Central Lancashire Online Knowledge (CLoK)

Title	Mapping the disease-specific LupusQoL to the SF-6D
Type	Article
URL	https://clok.uclan.ac.uk/11402/
DOI	https://doi.org/10.1007/s11136-014-0892-4
Date	2014
Citation	Meacock, Rachel, Harrison, Mark, McElhone, Kathleen, Abbott, Janice, Haque, Sahena, Bruce, Ian and Teh, Lee-Suan (2014) Mapping the disease-specific LupusQoL to the SF-6D. Quality of Life Research. ISSN 0962-9343
Creators	Meacock, Rachel, Harrison, Mark, McElhone, Kathleen, Abbott, Janice, Haque, Sahena, Bruce, Ian and Teh, Lee-Suan

It is advisable to refer to the publisher's version if you intend to cite from the work.
https://doi.org/10.1007/s11136-014-0892-4

For information about Research at UCLan please go to http://www.uclan.ac.uk/research/

Mapping the disease-specific LupusQoL to the SF-6D

Meacock R, Harrison M, McElhone K, Abbott J, Haque S, Bruce I, Teh LS.

ABSTRACT

Purpose

To derive a mapping algorithm to predict SF-6D utility scores from the non-preference-based LupusQoL and test the performance of the developed algorithm on a separate independent validation data set.

Method

LupusQoL and SF-6D data were collected from 320 patients with systemic lupus erythematosus (SLE) attending routine rheumatology outpatient appointments at seven centres in the UK. Ordinary least squares (OLS) regression was used to estimate models of increasing complexity in order to predict individuals' SF-6D utility scores from their responses to the LupusQoL questionnaire. Model performance was judged on predictive ability through the size and pattern of prediction errors generated. The performance of the selected model was externally validated on an independent data set containing 113 female SLE patients who had again completed both the LupusQoL and SF-36 questionnaires.

Results

Four of the eight LupusQoL domains (physical health, pain, emotional health, and fatigue) were selected as dependent variables in the final model. Overall model fit was good, with R^2 0.7219, MAE 0.0557, and RMSE 0.0706 when applied to the estimation data set, and R^2 0.7431, MAE 0.0528, and RMSE 0.0663 when applied to the validation sample.

Conclusion

This study provides a method by which health state utility values can be estimated from patient responses to the non-preference-based LupusQoL, generalisable beyond the data set upon which it was estimated. Despite concerns over the use of OLS to develop mapping algorithms, we find this method to be suitable in this case due to the normality of the SF-6D data.

INTRODUCTION

Economic evaluations are increasingly needed to guide decisions on how to best allocate scarce health care resources. In order to compare across different health conditions and interventions, the results of such evaluations must be expressed in a common metric. International advisory bodies such as The National Institute for Health and Care Excellence (NICE) in England and Wales[1], The Academy of Managed Care Pharmacy (AMCP) in the US[2] and The Canadian Agency for Drugs and Technologies in Health (CADTH)[3] have specified cost-utility analysis (CUA) as the preferred form of economic evaluation, with quality-adjusted life years (QALYs) therefore being used as the common metric of health benefit. The QALY is a measure of health which combines length and health-related quality of life (HRQoL) into a single number. The quality dimension of the QALY represents the utility associated with a health state, and can be estimated using a preference-based measure of health. These preference-based measures attribute a pre-scored health state classification system which has been valued using preferences elicited from samples of the general public, referred to as utility, to responses to patient-reported outcome measures (PROMs). Despite this need for utility data, many trials fail to include a preference-based measure of health. Non-preference based disease-specific measures are often favoured as these instruments are thought to be more relevant to the patient populations being examined, and more responsive to changes in health resulting from treatment[4].

One solution increasingly being employed by analysts faced with a lack of available utility data is empirical mapping[5, 6]. Mapping involves estimating a statistical relationship between a non-preference based measure and a preference-based measure of HRQoL, using a dataset in which both measures have been administered to the same patients. The resulting model or algorithm can then be used to predict health state utility values for patients that have only completed the non-preference based instrument. Various model specifications can be used, ranging from a simple linear function where the utility score of the preference-based measure is regressed onto the total score of the non-preference based measure, to complex 'response mapping' models. Response mapping models predict how patients would have responded to the preference-based questionnaire using separate

models to predict the response for each dimension of a PROM[7]. Brazier and colleagues provide a useful review of published mapping studies for those unfamiliar with the approach[5].

The LupusQoL[8] is one such disease-specific non-preference based measure, developed in order to specifically assess the impact of systemic lupus erythematosus (SLE) and its treatment upon the HRQoL of people with the condition SLE[8]. SLE is a relapsing/remitting autoimmune multi-system disease of unknown aetiology[9]. The nature of the disease is such that it can affect different organs and systems in different patients, and can evolve over time. SLE has various manifestations including joint and muscle ache, rashes, extreme fatigue, hair loss, ulcers, anxiety and depression[10]. Around 90% of patients are women[10], with incidence highest in those of child-bearing age, and in non-Caucasian populations[11]. There is no cure for the condition, which has been shown to impose a substantial burden in term of HRQoL[10, 12], and treatment is therefore aimed at reducing disease activity and frequency of relapse. Few medications are currently licensed for the treatment of SLE, and the commonly used immune-suppressants can be associated with significant adverse events. However, the new age of biological therapies has seen a number of novel therapies being developed to treat the condition, showing promising results in terms of clinical-effectiveness[13]. The impact of these new treatments upon HRQoL will need to be evaluated, along with their cost-effectiveness, if they are to be widely implemented into clinical practice.

The objective of this article was to derive a mapping algorithm to predict SF-6D utility scores from the non-preference based LupusQoL, and test the performance of the developed algorithm on a separate independent validation data set.

METHODS

Instruments

The LupusQoL[8]

The LupusQoL provides a profile of scores across 8 domains of HRQoL: physical health (8 items), pain (3 items), planning (3 items), intimate relationships (2 items), burden to others (3 items),

emotional health (6 items), body image (5 items), and fatigue (4 items) (Appendix A1). These items were derived using qualitative interviews with clinicians and patients to identify areas most relevant for SLE populations, and so provide a rich source of information. Each item is scored on a 5 point Likert scale and mean raw domain scores are transformed onto a scale of 0-100, with higher numbers representing better HRQoL. Transformed domain scores can be calculated providing at least 50% of item answers are completed.

The instrument has acceptable ceiling and minimal floor effects, and demonstrates good internal reliability (Cronbach's alpha 0.88 to 0.96), test-retest reliability (r 0.72 to 0.93), concurrent validity with the comparable domains of the SF-36 (r 0.71 to 0.79), and discriminant validity for different levels of disease activity as measures by the British Isles Lupus Assessment Group (BILAG) index [8]. The sensitivity to change has been shown to be comparable to that detected by the SF-36[14]. The LupusQoL has been translated into 77 languages for use in 52 countries, and been used successfully in a number of clinical trials[15]. However, in its current form the LupusQoL is not amenable for use in CUA as it is not preference-based[4, 16].

The SF-6D[17–19]

The SF-6D is a generic preference-based measure of HRQoL derived from the SF-36[20, 21] or SF-12[22]. It covers 6 domains of HRQoL: physical functioning, role limitations, social functioning, pain, mental health, and vitality. Preference weights can be applied to the SF-6D health state classification to produce a single index utility score ranging from 0.296-1 when derived from the SF-36, where 1 represents a state of full health. The SF-6D has been found to be an appropriate measure of HRQoL for people with SLE [23], demonstrating acceptable psychometric properties [24] and displaying construct and criterion validity [25].

Both instruments are self-reported questionnaires which ask respondents to consider their HRQoL over the last 4 weeks.

Data

Estimation data set

Data collected from 320 patients during the LupusQoL development and validation study was used to estimate the mapping algorithm[8]. Patients were approached whilst attending routine rheumatology outpatient appointments at seven centres in the UK, and so characterise a representative sample of the general SLE population. All patients met four or more of the revised American College of Rheumatology (ACR) criteria for SLE[26, 27], were at least 18 years old, literate in English and without major unstable psychiatric disease.

Validation data set

A separate independent data set from a study investigating carotid atherosclerosis in SLE was used to test the validity of the developed mapping function[25, 28]. Data were collected from 113 female patients over the age of 18 who again met four or more of the revised American College of Rheumatology (ACR) criteria for SLE[26, 27]. Patients were recruited from outpatient clinics in the North West of England. As the original study was designed to collect DNA for genetic studies, recruitment was restricted to white British women.

Patients in both data sets completed both the LupusQoL and the SF-36 questionnaires at the same follow-up assessments.

Statistical analysis

The correlation between paired observations was examined using Pearson's product-moment correlation coefficients, as the performance of any mapping function is dependent upon the strength of the underlying relationship between the two measures used. Whilst LupusQoL scores are normally reported as a profile of 8 domain scores, it was agreed with the developers that for the purpose of this exercise domain scores could be summed to create a total score ranging from a possible $0 - 800$ (KM, L-ST JA). A higher total score represents better HRQoL.

Although ordinary least squares (OLS) regression is the most widely used method for developing mapping algorithms[5], there has been debate in the recent literature over its use to analyse HRQoL data[6, 29]. These issues, however, are mainly directed at the suitability of OLS for predicting EQ-5D utility scores due its bi- or often tri-modal distribution[29] and the well documented ceiling effect of the instrument[30, 31], which typically results in a concentration of observations at 1. A number of studies have found OLS to be appropriate when mapping to the SF-6D, often finding this method to demonstrate superior performance to alternatives such as Tobit, CLAD, two-part models and response mapping methods in this context [32–34]. The distribution of the SF-6D scores in both the estimation and validation data sets were therefore tested for normality in order to assess the appropriateness of OLS to model this particular relationship.

Various model specifications were estimated in an attempt to find the best fit, upon which the final mapping algorithm was then based. The SF-6D utility score was used as the dependent variable throughout, and models estimated first using LupusQoL total score and then domain scores as explanatory variables. Squared and first degree interaction terms, plus age and gender patient characteristics were then added to investigate whether these improved model fit. OLS regression was used and a backward stepwise elimination procedure employed, retaining variables with a P-value < 0.05 and excluding those with P > 0.10. Given the size of the data set and the large number of possible responses to the LupusQoL (five possible responses on each of the 34 items), the data was not rich enough to accurately allow estimation of models using LupusQoL item scores as dependent variables or response mapping.

A number of different criteria were examined in order to assess model performance within the estimation data set. As the purpose of the function is to predict SF-6D utility scores from LupusQoL data, the size and pattern of the prediction errors is the primary concern[5, 6], therefore the mean absolute error (MAE) and root mean squared error (RMSE) were calculated at the individual patient level to assess the size of the prediction errors. Smaller MAE and RMSE values indicate greater predictive accuracy. As there are no current guidelines indicating what level of prediction errors are

7

acceptable, results are compared to those reported for other published mapping functions. In addition, observed SF-6D values were plotted against those predicted and the residuals of these predictions in order to examine the pattern of errors across the scale of the SF-6D. The range of predicted values from the developed algorithm are reported and compared to the range of observed SF-6D scores to assess the extent to which the model was affected by the commonly encountered problems of poor fit at the extremes if the distribution, systematic error patterns and compression of range[5]. Furthermore, the proportion of absolute prediction errors below 0.05 and 0.1 were calculated.

The same criteria were used to assess the model when applied to the independent validation data set, where we are again concerned with the size and pattern of the prediction errors. External validation assesses the performance of the mapping algorithm at its primary purpose; to predict SF-6D utility scores using LupusQoL responses in an unrelated data set. Additionally, the accuracy with which the developed model was able to predict the mean SF-6D score for the validation data set was examined. The ability to predict mean scores is important, as practical applications of mapping generally involve estimating utility values at this aggregate level rather than predicting individual health states[5, 6]. Furthermore, assessment of predicted mean scores is only relevant when applying the model to the independent validation data set, as OLS estimates the unknown parameters of a model so as to minimise the sum of the squared errors, and so the mean value of the predicted scores will equal that of the observed in the data upon which the model was estimated. Although written primarily to guide exercises mapping to the EQ-5D, throughout we follow the latest recommendations for good practice where applicable[6]. All analysis was undertaken using Stata version 13.

RESULTS

Descriptive statistics

Demographic and clinical characteristics of the patients in both the estimation and validation data sets are shown in Table 1. Patients in the estimation data set cover the full range of possible SF-6D scores, with a mean score of 0.615 (SD 0.130). A broad range of LupusQoL scores are observed, ranging from 26 – 800, with a mean of 509 (SD 182). The full range of possible scores on each LupusQoL

item are covered, apart from the physical health domain where observed scores range from 3 − 100. The correlation coefficient of r = 0.824 between total LupusQoL and SF-6D scores indicates a strong correlation between the two measures[35]. The strength of the correlation between the LupusQoL domain scores when considered individually and the SF-6D vary from 0.580 for the body image domain to 0.805 for the physical health domain. Mean SF-6D scores for patients in the validation sample (0.638 (SD 0.138)) were similar to those in the estimation sample, although the range of observed SF-6D scores was smaller (0.327 − 1). The range of total LupusQoL scores observed in the validation sample was also reduced, ranging from 106 − 800 with a mean of 494 (SD 194). The full range of possible scores are observed for six of the eight LupusQoL items, with scores on the emotional health domain ranging from 25 − 100 and from 8 − 100 for body image. A strong correlation exists between LupusQoL and SF-6D (correlation coefficient r = 0.847). The magnitudes of the correlations between the LupusQoL domain scores and the SF-6D are marginally larger for all domains in the validation sample, varying from 0.622 for the intimate relationships domain to 0.816 for the physical health domain.

Figure 1 shows the distribution of SF-6D scores in the estimation and validation samples. We fail to reject the null hypothesis of normality at either the 5% or 10% significance level in either data set (p=0.130 estimation sample, p=0.198 validation sample), indicating that OLS is appropriate in this situation.

Selected model and performance

The addition of squared and first degree interaction terms, age, and gender did not improve the predictive accuracy of the basic OLS model, and so these were not selected for inclusion in the final algorithm.

The specification for the selected model is shown in Table 2, alongside the model including all of the LupusQoL domains for comparison. There is very little difference between these two models in terms of the size of the coefficients, statistical significance, explanatory power, or the size of the prediction

errors generated. This indicates that the items not selected for the final model contribute very little to predicting SF-6D utilities, and so excluding these increases the efficiency of the final model. Four of the eight LupusQoL domains were selected for inclusion using the stepwise procedure: physical health (PH), pain (P), emotional health (EH), and fatigue (F). The coefficients relating to the constant and LupusQoL items displayed in the table are those produced by the stepwise regression, and are multiplied by the score on the corresponding LupusQoL item in order to calculate the predicted SF-6D score as follows:

$$SF - 6D\ utility$$
$$= 0.3040964 + 0.0014778 * PH + 0.0014531 * P + 0.0011405 * EH$$
$$+ 0.0008953 * F$$

The small magnitude of the coefficients reflects the fact that each LupusQoL item is scored 0-100. The performance statistics for the selected model are shown in Table 3. Within the estimation data set overall model fit was good, with more than 70% of the variation in SF-6D utility scores explained by the four selected items of the LupusQoL. MAE is 0.0557, which is towards the lower end of the range of MAEs reported in a recent systematic review of published mapping functions (MAE 0.0011 to 0.19)[5]. This MAE equates to an average prediction error in the magnitude of 8% of the total scale covered by the SF-6D ($0.296 - 1$). The RMSE was 0.0706, which is lower than that of any reported in the same systematic review (RMSE 0.084 to 0.2). Figure 2 plots the observed SF-6D values against those predicted by the model, and the prediction errors, illustrating the fit of the model with the estimation data set. The model predicts well for the majority of the SF-6D scale, but suffers from problems of under-prediction at the upper end. The under-prediction is caused by the inability of the model to predict utility scores larger than 0.801. The majority of errors were small in absolute magnitude, with 88% of predictions within 0.1 of the observed value and 52% within 0.05.

Model performance in the independent validation data set was comparable, which suggests that the developed mapping function between the LupusQoL and the SF-6D represents a strong statistical relationship which is applicable outside of the estimation data set. 74% of the variation in SF-6D

utility scores was explained by the model. The MAE was 0.0528 and RMSE 0.0663, both marginally lower than in the estimation data set. Figure 3 plots the observed SF-6D values in the validation data set against those predicted by the model, and the prediction errors, showing similar model fit to that observed in the estimation data set. The model again predicts well for the majority of the SF-6D scale, but the problem of under-prediction at the upper end persists. The majority of errors were again small in absolute magnitude, with 93% of predictions within 0.1 of the observed value and 54% within 0.05. Finally, the model performed extremely well when predicting the mean SF-6D utility score for the validation sample. The observed mean was 0.624 and the predicted mean 0.617.

CONCLUSIONS

This study provides a method by which health state utility values can be estimated from patient responses to the non-preference based disease-specific LupusQoL measure. The developed mapping algorithm performs well in comparison to other published studies[5], and so provides a practical solution for researchers seeking to use existing datasets where the LupusQoL but no preference-based utility measure is collected for economic evaluation. The encouraging performance of the mapping function may be a product of the conceptual similarity of the LupusQol and the SF-6D measures compared with other mapping functions which attempt to map to preference-based HRQoL measures from more narrow measures of outcome, such as the Heath Assessment Questionnaire which measures functional disability[36]. The validation of the model using an independent data set is also a strength of the study, demonstrating the generalizability of the developed algorithm beyond the data set upon which is was estimated.

Four of the eight LupusQoL dimensions were selected for inclusion in the final model: physical health, pain, emotional health, and fatigue. The inclusion of these domains is not surprising given the conceptual overlap of these dimensions with those of the SF-6D: physical functioning, role limitations, social functioning, pain, mental health, and vitality. Nor is it surprising that gender did not significantly improve the predictive accuracy of the model, given the predominance of females in the sample. The remaining four dimensions of the LupusQoL; planning, intimate relationships, burden to

others, and body image, are clearly of importance to people with SLE [8]. The omission of these dimensions from the final model does not dispute this, but reflects that these domains do not significantly impact upon utility as measured by the SF-6D. Whilst the collection of both a generic preference-based and a disease-specific measure of HRQoL remains the preferred method, the strong statistical relationship demonstrated means that researchers concerned about the burden placed upon patients by administering multiple measures may be able to obtain the necessary clinical information from just using the LupusQoL, and predict utility scores based upon these responses. However, the longitudinal validity of this algorithm, which has highlighted limitations in other mapping functions[37], has not yet been tested. Therefore, we reiterate previous recommendations to include at least one preference-based measure of HRQoL in all relevant clinical studies wherever possible[37–39]. The lack of interchangeability among different preference-based measures should also be considered by those wishing to use the algorithm presented here, as it has been developed to predict utility as measured by the SF-6D, and so may produce different values to HRQoL measured by another preference-based measure such as the EQ-5D[40]. The impact of using different preference-based measures on QALY estimates has been demonstrated and discussed in the field of rheumatology[41, 42].

Although the predictive accuracy of the mapping function is impressive, the model is unable to predict utility scores above 0.801. A recent study has highlighted that regression to the mean may be the cause of the reduction in variance and prediction range often observed in mapping studies [43]. However, this would affect both the upper and lower ends of the scale. As our function does not suffer from an inability to predict scores at the lower end of the SF-6D range, it appears more likely that the problem is due to a lack of observations in the estimation data set covering high SF-6D scores. This absence is a reflection of the large detrimental effect of SLE on HRQoL[10]. In our representative samples of outpatients, just 9% of the estimation and 10% of the validation sample have observed SF-6D scores above this level. In the context of economic evaluations it is the error at group level which is important[44, 45], which was shown to be minimal when the model was applied to the independent validation data. Caution should however be aired if applying the function to a particularly mild patient

population, or if treatment is expected to return HRQoL to near full health. The restriction of our validation sample to white British women means that we were unable to test the developed model on an entirely representative SLE population. However, previous analysis performed on our estimation sample found no significant differences in LupusQoL scores by sex or ethnicity as classified as Black-Caribbean, Asian, and White[46]. This restriction is therefore likely to have minimal implications for the validity of our exercise.

Despite concerns over the use of OLS, we find this method of estimation to be suitable in this case due to the normality of the SF-6D data. Alternative methods for mapping, such as the linking (scale-aligning)[43, 47], have been proposed to overcome the limitations of OLS, such as the inability to predict accurately in high and low regions of the scale due to regression to the mean. However, in our case we were unable to operationalise this approach as we considered some of the conditions required for this technique to be violated[43], most notably that there is no official overall scale for the LupusQoL to link to an overall SF-6D scale measuring the same construct of HRQoL. Whilst we used a 'total' LupusQoL score for exploratory purposes in this study, this is currently neither validated nor recommend for clinical use. The development of an overall LupusQoL score, especially if combined with preference-based values of the health states described, would represent an important development, both for the measure itself and in allowing further refinement of its relationship with the SF-6D.

In agreement with previous studies[5], we find a simple additive model with the utility score as the dependent variable and dimension scores as independent variables to be the most appropriate functional form, with the addition of squared terms, first degree interactions, and patient characteristics having little impact on model performance. Our results do, however, oppose previous observations that the degree of error tends to be larger when mapping from a disease-specific measure rather than another generic instrument[2]. The strong predictive results observed in this study are likely due to the nature of the disease in question. The breadth of symptoms resulting from SLE and its various manifestations mean that the LupusQoL domains cover a much broader range of HRQoL

elements than many other disease-specific measures. Mapping relies on the conceptual overlap between the two measures used, not only so that the generic measure is able to capture all of the relevant effects of the disease in question, but also to ensure that the disease-specific instrument is able to capture comorbidities and side effects of potential treatments. There appears to be sufficient overlap between the SF-6D and LupusQoL to allow the estimation of a useful mapping function.

REFERENCES

1. Guide to the methods of technology appraisal 2013 Foreword PMG9. (n.d.). Retrieved March 10, 2014, from http://publications.nice.org.uk/guide-to-the-methods-of-technology-appraisal-2013-pmg9

2. Sullivan, S., Lyles, A., Luce, B., & Grigar, J. (n.d.). AMCP guidance for submission of clinical and economic evaluation data to support formulary listing in U.S. health plans and pharmacy benefits management organizations. *Journal of Managed Care Pharmacy*, *7*(4), 272–282.

3. Canadian Agency for Drugs and Technologies in Health. (2006). *Guidelines for the Economic Evaluation of Health Technologies*. Canada. Retrieved from http://www.cadth.ca/media/pdf/186_EconomicGuidelines_e.pdf

4. Gold, M. R. (1996). *Cost-Effectiveness in Health and Medicine*. Oxford University Press.

5. Brazier, J. E., Yang, Y., Tsuchiya, A., & Rowen, D. L. (2010). A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. *The European journal of health economics: HEPAC: health economics in prevention and care*, *11*(2), 215–225. doi:10.1007/s10198-009-0168-z

6. Longworth, L., & Rowen, D. (2013). Mapping to obtain EQ-5D utility values for use in NICE health technology assessments. *Value in health: the journal of the International Society for Pharmacoeconomics and Outcomes Research*, *16*(1), 202–210. doi:10.1016/j.jval.2012.10.010

7. Gray, A. M., Rivero-Arias, O., & Clarke, P. M. (2006). Estimating the Association between SF-12 Responses and EQ-5D Utility Values by Response Mapping. *Medical Decision Making*, *26*(1), 18–29. doi:10.1177/0272989X05284108

8. McElhone, K., Abbott, J., Shelmerdine, J., Bruce, I. N., Ahmad, Y., Gordon, C., … Teh, L.-S. (2007). Development and validation of a disease-specific health-related quality of life measure, the LupusQol, for adults with systemic lupus erythematosus. *Arthritis and rheumatism*, *57*(6), 972–979. doi:10.1002/art.22881

9. O'Neill, S., & Cervera, R. (2010). Systemic lupus erythematosus. *Best practice & research. Clinical rheumatology*, *24*(6), 841–855. doi:10.1016/j.berh.2010.10.006

10. Meacock, R., Dale, N., & Harrison, M. J. (2013). The humanistic and economic burden of systemic lupus erythematosus : a systematic review. *PharmacoEconomics*, *31*(1), 49–61. doi:10.1007/s40273-012-0007-4

11. Petri, M. (2002). Epidemiology of systemic lupus erythematosus. *Best practice & research. Clinical rheumatology*, *16*(5), 847–858.

12. McElhone, K., Abbott, J., & Teh, L. S. (2006). A review of health related quality of life in systemic lupus erythematosus. *Lupus*, *15*(10), 633–643.

13. Bruce, I. N. (2010). Re-evaluation of biologic therapies in systemic lupus erythematosus. *Current opinion in rheumatology*, *22*(3), 273–277. doi:10.1097/BOR.0b013e3283374e78

14. Touma, Z., Gladman, D. D., Ibañez, D., & Urowitz, M. B. (2011). Is There an Advantage Over SF-36 with a Quality of Life Measure That Is Specific to Systemic Lupus Erythematosus? *The Journal of Rheumatology*, *38*(9), 1898–1905. doi:10.3899/jrheum.110007

15. Clinical Trials - Corporate Translations. (n.d.). Retrieved February 7, 2014, from http://www.corptransinc.com/Sites/LupusQoL/Instrument-Information/Clinical-Trials.aspx

16. Drummond, M. F., Wilson, D. A., Kanavos, P., Ubel, P., & Rovira, J. (2007). Assessing the economic challenges posed by orphan drugs. *International journal of technology assessment in health care*, *23*(1), 36–42. doi:10.1017/S0266462307051550

17. Brazier, J., Roberts, J., & Deverill, M. (2002). The estimation of a preference-based measure of health from the SF-36. *Journal of health economics*, *21*(2), 271–292.

18. Brazier, J. E., & Roberts, J. (2004). The estimation of a preference-based measure of health from the SF-12. *Medical care*, *42*(9), 851–859.

19. Brazier, J., Usherwood, T., Harper, R., & Thomas, K. (1998). Deriving a Preference-Based Single Index from the UK SF-36 Health Survey. *Journal of Clinical Epidemiology*, *51*(11), 1115–1128. doi:10.1016/S0895-4356(98)00103-6

20. Brazier, J. (1993). The SF-36 health survey questionnaire—a tool for economists. *Health Economics*, *2*(3), 213–215. doi:10.1002/hec.4730020304

21. Ware, J. E., Snow, K. K., Kosinski, M., Gandek, B., & Institute, N. E. M. C. H. H. (1993). *SF-36 health survey: manual and interpretation guide*. The Health Institute, New England Medical Center.

22. Ware, J., Jr, Kosinski, M., & Keller, S. D. (1996). A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Medical care*, *34*(3), 220–233.

23. Sanchez, M, McGwin, G, Duran, S, Fernandez, M, Reveille, J, Vila, L, & Alarcon, G. (2009). Factors predictive of overall health over the course of the disease in patients with systemic lupus erythematosus

from the LUMINA cohort (LXII): use of the SF-6D. *Clinical and Experimental Rheumatology*, *27*, 64–71.

24. Aggarwal, R., Wilke, C. T., Pickard, A. S., Vats, V., Mikolaitis, R., Fogg, L., … Jolly, M. (2009). Psychometric Properties of the EuroQol-5D and Short Form-6D in Patients with Systemic Lupus Erythematosus. *The Journal of Rheumatology*, *36*(6), 1209–1216. doi:10.3899/jrheum.081022

25. Harrison, M. J., Ahmad, Y., Haque, S., Dale, N., Teh, L.-S., Snowden, N., … Bruce, I. N. (2012). Construct and criterion validity of the short form-6D utility measure in patients with systemic lupus erythematosus. *The Journal of rheumatology*, *39*(4), 735–742. doi:10.3899/jrheum.110648

26. Tan, E. M., Cohen, A. S., Fries, J. F., Masi, A. T., McShane, D. J., Rothfield, N. F., … Winchester, R. J. (1982). The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis and rheumatism*, *25*(11), 1271–1277.

27. Hochberg, M. C. (1997). Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis and rheumatism*, *40*(9), 1725. doi:10.1002/1529-0131(199709)40:9<1725::AID-ART29>3.0.CO;2-Y

28. Ahmad, Y., & Bruce, I. (2004). Sunclinical atherosclerosis in systemic lupus erythematosus. *The Journal of Rheumatology*, *31*, 841–3.

29. Hernández Alava, M., Wailoo, A. J., & Ara, R. (2012). Tails from the Peak District: Adjusted Limited Dependent Variable Mixture Models of EQ-5D Questionnaire Health State Utility Values. *Value in Health*, *15*(3), 550–561. doi:10.1016/j.jval.2011.12.014

30. Pullenayegum, E. M., Tarride, J.-E., Xie, F., Goeree, R., Gerstein, H. C., & O'Reilly, D. (2010). Analysis of Health Utility Data When Some Subjects Attain the Upper Bound of 1: Are Tobit and CLAD Models Appropriate? *Value in Health*, *13*(4), 487–494. doi:10.1111/j.1524-4733.2010.00695.x

31. Janssen, M. F., Pickard, A. S., Golicki, D., Gudex, C., Niewada, M., Scalone, L., … Busschbach, J. (2013). Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation*, *22*(7), 1717–1727. doi:10.1007/s11136-012-0322-4

32. Brazier, J., Connell, J., Papaioannou, D., Mukuria, C., Mulhern, B., Peasgood, T., … Parry, G. (2014). A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures. *Health Technology Assessment*, *18*(34). doi:10.3310/hta18340

33. Browne, C., Brazier, J., Carlton, J., Alavi, Y., & Jofre-Bonet, M. (2012). Estimating quality-adjusted life years from patient-reported visual functioning. *Eye*, *26*(10), 1295–1301. doi:10.1038/eye.2012.137

34. Yang, Y., Wong, M. Y., Lam, C. L. K., & Wong, C. K. H. (2014). Improving the mapping of condition-specific health-related quality of life onto SF-6D score. *Quality of Life Research*, *23*(8), 2343–2353. doi:10.1007/s11136-014-0668-x

35. Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Psychology Press.

36. Bansback, N., Marra, C., Tsuchiya, A., Anis, A., Guh, D., Hammond, T., & Brazier, J. (2007). Using the health assessment questionnaire to estimate preference-based single indices in patients with rheumatoid arthritis. *Arthritis Care & Research*, *57*(6), 963–971. doi:10.1002/art.22885

37. Harrison, M. J., Lunt, M., Verstappen, S. M., Watson, K. D., Bansback, N. J., & Symmons, D. P. (2010). Exploring the validity of estimating EQ-5D and SF-6D utility values from the health assessment questionnaire in patients with inflammatory arthritis. *Health and Quality of Life Outcomes*, *8*(1), 21. doi:10.1186/1477-7525-8-21

38. Gabriel, S., Drummond, M., Maetzel, A., Boers, M., Coyle, D., Welch, V., & Tugwell, P. (2003). OMERACT 6 Economics Working Group report: a proposal for a reference case for economic evaluation in rheumatoid arthritis. *The Journal of Rheumatology*, *30*(4), 886–890.

39. Barton, G. R., Sach, T. H., Jenkinson, C., Avery, A. J., Doherty, M., & Muir, K. R. (2008). Do estimates of cost-utility based on the EQ-5D differ from those based on the mapping of utility scores? *Health and Quality of Life Outcomes*, *6*(1), 51. doi:10.1186/1477-7525-6-51

40. Feeny, D., Spritzer, K., Hays, R. D., Liu, H., Ganiats, T. G., Kaplan, R. M., … Fryback, D. G. (2012). Agreement About Identifying Patients Who Change Over Time: Cautionary Results in Cataract and Heart Failure Patients. *Medical decision making : an international journal of the Society for Medical Decision Making*, *32*(2), 273–286. doi:10.1177/0272989X11418671

41. Marra, C. A., Marion, S. A., Guh, D. P., Najafzadeh, M., Wolfe, F., Esdaile, J. M., … Anis, A. H. (2007). Not all "quality-adjusted life years" are equal. *Journal of Clinical Epidemiology*, *60*(6), 616–624. doi:10.1016/j.jclinepi.2006.09.006

42. Kirwan, J. R., Boonen, A., Harrison, M. J., Hewlett, S. E., Wells, G. A., Singh, J. A., … Dworkin, R. H. (2011). OMERACT 10 Patient Perspective Virtual Campus: Valuing Health; Measuring Outcomes in Rheumatoid Arthritis Fatigue, RA Sleep, Arthroplasty, and Systemic Sclerosis; and Clinical

Significance of Changes in Health. *The Journal of Rheumatology*, *38*(8), 1728–1734.

doi:10.3899/jrheum.110393

43. Fayers, P. M., & Hays, R. D. (2014). Should Linking Replace Regression When Mapping from Profile-Based Measures to Preference-Based Measures? *Value in Health*, *17*(2), 261–265.

doi:10.1016/j.jval.2013.12.002

44. Rowen, D., Brazier, J., & Roberts, J. (2009). Mapping SF-36 onto the EQ-5D index: how reliable is the relationship? *Health and Quality of Life Outcomes*, *7*(1), 27. doi:10.1186/1477-7525-7-27

45. Ara, R., & Brazier, J. (2008). Deriving an algorithm to convert the eight mean SF-36 dimension scores into a mean EQ-5D preference-based score from published studies (where patient level data are not available). *Value in health: the journal of the International Society for Pharmacoeconomics and Outcomes Research*, *11*(7), 1131–1143. doi:10.1111/j.1524-4733.2008.00352.x

46. McELHONE, K., Castelino, M., Abbott, J., Bruce, I. N., Ahmad, Y., Shelmerdine, J., … Teh, L.-S. (2010). The LupusQoL and Associations with Demographics and Clinical Measurements in Patients with Systemic Lupus Erythematosus. *The Journal of Rheumatology*, *37*(11), 2273–2279. doi:10.3899/jrheum.091277

47. Fayers, P. M., & Hays, R. D. (2014). Don't middle your MIDs: regression to the mean shrinks estimates of minimally important differences. *Quality of Life Research*, *23*(1), 1–4. doi:10.1007/s11136-013-0443-4

Table 1 Characteristics of the estimation and validation samples

Characteristic	Estimation sample (n = 320)	Validation sample (n = 113)
Age, years, mean (SD)	44.8 (13.6)	48.6 (9.3)
Female, %	95	100
Disease duration, years, mean (SD)	10.5 (8.7)	12.6 (9.7)
SF-6D index score, mean (SD), range	0.615 (0.130) 0.296 - 1	0.638 (0.138) 0.327 - 1
LupusQoL total score, mean (SD), range	509 (182) 26 - 800	494 (194) 106 - 800
LupusQoL domain scores, mean (SD), range		
Physical Health	61 (27) 3 - 100	58 (27) 0 - 100
Pain	65 (27) 0 - 100	63 (29) 0 - 100
Planning	66 (30) 0 - 100	67 (29) 0 - 100
Intimate Relationships	61 (33) 0 - 100	59 (34) 0 - 100
Burden to Others	58 (28) 0 - 100	62 (29) 0 - 100
Emotional Health	73 (21) 0 - 100	77 (19) 25 - 100
Body Image	72 (26) 0 - 100	70 (26) 8 - 100
Fatigue	52 (26) 0 - 100	53 (28) 0 - 100
Pearson's correlation coefficient between paired LupusQol and SF-6D scores, r	0.824	0.847
Bivariate Pearson's correlation coefficients between paired LupusQoL domain and SF-6D utility scores, r		
Physical Health	0.805	0.816
Pain	0.789	0.811
Planning	0.692	0.697
Intimate Relationships	0.611	0.622
Burden to Others	0.636	0.684
Emotional Health	0.667	0.761
Body Image	0.580	0.690
Fatigue	0.737	0.747

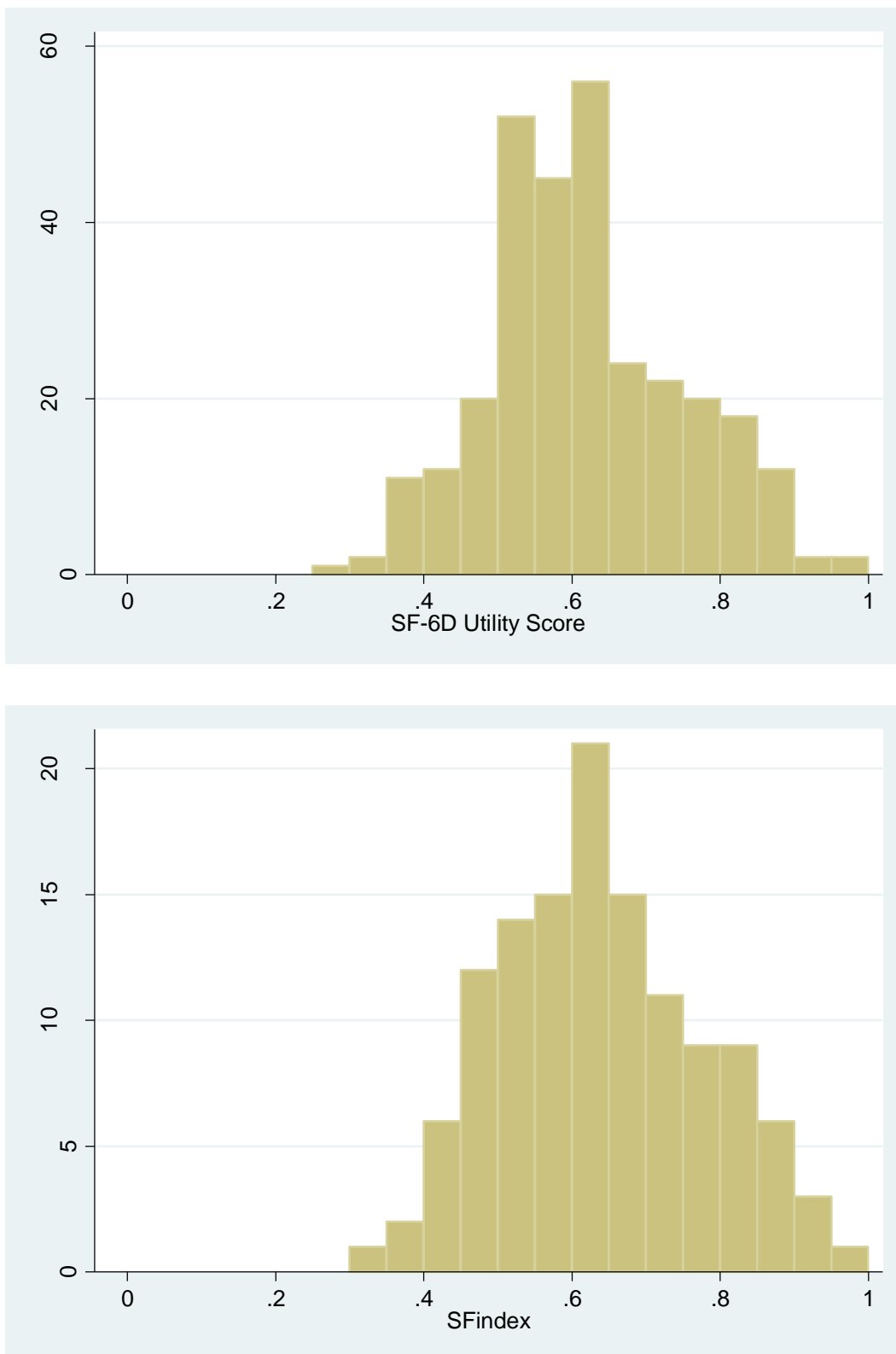**Figure 1 Distribution of SF-6D scores in the estimation (upper) and validation (lower) samples**

Table 2 Model containing all LupusQoL domains and selected model for predicting SF-6D utility scores from patients' LupusQoL scores

	Model including all LupusQoL domains			Selected stepwise model		
Domain	β	Standard Error	*p* value	β	Standard Error	*p* value
Physical Health (PH)	0.0014747	0.0004132	< 0.001	0.0014778	0.0003839	< 0.001
Pain (P)	0.0014480	0.0003319	< 0.001	0.0014531	0.0003247	< 0.001
Planning	-0.0001061	0.0002853	0.710			
Intimate Relationships	0.0000107	0.0001952	0.957			
Burden to Others	0.0001495	0.0002392	0.533			
Emotional Health (EH)	0.0011993	0.0003634	0.001	0.0011405	0.0002972	<0.001
Body Image	-0.0000789	0.0002653	0.766			
Fatigue (F)	0.0008937	0.0003180	0.005	0.0008953	0.0003046	0.004
Constant	0.3037726	0.0168842	< 0.001	0.3040964	0.0162895	< 0.001
Adjusted R^2			0.7136			0.7175
RMSE			0.0711			0.0706

Table 3 Performance of selected model in the estimation and validation data sets

Performance	Estimation data	Validation data
No. of observations	256	109
R^2	0.7219	0.7431
MAE	0.0557	0.0528
RMSE	0.0706	0.0663
Predictions within ± 0.05 of observed value	52%	54%
Predictions within ± 0.1 of observed value	88%	93%
Range of observed values	0.296 – 1.00	0.327 – 0.938
Range of predicted values	0.324 – 0.801	0.392 – 0.801
Mean observed SF-6D (SD)		0.624 (0.130)
Mean predicted SF-6D (SD)		0.617 (0.116)

Figure 2 Observed and predicted SF-6D values and prediction errors for the model applied to the estimation data set
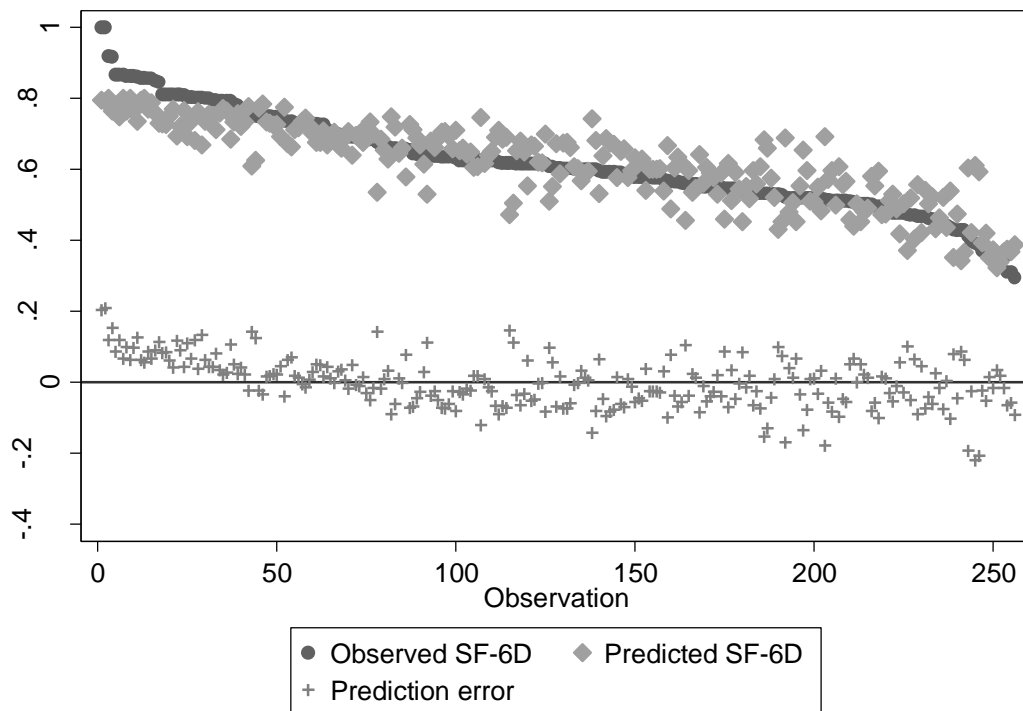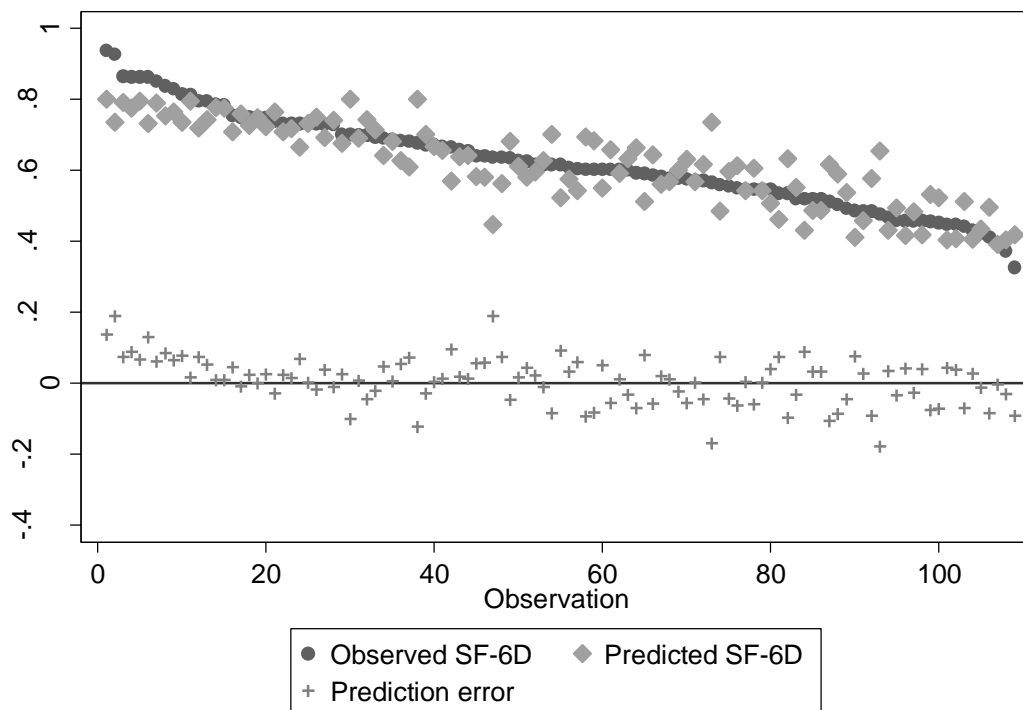

Figure 3 Observed and predicted SF-6D values and prediction errors for the model applied to the validation data set

APPENDIX

Table A1. Overview of the domains and items of the LupusQoL

LupusQol Domain	Aspects of life covered by each LupusQoL item
Physical Health	1. Help needed to do heavy physical jobs e.g. digging the garden, decorating, moving furniture 2. Help needed to do moderate physical jobs e.g. vacuuming, ironing, shopping, cleaning the bathroom 3. Help needed to do light physical jobs e.g. cooking, opening jars, dusting, combing my hair, attending to personal hygiene 4. Unable to perform everyday tasks as well as I would like to e.g. my job, childcare, housework 5. Difficulty climbing stairs 6. Have lost some independence and am reliant on others 7. Have to do things at a slower pace 8. Sleep pattern is disturbed
Pain	9. Prevented from performing activities the way I would like because of pain 10. Pain experienced interferes with the quality of my sleep 11. Pain is so severe it limits my mobility
Planning	12. I avoid planning to attend events in the future 13. Unable to organise my life efficiently due to the unpredictability of my Lupus 14. Difficult to commit to social arrangements because symptoms vary from day to day
Intimate Relationships	15. Less interested in a sexual relationship due to the pain I experience 16. Not interested in sex
Burden to Others	17. Concerned that my Lupus is stressful for those close to me 18. Concerned that I cause worry to those close to me 19. Feel that I am a burden to my friends and/or family
Emotional Health	20. Resentful 21. So fed up nothing can cheer me up 22. Sad 23. Anxious 24. Worried 25. Lacking in self-confidence
Body Image	26. Physical appearance interferes with my enjoyment of life 27. My appearance (e.g. rash, weight gain/loss) makes me avoid social situations 28. Skin rashes make me feel less attractive 29. Hair loss I have experienced makes me feel less attractive 30. Weight gain I have experienced because of my treatment makes me feel less attractive
Fatigue	31. Cannot concentrate for long periods of time 32. Feel worn out and sluggish 33. Need to have early nights 34. Often exhausted in the morning

Respondents are asked to read the statement for each item, and choose one of five responses which most closely relates to how they feel. Each statement is phrased in relation to a respondent's Lupus, and asks them to consider the previous four week period. A copy of the LupusQoL questionnaire can be found in [8].