TA, Kavvoura FK, *et al.* Local literature bias in genetic epidemiology: an empirical evaluation of the Chinese literature. *PLoS Med* 2005;2:e334.
- Vickers A, Goyal N, Harland R, *et al.* Do certain countries produce only positive results? A systematic review of controlled trials. *Control Clin Trials* 1998;19:159–66.
- Zhang D, Freemantle N, Cheng KK. Are randomized trials conducted in China or India biased? A comparative empirical analysis. *J Clin Epidemiol* 2011;64:90–5.
- O'Shea JC, DeMets DL. Statistical issues relating to international differences in clinical trials. *Am Heart J* 2001;142:21–8.
- Chang WC, Midodzi WK, Westerhout CM, *et al.* Are international differences in the outcomes of acute coronary syndromes apparent or real? A multilevel analysis. *J Epidemiol Community Health* 2005;59:427–33.
- White HD. International differences: selection, noise, or real? *Eur Heart J* 2000;21:339–42.
- Fox KA, Goodman S, Bigonzi F, *et al.* Inter-regional differences and outcome in unstable angina; analysis of the international ESSENCE trial. Efficacy and Safety of Subcutaneous Enoxaparin in Non-Q-wave Coronary Events. *Eur Heart J* 2000;21:1433–9.
- Sinnaeve PR, Van de Werf FJ. Global patterns of health care for acute coronary syndromes. *Curr Opin Cardiol* 2004;19:625–30.

40. Faergeman O, Kjekshus J, Cook T, *et al.* Differences in the treatment of coronary heart disease between countries as revealed in the Scandinavian Simvastatin Survival Study (4S). *Eur Heart J* 1998;19:1531–7.
41. Berger JS, Stebbins A, Granger CB, *et al.* Initial aspirin dose and outcome among ST-elevation myocardial infarction patients treated with fibrinolytic therapy. *Circulation* 2008;117:192–9.
42. Welsh RC, Chang W, Goldstein P, *et al.* Time to treatment and the impact of a physician on prehospital management of acute ST elevation myocardial infarction: insights from the ASSENT-3 PLUS trial. *Heart* 2005;91:1400–6.
43. Yusuf S, Flather M, Pogue J, *et al.* Variations between countries in invasive cardiac procedures and outcomes in patients with suspected unstable angina or myocardial infarction without initial ST elevation. OASIS (Organisation to Assess Strategies for Ischaemic Syndromes) Registry Investigators. *Lancet* 1998;352:507–14.
44. Bleyer AJ, Hylander B, Sudo H, *et al.* An international study of patient compliance with hemodialysis. *JAMA* 1999;281:1211–13.