

## Original article

# Causal cascade of direct and indirect effects of anifrolumab on patient-reported outcomes: structural equation modelling of two Phase 3 trials

Donald Stull<sup>1</sup>, Sean O'Quinn<sup>2</sup>, Betsy Williams<sup>1</sup>, Stephanie Bean<sup>1</sup>, Erik Schwetje<sup>3</sup>, Gabriel Abreu<sup>4</sup> and Raj Tummala <sup>3</sup>

## Abstract

**Objectives.** SLE significantly impairs health-related quality of life (HRQoL). In this *post hoc* analysis, structural equation modelling was used to examine the 'causal cascade' of interaction between anifrolumab, disease activity and patient-reported outcomes (PROs) in pooled data from the phase 3 TULIP-1 and TULIP-2 trials.

**Methods.** Data were pooled from the TULIP-1 ( $n=364$ ) and TULIP-2 ( $n=362$ ) randomized, placebo-controlled, 52-week trials of intravenous anifrolumab (300 mg every 4 weeks for 48 weeks). We evaluated changes from baseline to week 24 and week 52 in four clinical (BICLA, BILAG-2004, SLEDAI-2K and changes in glucocorticoid dosage) and six PRO measures (SF-36, FACIT-F, EQ-5D, LupusQoL, PHQ-8 and pain NRS) in our hypothesized model of interactions.

**Results.** Our hypothesized model had an acceptable fit to the pooled TULIP trial data. At week 24, significant paths revealed that when compared with placebo, anifrolumab treatment improved disease activity as measured by BICLA, BILAG-2004, SLEDAI-2K and changes to glucocorticoid dosage. In turn, these clinical measures reduced pain, which improved fatigue, physical functioning, mood/emotions and HRQoL. When the model incorporated number of glucocorticoid tapers as the measure of change in glucocorticoid dosage, treatment effects of anifrolumab on glucocorticoid tapers were not retained at week 52. However, at week 52 treatment indirectly improved HRQoL through its direct effects on BICLA.

**Conclusions.** Anifrolumab is associated with significant patient-reported improvements in aspects of HRQoL including pain, fatigue, mood and physical function. These benefits are from the direct effect of anifrolumab treatment on disease activity and reduction in glucocorticoid dosage.

**Key words:** SLE and autoimmunity, quality of life, biological therapies, clinical trials and methods, depression

### Rheumatology key messages

- We evaluated the impact of anifrolumab treatment on patient-reported outcomes (PROs) with structural equation modelling.
- Improvements in BICLA, BILAG-2004 and SLEDAI-2K and reduced glucocorticoid dosage lessened pain and improved PROs.
- Anifrolumab treatment positively impacts overall health-related quality of life for patients with SLE.

<sup>1</sup>Patient Centered Endpoints, IQVIA, New York City, NY, <sup>2</sup>BioPharmaceuticals Medical, <sup>3</sup>BioPharmaceuticals R&D, AstraZeneca, Gaithersburg, MD, USA and <sup>4</sup>BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden

Submitted 16 July 2021; accepted 25 February 2022

Correspondence to: Raj Tummala, Global Clinical Head, Anifrolumab, BioPharmaceuticals R&D, 1 MedImmune Way, Gaithersburg, MD 20878, USA. E-mail: raj.tummala@astrazeneca.com

## Introduction

SLE is a complex, chronic and heterogeneous autoimmune disease that can affect any organ system, and patients with SLE present with a variety of clinical manifestations [1]. These clinical manifestations are devastating for most patients and can lead to reduced physical function, loss of employment, a major impact on health-related quality of life (HRQoL), frequent hospitalizations, cumulative and irreversible organ damage, and early mortality [2–6]. Organ damage accumulates from SLE disease activity itself and the treatment-related adverse effects associated with chronic use of glucocorticoids and other immunosuppressive agents [7, 8]. There remains a substantial unmet medical need for novel treatments in SLE with disease-specific mechanisms of action that can reduce overall disease activity and concomitant use of steroids and other non-specific immunosuppressive agents, while also reducing flares, comorbidity and long-term organ damage [9].

Multiple lines of evidence indicate a role of type I IFNs in the pathogenesis of SLE and other autoimmune diseases [10–12]. Anifrolumab is a human immunoglobulin (Ig) G1 kappa monoclonal antibody that inhibits type I IFN signalling and blocks the biologic activity of type I IFNs. The efficacy of anifrolumab in patients with moderate to severe SLE was evaluated in two phase 3 randomized, placebo-controlled, double-blind, 52-week trials, TULIP-1 (NCT02446912) and TULIP-2 (NCT02446899) [13, 14]. Improvement in favour of anifrolumab was observed on several measures of disease activity and reductions in glucocorticoid dosage were reported.

Most SLE clinical studies evaluating new therapeutic options include a composite end point comprising clinical and laboratory measures of global and organ-specific disease and clinician-graded severity assessments as a primary end point. Although improvement in these clinician-rated outcomes provides evidence to support the therapeutic value of an investigational drug, they do not directly capture how patients feel and function. As such, patient-reported outcomes (PROs) are an important complement to composite measures of disease activity and damage when evaluating the efficacy of SLE treatments. The importance of PROs in SLE clinical trials is highlighted in the consensus recommendations from the OMERACT group which recommends that studies include generic and disease-specific HRQoL as well as measures of symptom severity [15]. The inclusion of these PRO measures in the TULIP-1 and TULIP-2 studies offers the opportunity to evaluate how improvements in clinical measures, including disease activity and oral glucocorticoid dosage reduction, relate to the SLE patient experience.

Causal cascades of effects of symptoms on functioning and HRQoL have been presented in multiple indications, including non-small-cell lung cancer [16], metastatic non-small-cell lung cancer and breast cancer [17], chemotherapy-induced anaemia [18], multidimensional

physical health measures, and use of long-term care services [19]. Based on this, the present study hypothesized a ‘causal cascade’ of effects such that treatment with anifrolumab would result in improvements in clinical assessments of disease activity, reduced use of glucocorticoids, and improved HRQoL. In this hypothesized ‘causal cascade’ it is important to consider the possible mechanism of action and the intervening variables in the causal chain between treatment and patient function and treatment. Thus, while no significant *direct* relationship may exist between treatment and HRQoL, there may be a significant *indirect* relationship between treatment and HRQoL through these intervening variables.

The goal of the present analysis was to examine a causal cascade (i.e. direct and indirect relationships) among clinical assessments, symptoms, patient functioning, and HRQoL for patients with SLE treated with anifrolumab vs placebo in the TULIP trials.

## Methods

This was a *post hoc* analysis of pooled data from the phase 3 randomized, placebo-controlled, double-blind, 52-week TULIP-1 and TULIP-2 trials [13, 14]. Patients were aged 18–70 years and fulfilled the American College of Rheumatology revised classification criteria for SLE [20]. Patients were randomized to receive intravenous infusions of placebo or anifrolumab (300 mg) every 4 weeks for 48 weeks in addition to standard therapy. For patients receiving oral glucocorticoids  $\geq 10$  mg/day (prednisone or equivalent) at baseline, a protocol-mandated attempt to taper to  $\leq 7.5$  mg/day was required between Weeks 8 and 40; tapering was also permitted for patients receiving oral glucocorticoids  $< 10$  mg/day at baseline. Stable oral glucocorticoid dosage was required for all patients between Weeks 40 and 52.

## Outcomes

We evaluated four clinical and six PRO measures. The clinical measures included the British Isles Lupus Assessment Group-based Composite Lupus Assessment (BICLA) [21, 22], BILAG-2004 [23], SLEDAI 2000 (SLEDAI-2K) [24, 25], and changes to glucocorticoid dosage measured in two different ways: (1) number of glucocorticoid dosage tapers and (2) percentage glucocorticoid dosage change from baseline. The PRO measures included the Short Form 36 Health Survey (SF-36) version 2 (acute) [26], Functional Assessment of Chronic Illness Therapy–Fatigue (FACIT-F), EuroQol 5 Dimension 5 Level (EQ-5D-5L) and visual analogue scale (EQ-5D VAS) [27–29], Lupus Quality of Life (LupusQoL) [30], eight-item Patient Health Questionnaire depression scale (PHQ-8) [31, 32] and a pain Numerical Rating Scale (NRS). As a sensitivity analysis, in place of the PRO pain measures, analyses included a count of swollen and tender joints. The purpose of including these variables in place of PRO measures was to assess the effects of anifrolumab on pain related to joints and the

**TABLE 1** Concepts and their respective measures for the planned analyses

Concepts	Measure(s)
Pain	Pain NRS, SF-36 Bodily Pain, LupusQoL Pain
Fatigue	FACIT-F, SF-36 Vitality, LupusQoL Fatigue
Physical functioning	SF-36 Physical Functioning, SF-36 Role Physical, LupusQoL Physical Health
Mood/emotions	SF-36 Mental Health, PHQ-8, LupusQoL Emotional Health
HRQoL	EQ-5D-5L, EQ-5D VAS

Note: For the sensitivity analyses using count of swollen and tender joints, the Pain concept was measured with two individual indicators: count of tender joints, and count of swollen joints at Weeks 24 and 52. EQ-5D-5L: EuroQol 5 Dimension 5 Level; EQ-5D VAS: EuroQol 5 Dimensions Visual Analogue Scale; FACIT-F: Functional Assessment of Chronic Illness Therapy–Fatigue; HRQoL: health-related quality of life; LupusQoL: Lupus Quality of Life; NRS: Numerical Rating Scale; PHQ-8: eight-item Patient Health Questionnaire Depression Scale; PRO: patient-reported outcomes; SF-36: Short Form 36 Health Survey.

effects of swollen and tender joints on patient functioning and HRQoL. Details of each of the clinical and PRO measurements are provided in the [Supplementary Materials \(Supplementary Data S1, available at \*Rheumatology\* online\)](#).

#### Patient-reported concepts

The PROs were further divided into five key patient-reported concepts of HRQoL (pain, fatigue, physical functioning, mood/emotions and HRQoL) and were measured using a multiple indicator modelling approach where multiple PROs and PRO domains measured each underlying concept (Table 1). Of the eight domains captured by SF-36, physical functioning, role physical, bodily pain, vitality and mental health were used in the key concepts. The LupusQoL also captures eight domains, of which, physical health, pain, fatigue and emotional health were used in the key concepts.

#### Analyses

Structural equation modelling (SEM) path analysis was used to evaluate the direct and indirect effects of the clinical and PRO variables [33, 34]. Analyses were first conducted in each trial separately and included only patients in the anifrolumab 300 mg and placebo arms. TULIP-1 included an anifrolumab 150 mg group that was not included in this analysis. If results were similar in each trial, then the two trials would be pooled. Analyses were conducted using data at baseline and Weeks 24 and 52, because these three time points included all PROs of interest. All SEM analyses were conducted using Mplus version 8.4 (Muthen & Muthen, Los Angeles, CA, USA).

#### Hypothesized model

Fig. 1 presents the hypothesized model that was the basis of the present analyses. After baseline, treatment was hypothesized as having direct effects on the clinical measures of BICLA responses and BILAG-2004 and SLEDAI-2K global scores, and was anticipated to result in reductions in glucocorticoid dosage. In addition, we anticipated a correlation (the double-headed arrow) between clinical

measures and glucocorticoid reduction, such that if treatment results in a BICLA response and an improved SLEDAI-2K/BILAG-2004 score, this should also be associated with a reduction in glucocorticoid dosage.

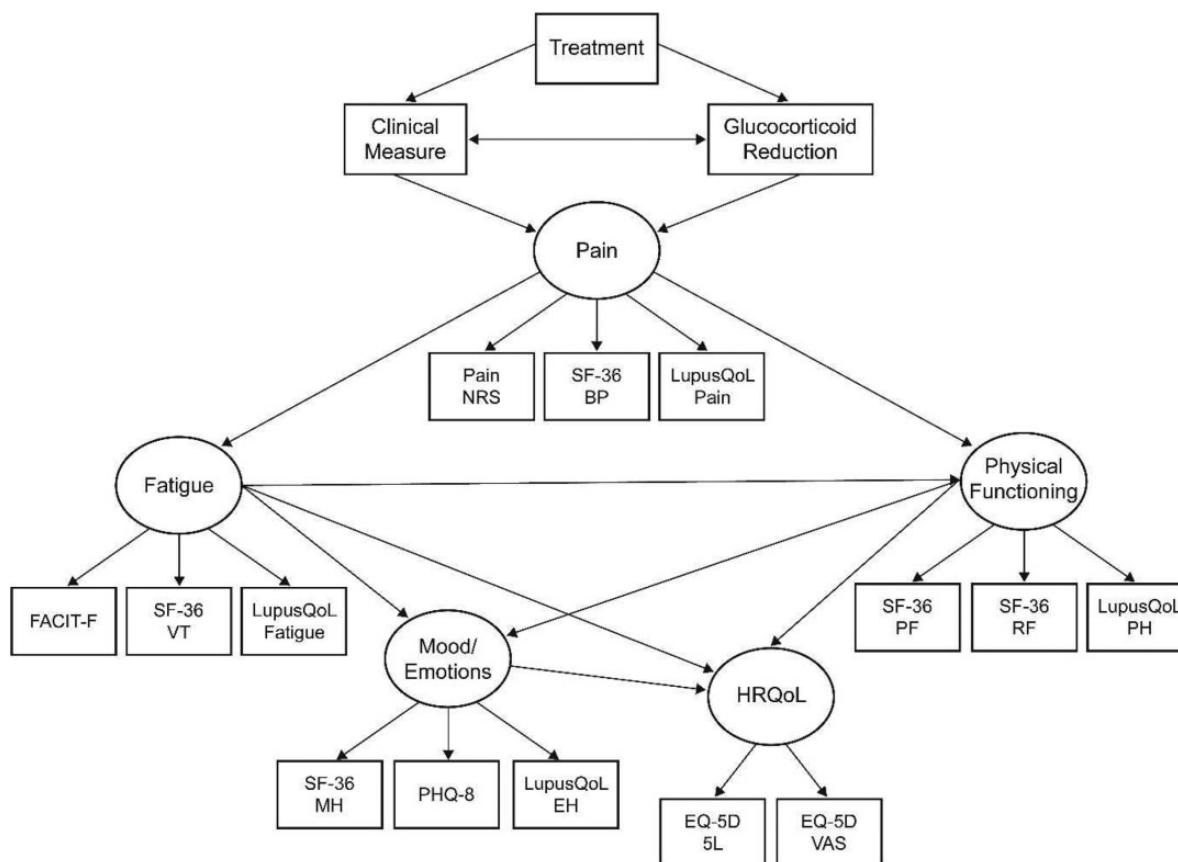
In this model, the clinical measures and glucocorticoid dosage reduction were expected to directly affect pain. That is, if treatment worked as anticipated, it would result in improved clinical scores and glucocorticoid dosage reductions, which would lead to reduced patient-reported pain. With reductions in patient-reported pain, there would be reductions in fatigue and improved physical function. With less fatigue and improved physical functioning, we expected to see improvement in mood and ultimately improved HRQoL (Fig. 1). Likewise, for the sensitivity analyses using counts of swollen and tender joints, treatment with anifrolumab should result in fewer swollen and tender joints, which would lead to reductions in fatigue, improved physical functioning, better mood and overall improved HRQoL.

This model was applied to post-baseline visits at week 24 and 52 to examine the differential effects of treatment. No baseline model or analysis is presented because no treatment effect is observed at baseline. Four tests of goodness-of-fit (model  $\chi^2$  with its degrees of freedom and *P*-value, comparative fit index [CFI] [35], root mean square error of approximation [RMSEA] [36], and standardized root mean square residual [SRMR] [34]) were used to evaluate the correspondence of the hypothesized model to the observed data.

The TULIP trials were conducted in accordance with the principles of the Declaration of Helsinki and the International Conference on Harmonization Guidelines for Good Clinical Practice. As this was a *post hoc* analysis of anonymised data, no ethics committee or institutional review board approvals were required. All such approvals were obtained in the original trials [13, 14].

## Results

In total, 726 patients were included in the model, 364 from TULIP-1 and 362 from TULIP-2. Of these 726 patients, 366 received placebo (184 in TULIP-1, 182 in TULIP-2) and 360

**Fig. 1** Hypothesized structural equation path model of treatment, clinical measures and PROs

BP: bodily pain; BILAG: British Isles Lupus Assessment Group; EQ-5D-5L: EuroQol 5 Dimension 5 Level; EQ-5D VAS: EuroQol 5 Dimensions visual analogue scale; FACIT-F: Functional Assessment of Chronic Illness Therapy–Fatigue; HRQoL: health-related quality of life; LupusQoL: Lupus Quality of Life; MH: mental health; EH: emotional health; NRS: Numerical Rating Scale; PHQ-8: 8-item Patient Health Questionnaire depression scale; PF: physical functioning; PH: physical health; PRO: patient-reported outcomes; RF: role physical; SF-36: Short Form 36 Health Survey; Tx: treatment; VT: vitality.

received anifrolumab 300 mg (180 patients in each trial). Descriptions of trial participants have been presented for TULIP-1 and TULIP-2 [13, 14]. Pooled patient demographics and clinical characteristics were generally balanced across treatment groups (anifrolumab 300 mg and placebo) in both individual trials and pooled patient data (Supplementary Table S1, available at *Rheumatology* online) [13, 14]. In the anifrolumab and placebo groups at baseline, 80.8% and 83.1% of patients, respectively, were receiving glucocorticoids, and 52.8% and 50.5% were receiving glucocorticoids  $\geq 10$  mg/day.

The hypothesized model had an acceptable fit to the data for each trial separately, was consistent across both trials and time points, and was of acceptable size according to CFI, RMSEA and SRMR estimates (CFI ranged from 0.931–0.953; RMSEA ranged from 0.065–0.087; SRMR ranged from 0.036–0.049). In addition, the relationships among the constructs were consistent across trials and time points (data not shown). Therefore, the TULIP trials were pooled for these analyses. For the pooled analyses,

the fit statistics were of acceptable size for the models using data at week 24 and week 52 that included BICLA response and the number of glucocorticoid tapers, and BICLA response and percentage change in glucocorticoid dosage (Table 2). The RMSEA at week 24 was slightly elevated (preferred value  $\leq 0.08$ ) relative to the CFI and SRMR, both of which indicated good fit (Table 2).

Table 3 presents the model-generated mean (s.d.) PRO scores, number of glucocorticoid tapers, and percent change from baseline in glucocorticoid dosage, at baseline (PROs only) and Weeks 24 and 52. The greatest differences in PRO scores occurred between baseline and week 24, with only a small incremental change between Weeks 24 and 52 for PROs.

#### Relationship of PROs with BICLA response and number of glucocorticoid dosage tapers

Fig. 2A presents the model for week 24 pooled data that included BICLA response and number of glucocorticoid dosage tapers between Weeks 8 and 24 as the clinical

TABLE 2 Goodness-of-fit statistics

Fit statistic:	BICLA response and count of glucocorticoid tapers		BICLA response and percentage reduction in glucocorticoid dosage	
	week 24	week 52	week 24	week 52
$\chi^2$ (df)	613.245 (103)***	453.319 (103)***	616.229 (103)***	446.752 (103)***
CFI <sup>a</sup>	0.936	0.949	0.936	0.951
RMSEA <sup>b</sup>	0.083	0.068	0.083	0.068
90% CI for RMSEA	0.076–0.089	0.062–0.075	0.077–0.089	0.061–0.074
SRMR <sup>c</sup>	0.039	0.035	0.039	0.034

For more complex models, the CFI may be smaller and the RMSEA may be larger than for more parsimonious models and with larger sample sizes. Goodness-of-fit statistics reported for pooled TULIP-1 and TULIP-2 data. <sup>a</sup>CFI generally preferred value  $\geq 0.9$  or 0.95. <sup>b</sup>RMSEA generally preferred value  $\leq 0.08$ . <sup>c</sup>SRMR generally preferred value  $\leq 0.10$ . \*\*\* $P \leq 0.001$ . BICLA: BILAG-based Composite Lupus Assessment; CFI: comparative fit index; df: degrees of freedom; RMSEA: root mean square error of approximation; SRMR: standardized root mean square residual.

TABLE 3 Model-generated mean scores for PROs and glucocorticoid dosage at baseline and weeks 24 and 52

<i>n</i> = 726	Baseline	Week 24	Week 52
Pain, mean (s.d.)			
Pain NRS	5.5 (2.4)	4.4 (2.6)	4.2 (2.6)
SF-36 bodily pain	43.9 (22.0)	54.3 (24.5)	55.1 (24.0)
LupusQoL pain	52.0 (28.8)	64.6 (27.4)	65.8 (26.1)
Fatigue, mean (s.d.)			
FACIT-F	25.8 (12.1)	30.8 (12.9)	31.2 (12.8)
SF-36 vitality	38.1 (21.2)	46.2 (23.1)	46.6 (23.3)
LupusQoL fatigue	48.3 (26.5)	58.0 (27.1)	58.2 (27.1)
Physical functioning, mean (s.d.)			
SF-36 physical functioning	51.5 (26.6)	61.8 (26.0)	63.4 (25.9)
SF-36 role physical	44.8 (25.9)	55.5 (26.5)	57.0 (26.0)
LupusQoL physical health	53.8 (25.7)	63.3 (25.2)	64.0 (24.6)
Mood/Emotions, mean (s.d.)			
SF-36 mental health	62.4 (21.0)	65.0 (21.6)	66.4 (21.4)
PHQ-8	9.7 (6.2)	7.9 (6.0)	7.8 (6.0)
LupusQoL emotional health	65.9 (24.8)	72.1 (23.6)	73.2 (23.6)
HRQoL, mean (s.d.)			
EQ-5D-5L	0.6 (0.2)	0.7 (0.2)	0.7 (0.2)
EQ-5D VAS	55.6 (20.8)	63.8 (21.0)	64.7 (21.3)
Glucocorticoid tapers, <i>n</i> , mean (s.d.)			
Glucocorticoid dosage, mg/day, change from baseline, %		1.1 (1.5)	2.0 (2.2)
		-20.0 (44.2)	-22.6 (52.1)

EQ-5D VAS: EuroQol 5 Dimensions visual analogue scale; EQ-5D-5L: EuroQol 5 Dimension 5 Level; FACIT-F: Functional Assessment of Chronic Illness Therapy-Fatigue; HRQoL: health-related quality of life; LupusQoL: Lupus Quality of Life; NRS: Numerical Rating Scale; PHQ-8: 8-item Patient Health Questionnaire depression scale; PRO: patient-reported outcomes; SF-36: Short Form 36 Health Survey.

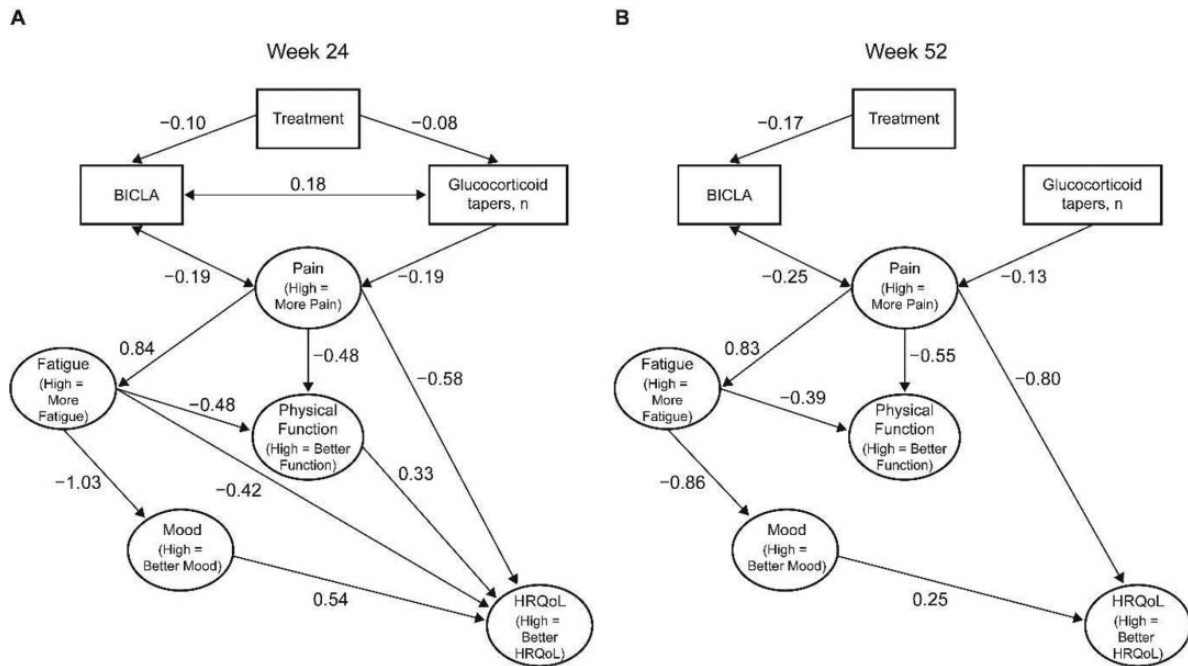
measures. Only paths that are significant ( $P < 0.05$ ) are shown. Betas larger than 1.0 are mathematically possible and not indicative of something wrong with the model [37]. The model shows that patients in the anifrolumab arm were more likely to be BICLA responders and to have more glucocorticoid dosage tapers during this period. Both BICLA response and a higher number of

glucocorticoid dosage tapers resulted in less patient-reported pain.

With a reduction in pain, there was a reduction in fatigue ( $\beta = 0.84$ ), an increase in physical functioning ( $\beta = -0.48$ ), and an improvement in HRQoL ( $\beta = -0.58$ ). When patient-reported fatigue was lessened, there was a corresponding improvement in physical functioning

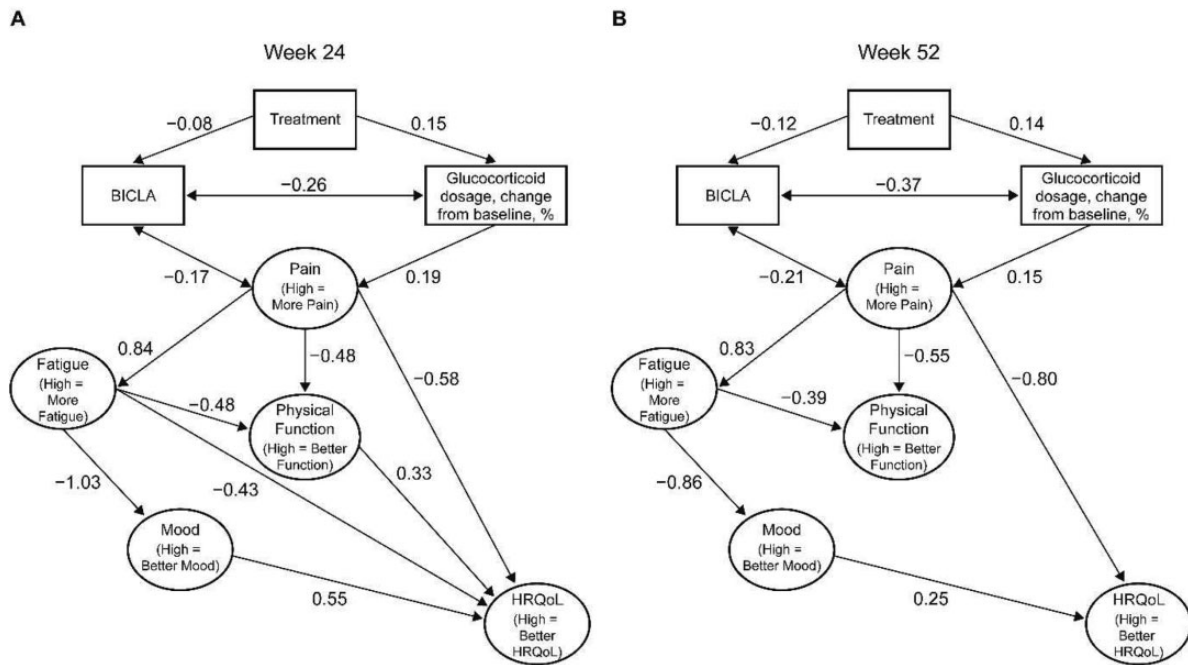


Fig. 2 BICLA and number of glucocorticoid tapers (A) week 24 and (B) week 52



BICLA: BILAG-based Composite Lupus Assessment; HRQoL: health-related quality of life. Models of the results from pooled TULIP-1 and TULIP-2 data. Only paths significant at  $P \leq 0.05$  are shown. Coefficients are equivalent to standardized regression coefficients (Betas).

Fig. 3 BICLA and percentage change from baseline in glucocorticoid dosage (A) week 24 and (B) week 52



BICLA: BILAG-based Composite Lupus Assessment; HRQoL: health-related quality of life. Models of the results from pooled TULIP-1 and TULIP-2 data. Only paths significant at  $P \leq 0.05$  are shown. Coefficients are equivalent to standardized regression coefficients (Betas).

( $\beta = -0.48$ ), an improvement in patient mood ( $\beta = -1.03$ ), and an improvement in HRQoL ( $\beta = -0.42$ ). When patient physical functioning was improved, there was an improvement in HRQoL ( $\beta = 0.33$ ). Finally, when mood was improved, there was an improvement in HRQoL ( $\beta = 0.54$ ) (Fig. 2A).

At week 52, some paths did not retain the significance observed at week 24, including between treatment and number of glucocorticoid tapers or between BICLA and number of glucocorticoid tapers (Fig. 2B). In addition, relationships between HRQoL and physical functioning and fatigue were no longer significant. Of the remaining significant paths, the model shows that patients treated with anifrolumab were more likely than patients in the placebo arm to be BICLA responders, had less pain, less fatigue, better physical functioning and better mood and HRQoL, similar to week 24.

#### Relationship of PROs with BICLA response and percentage change in glucocorticoid dosage

Significant paths observed in our model for week 24 pooled data that included BICLA response and percentage change in glucocorticoid dosage show that patients in the anifrolumab arm were more likely to be BICLA responders and to have a greater decrease in glucocorticoid dosage from baseline to week 24 (Fig. 3A). In addition, the correlation between BICLA response and percentage change from baseline in glucocorticoid dosage indicates that patients who were BICLA responders also had a greater decrease in glucocorticoid dosage during this period. Both BICLA response and a greater decrease in glucocorticoid dosage resulted in less patient-reported pain.

With a reduction in pain, there was a reduction in fatigue ( $\beta = 0.84$ ), an increase in physical functioning ( $\beta = -0.48$ ) and an improvement in HRQoL ( $\beta = -0.58$ ). When patient-reported fatigue was lessened, there was a corresponding improvement in physical functioning ( $\beta = -0.48$ ), an improvement in patient mood ( $\beta = -1.03$ ) and an improvement in HRQoL ( $\beta = -0.43$ ). When patient physical functioning was improved, there was an improvement in HRQoL ( $\beta = 0.33$ ). Finally, when mood was improved, there was an improvement in HRQoL ( $\beta = 0.55$ ) (Fig. 3A).

In contrast to the model including BICLA response and the number of glucocorticoid dosage tapers, at week 52 there were significant relationships between treatment and percentage change from baseline in glucocorticoid dosage and between BICLA response and percentage change from baseline in glucocorticoid dosage (Fig. 3B). However, relationships between HRQoL and physical functioning and fatigue were no longer significant.

Despite these differences in the significance of the model paths, at week 52 the model shows that patients treated with anifrolumab are more likely than patients in the placebo arm to be BICLA responders, have greater decreases in glucocorticoid dosage, have less pain, less

fatigue, better physical functioning, and better mood and HRQoL (Fig. 3B).

#### Relationship of PROs with BILAG-2004 and SLEDAI-2K

We also evaluated the clinical measures BILAG-2004 and SLEDAI-2K in our hypothesized model at week 24 and week 52 for both number of glucocorticoid tapers and percentage reduction in glucocorticoid dosage (Supplementary Figs S1–S4, available at *Rheumatology* online). Results from the model were nearly identical to those presented for BICLA (Figs 2 and 3), including significant indirect relationships between clinical response and less pain, less fatigue, better physical functioning, and better mood and HRQoL in patients treated with anifrolumab compared with patients in the placebo group (Supplementary Figs S1–S4, available at *Rheumatology* online).

#### Sensitivity analyses: swollen and tender joint counts in place of PRO for pain

When the pain PRO concept was replaced with a multiple indicator factor (swollen and tender joint counts) in the model for BICLA response and percentage change in glucocorticoid dosage, the sizes of some path coefficients changed but the results were virtually identical to the analyses including the pain PRO measure (Supplementary Fig. S5, available at *Rheumatology* online). Significant indirect relationships between treatment and clinical response resulted in fewer swollen and tender joints, less fatigue, better physical functioning, better mood and improved HRQoL in patients treated with anifrolumab compared with patients in the placebo group.

## Discussion

In this analysis, we aimed to understand the relationship of treatment and disease activity with HRQoL and symptoms in patients with moderate to severe SLE enrolled in the TULIP-1 and TULIP-2 trials. The evaluation of our hypothesized model consistently showed a causal cascade of effects between treatment with anifrolumab, clinical measures of disease activity and multiple domains of HRQoL. Anifrolumab treatment did not have a direct effect on patient-reported pain or patient-reported fatigue, physical functioning, mood or HRQoL. Instead, treatment with anifrolumab, through its effect on the type 1 IFN pathway, indirectly affected these variables by way of its effect on clinical measures of disease activity (BICLA, BILAG-2004, SLEDAI-2K and changes in glucocorticoid dosage).

The TULIP trials included the generic and specific HRQoL measures SF-36 and LupusQoL, and additional HRQoL measures to capture the full complement of patient health status including fatigue (FACIT-F), pain (Pain NRS) and depression (PHQ-8). These measures were incorporated into the five key patient-reported concepts in our model to evaluate patient-reported efficacy of anifrolumab. Our findings that anifrolumab treatment

improves both clinician-rated disease activity and PROs highlights the value of multiple measures of treatment efficacy as recommended by OMERACT.

Measuring direct treatment benefit with PROs in patients with SLE in a clinical trial setting is difficult and complex. Both poor correlation between PROs and disease activity indices [38–43] and discordance between patient and physician assessments of disease activity are well known in SLE [44, 45]. Multiple factors including background medications, glucocorticoid use, flares and comorbidity can confound treatment response and significantly impact HRQoL [46–48]. Patients in the TULIP trials were receiving standard therapy including combinations of glucocorticoids, immunosuppressants and antimalarials, while also attempting to taper glucocorticoids. By week 52, patients had a mean of ~2 attempts to taper glucocorticoids. Additionally, 33% treated with anifrolumab and 43% of patients in the placebo group had one or more flares during the trial [49]. Thus, these factors are likely to influence a direct assessment of anifrolumab's treatment effect on PROs. These analyses attempted to bridge the connections between treatment and PROs with hypothesized models of direct and indirect treatment effects. The significant pathways observed at week 24 and week 52 showed that patients treated with anifrolumab were more likely than patients in the placebo group to be BICLA responders and have a greater decrease in oral glucocorticoid dosage, have less pain, less fatigue, better physical functioning, and better mood and improved HRQoL. In addition, we observed improvements in the number of swollen and tender joints that were also related to less fatigue, better physical functioning, mood and overall HRQoL. These results suggest that treatment with anifrolumab is effective at reducing patient-reported symptoms, swollen and tender joints, and improving HRQoL, though these effects are modelled indirectly.

We observed a loss of some significant pathways at week 52 compared with week 24 in our analysis of BICLA response and number of glucocorticoid tapers. Most significant was the loss of the pathway between treatment and glucocorticoid taper. This loss was likely driven by a smaller mean number of tapers between Weeks 24 and 52 than between baseline and week 24, the period in which most of the effects occurred. Fewer tapers from week 24 to week 52 for both the anifrolumab and placebo groups likely attenuated the effects observable in the model, suggesting that outcomes with little variability limit the ability to detect a significant correlation. These findings also suggest that the greatest impact of anifrolumab treatment on the use of glucocorticoids is during the first 24 weeks. Examining reductions in oral glucocorticoids in terms of percentage change from baseline yielded somewhat more robust results at week 52. The significant pathway between treatment and percentage change in glucocorticoid dosage presented at week 24 was preserved at week 52 and supports examining glucocorticoid dosage changes in multiple ways.

Several factors may have limited the robustness of some variables in the model. The variable for the percentage change from baseline in glucocorticoid dosage was

calculated using dosing information collected only at baseline and Weeks 24 and 52, and did not account for dosage adjustments between these time points. Similarly, the number of glucocorticoid dosage tapers did not account for increases in dosage. Although neither of these approaches measuring changes in glucocorticoid use can accurately characterize the glucocorticoid burden, both yielded results that were comparable and in line with expectations from the mechanism of action of anifrolumab, which lends confidence that the results are real and meaningful. Fibromyalgia is an important consideration for patients with SLE, as it may significantly worsen pain, fatigue and HRQoL. Few patients in the TULIP trials had a diagnosis of fibromyalgia [50], so we did not consider this in our analyses. The clinical importance of improvements in patient-reported concepts cannot be determined in the context of other PRO studies, as minimal clinically important differences were not considered in this analysis, only changes to scores. Nonetheless, the strength of the significant pathways suggests that improvements in PROs are an important outcome of treatment with anifrolumab. Finally, no formal, empirical test of model invariance was conducted to check whether the hypothesized models fit the same for each trial. However, the size of parameter estimates and comparable fit statistics, and the consistency of results across the two trials gave us confidence that the models were similar for both trials, and pooling patient data was justified.

To conclude, PRO outcomes in TULIP trials assessed by the hypothesized model in these analyses show that anifrolumab is associated with significant reported improvements in aspects of HRQoL, including pain, fatigue, mood and physical function. Nearly identical results were obtained when replacing patient-reported pain with counts of swollen and tender joints. The benefits to patients are from the direct effect of anifrolumab treatment on disease activity and reduction in glucocorticoid dosage. These analyses support the utility of considering the effects of treatment on proximal clinical measures and more distal measures of patient-reported symptoms and HRQoL when evaluating treatment benefit in terms of the patient experience in SLE.

## Acknowledgements

Medical writing support was provided by Rebecca S. Jones, PhD, of JK Associates Inc., part of Fishawack Health.

*Role of the study sponsor:* The study was sponsored by AstraZeneca. All authors interpreted the data, critically reviewed the manuscript for important intellectual content, approved the final draft, and agreed to its submission.

*Funding:* This work was supported by AstraZeneca.

*Disclosure statement:* D.S., B.W. and S.B. are employees of IQVIA and were paid by AstraZeneca to conduct this research study. S.O., E.S., G.A. and R.T. are employees of AstraZeneca.



## Data availability statement

Data underlying the findings described in this manuscript may be obtained in accordance with AstraZeneca's data sharing policy described at <https://astrazenecagrouptrials.pharmacm.com/ST/Submission/Disclosure>.

## Supplementary data

Supplementary data are available at *Rheumatology* online.

## References

- Kaul A, Gordon C, Crow MK *et al*. Systemic lupus erythematosus. *Nat Rev Dis Primers* 2016;2:39.
- Alarcon GS, McGwin G Jr, Uribe A *et al*. Systemic lupus erythematosus in a multiethnic lupus cohort (LUMINA). XVII. Predictors of self-reported health-related quality of life early in the disease course. *Arthritis Rheum* 2004;51:465–74.
- Nived O, Jonsen A, Bengtsson AA, Bengtsson C, Sturfelt G. High predictive value of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology damage index for survival in systemic lupus erythematosus. *J Rheumatol* 2002;29:1398–400.
- Drenkard C, Bao G, Dennis G *et al*. Burden of systemic lupus erythematosus on employment and work productivity: data from a large cohort in the southeastern United States. *Arthritis Care Res* 2014;66:878–87.
- Holloway L, Humphrey L, Heron L *et al*. Patient-reported outcome measures for systemic lupus erythematosus clinical trials: a review of content validity, face validity and psychometric performance. *Health Qual Life Outcomes* 2014;12:116.
- Jorge AM, Lu N, Zhang Y, Rai SK, Choi HK. Unchanging premature mortality trends in systemic lupus erythematosus: a general population-based study (1999–2014). *Rheumatology* 2018;57:337–44.
- Al Sawah S, Zhang X, Zhu B *et al*. Effect of corticosteroid use by dose on the risk of developing organ damage over time in systemic lupus erythematosus—the Hopkins Lupus Cohort. *Lupus Sci Med* 2015;2:e000066.
- Thamer M, Hernan MA, Zhang Y, Cotter D, Petri M. Prednisone, lupus activity, and permanent organ damage. *J Rheumatol* 2009;36:560–4.
- van Vollenhoven RF, Mosca M, Bertsias G *et al*. Treat-to-target in systemic lupus erythematosus: recommendations from an international task force. *Ann Rheum Dis* 2014;73:958–67.
- Kirou KA, Lee C, George S *et al*. Activation of the interferon-alpha pathway identifies a subgroup of systemic lupus erythematosus patients with distinct serologic features and active disease. *Arthritis Rheum* 2005;52:1491–503.
- Lichtman EI, Helfgott SM, Krieger MA. Emerging therapies for systemic lupus erythematosus—focus on targeting interferon-alpha. *Clin Immunol* 2012;143:210–21.
- Psarras A, Emery P, Vital EM. Type I interferon-mediated autoimmune diseases: pathogenesis, diagnosis and targeted therapy. *Rheumatology* 2017;56:1662–75.
- Furie R, Morand E, Bruce I *et al*. Type I interferon inhibitor anifrolumab in active systemic lupus erythematosus (TULIP-1): a randomised, controlled, phase 3 trial. *Lancet Rheumatol* 2019;1:E208–19.
- Morand EF, Furie R, Tanaka Y *et al*. Trial of anifrolumab in active systemic lupus erythematosus. *N Engl J Med* 2020;382:211–21.
- Strand V, Gladman D, Isenberg D *et al*. Endpoints: consensus recommendations from OMERACT IV. Outcome Measures in Rheumatology. *Lupus* 2000;9:322–7.
- Stull DE, Houghton K, Ainsworth C *et al*. “Causal cascade” among outcomes in non-small cell lung cancer: assessing the direct and indirect effects of symptoms on health-related quality of life (HRQL) outcomes. *Value Health J* 2016;19:A748.
- Stull DE, Houghton K, Ainsworth C, Price G, Boye ME. The effects of disease and treatment-associated cancer symptoms on health-related quality-of-life: the mediating effect of fatigue in non-small cell lung cancer and metastatic breast cancer. Poster presented at the 2017 ISPOR 22nd annual international meeting; 22 May 2017. Boston, MA: Value Health, 2017;20:A121.
- Stull D, Vernon M, Legg J *et al*. Use of latent growth curve models for assessing the effects of darbepoetin alfa on hemoglobin and fatigue. *Contemp Clin Trials* 2010;31:172–9.
- Stull D, Kercher K, Kosloski K. Physical health and long-term care: a multidimensional approach. *Am Behav Sci* 1996;39:317–35.
- Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1997;40:1725.
- Murphy CL, Yee CS, Gordon C, Isenberg D. From BILAG to BILAG-based combined lupus assessment-30 years on. *Rheumatology* 2016;55:1357–63.
- Wallace D, Strand V, Furie R *et al*. Evaluation of treatment success in systemic lupus erythematosus clinical trials: development of the British Isles Lupus Assessment Group-Based Composite Lupus Assessment endpoint. Presented as a poster at *American College of Rheumatology congress (ACR)*, 2011. Poster number 2265.
- Yee CS, Farewell V, Isenberg DA *et al*. British Isles Lupus Assessment Group 2004 index is valid for assessment of disease activity in systemic lupus erythematosus. *Arthritis Rheum* 2007;56:4113–9.
- Bombardier C, Gladman DD, Urowitz MB, Caron D, Chang CH. Derivation of the SLEDAI. A disease activity index for lupus patients. The Committee on Prognosis Studies in SLE. *Arthritis Rheum* 1992;35:630–40.
- Gladman DD, Ibanez D, Urowitz MB. Systemic lupus erythematosus disease activity index 2000. *J Rheumatol* 2002;29:288–91.
- Ware J, Kosinski M, Bjorner J, Turner-Bowker D, Gandek B, Maruish M. User's manual for the 36v2 health survey. Lincoln, RI: QualityMetric Incorporated; 2007.
- Hurst NP, Jobanputra P, Hunter M *et al*. Validity of Euroqol—a generic health status instrument—in patients with rheumatoid arthritis. *Economic and Health*

- Outcomes Research Group. *Br J Rheumatol* 1994;33:655–62.
- 28 Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). *Br J Rheumatol* 1997;36:551–9.
- 29 EuroQol. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy* 1990;16:199–208.
- 30 McElhone K, Abbott J, Shelmerdine J *et al.* Development and validation of a disease-specific health-related quality of life measure, the LupusQol, for adults with systemic lupus erythematosus. *Arthritis Rheum* 2007;57:972–9.
- 31 Kroenke K, Spitzer RL. The PHQ-9: a new depression diagnostic and severity measure. *Psychiatric Annals* 2002;32:509–15.
- 32 American Psychiatric Association. *Diagnostic and statistical manual of mental disorders: DSM-IV*. 4th ed. Washington (DC). 1994.
- 33 Tomarken AJ, Waller NG. Structural equation modeling: strengths, limitations, and misconceptions. *Annu Rev Clin Psychol* 2005;1:31–65.
- 34 Kline R. *Methodology in the social sciences. Principles and practice of structural equation modeling*. 4th ed. New York, NY: The Guilford Press. 2016.
- 35 Bentler PM. Comparative fit indexes in structural models. *Psychol Bull* 1990;107:238–46.
- 36 Steiger JH. Structural model evaluation and modification: an interval estimation approach. *Multivariate Behav Res* 1990;25:173–80.
- 37 Jöreskog KG. How large can a standardized coefficient be? 1999. <http://www.statmodel.com/download/Joreskog.pdf> (22 April 2021, date last accessed).
- 38 Alarcon GS, McGwin G Jr, Brooks K *et al.* Systemic lupus erythematosus in three ethnic groups. XI. Sources of discrepancy in perception of disease activity: a comparison of physician and patient visual analog scale scores. *Arthritis Rheum* 2002;47:408–13.
- 39 Annapureddy N, Devilliers H, Jolly M. Patient-reported outcomes in lupus clinical trials with biologics. *Lupus* 2016;25:1111–21.
- 40 Dua AB, Touma Z, Toloza S, Jolly M. Top 10 recent developments in health-related quality of life in patients with systemic lupus erythematosus. *Curr Rheumatol Rep* 2013;15:380.
- 41 Giangreco D, Devilliers H, Annapureddy N, Block JA, Jolly M. Lupus Impact Tracker is responsive to physician and patient assessed changes in systemic lupus erythematosus. *Lupus* 2015;24:1486–91.
- 42 Gladman DD, Urowitz MB, Ong A, Gough J, MacKinnon A. A comparison of five health status instruments in patients with systemic lupus erythematosus (SLE). *Lupus* 1996;5:190–5.
- 43 Merrill J, Buyon J, Furie R *et al.* Assessment of flares in lupus patients enrolled in a phase II/III study of rituximab (EXPLORER). *Lupus* 2011;20:709–16.
- 44 Neville C, Clarke AE, Joseph L *et al.* Learning from discordance in patient and physician global assessments of systemic lupus erythematosus disease activity. *J Rheumatol* 2000;27:675–9.
- 45 Yen JC, Abrahamowicz M, Dobkin PL *et al.* Determinants of discordance between patients and physicians in their assessment of lupus disease activity. *J Rheumatol* 2003;30:1967–76.
- 46 Bexelius C, Wachtmeister K, Skare P, Jönsson L, Vollenhoven R. V. 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;22:793–801.
- 47 Katz P, Nelson WW, Daly RP *et al.* Patient-reported lupus flare symptoms are associated with worsened patient outcomes and increased economic burden. *J Manag Care Spec Pharm* 2020;26:275–83.
- 48 Katz P, Wan GJ, Daly P *et al.* Patient-reported flare frequency is associated with diminished quality of life and family role functioning in systemic lupus erythematosus. *Qual Life Res* 2020;29:3251–61.
- 49 Furie R, Morand EF, Askanase AD *et al.* Anifrolumab reduces flare rates in patients with moderate to severe systemic lupus erythematosus. *Lupus* 2021;30:1254–63.
- 50 Tummala R, Abreu G, Pineda L, Michaels MA *et al.* Safety profile of anifrolumab in patients with active SLE: an integrated analysis of phase II and III trials. *Lupus Sci Med* 2021;8:e000464.