OPEN ACCESS

EXTENDED REPORT

Common measure of quality of life for people with systemic sclerosis across seven European countries: a cross-sectional study

Mwidimi Ndosi,[1,2] Begonya Alcacer-Pitarch,[3,4] Yannick Allanore,[5] Francesco del Galdo,[3,4] Marc Frerix,[6] Sílvia García-Díaz,[7] Roger Hesselstrand,[8] Christine Kendall,[6] Marco Matucci-Cerinic,[9,10] Ulf Mueller-Ladner,[6] Gunnel Sandqvist,[8] Vicenç Torrente-Segarra,[7] Tim Schmeiser,[6,11] Matylda Sierakowska,[12] Justyna Sierakowska,[13] Stanslaw Sierakowski,[14] Anthony Redmond[3,4]

Handling editor Josef S Smolen

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2017-212412).

For numbered affiliations see end of article.

Correspondence to
Dr Mwidimi Ndosi, Department of Nursing and Midwifery, University of the West of England, Bristol, UK ; mwidimi.ndosi@uwe.ac.uk

Received 19 September 2017
Revised 9 January 2018
Accepted 30 January 2018
Published Online First 20 February 2018

ABSTRACT

Objectives The aim of this study was to adapt the Systemic Sclerosis Quality of Life Questionnaire (SScQoL) into six European cultures and validate it as a common measure of quality of life in systemic sclerosis (SSc).

Methods This was a seven-country (Germany, France, Italy, Poland, Spain, Sweden and UK) cross-sectional study. A forward–backward translation process was used to adapt the English SScQoL into target languages. SScQoL was completed by patients with SSc, then data were validated against the Rasch model. To correct local response dependency, items were grouped into the following subscales: function, emotion, sleep, social and pain and reanalysed for fit to the model, unidimensionality and cross-cultural equivalence.

Results The adaptation of the SScQoL was seamless in all countries except Germany. Cross-cultural validation included 1080 patients with a mean age 58.0 years (SD 13.9) and 87% were women. Local dependency was evident in individual country data. Grouping items into testlets corrected the local dependency in most country specific data. Fit to the model, reliability and unidimensionality was achieved in six-country data after cross-cultural adjustment for Italy in the social subscale. The SScQoL was then calibrated into an interval level scale.

Conclusion The individual SScQoL items have translated well into five languages and overall, the scale maintained its construct validity, working well as a five-subscale questionnaire. Measures of quality of life in SSc can be directly compared across five countries (France, Poland Spain, Sweden and UK). Data from Italy are also comparable with the other five countries although require an adjustment.

INTRODUCTION

Systemic sclerosis (SSc) is a heterogeneous connective tissue disease characterised by vasculopathy, immune activation and fibrosis.[1–3] The multisystem involvement in the disease has severe physical and psychosocial impact affecting the patients' quality of life (QoL). QoL is a complex interaction between the ways in which people perceive their health and how it relates to other aspects of their lives that are less directly health-specific.

Several tools have been used in different studies to capture QoL in people with SSc, such as the SF-36 and the EuroQol 5-Domain health questionnaire,[4–6] however, these tools are not disease-specific and can be less sensitive to the more directly disease-related factors. To capture the true psychosocial impact of the disease, a needs-based disease-specific QoL is the gold standard. The Systemic Sclerosis Quality of Life Questionnaire (SScQoL), developed by Reay[7] and translated into six languages in this paper, was developed according to this principle. The SScQoL tool measures the disease impact on health and well-being, and has been developed using a needs-based quality of life model, which is based on the understanding that individuals are driven or motivated by their needs and that life gains its quality from the ability and capacity of individuals to satisfy their needs.[8 9]

During its development, the original SScQoL[7] was subject to strict principles of item response theory to ensure the highest quality measure of needs-based patient-reported QoL reporting in people with SSc. The SScQoL joins a stable of measures including the Rheumatoid Arthritis Quality of Life,[10] Osteoarthritis Quality of Life[11] and Ankylosing Spondylitis Quality of Life[12] developed at the University of Leeds and forming the cornerstone of patient-reported outcome measurement (PROM) in many rheumatological conditions.

The SScQoL is a self-completed questionnaire comprising 29 questions exploring the impact of SSc on health and well-being, covering four themes identified by patients with SSc: emotion, physical adaptation, impact on/with others and impact on self. It takes the patient approximately 5 mins to complete and provides quantitative data that enables the health professional involved to accurately evaluate the impact of SSc on an individual patient or groups of people with the disease. Due to its robust validation, the SScQoL can also be used with confidence as a research tool to evaluate pharmacological and non-pharmacological interventions.

Initial development and testing demonstrated the reliability, validity and the patient acceptance of the instrument and the original English language version of the tool has been subjected to Rasch

To cite: Ndosi M, Alcacer-Pitarch B, Allanore Y, *et al*. *Ann Rheum Dis* 2018;**77**:1032–1038.

Analysis to ensure its construct validity, unidimensionality and absence of differential item functioning.[7]

The relevance of a common measure

SSc is considered a rare disease due to its prevalence (82 per 1 000 000 adjusted for the UK population).[13] The small numbers of people affected by SSc causes methodological problems, particularly when developing research studies requiring large sample sizes. To overcome these problems, there is a need for multicentre and international studies, using common outcome measures, which have demonstrable cross-cultural relevance and measurement equivalence, which in turn allow researchers to obtain reliable results that are comparable across countries. In addition, the existing European collaborations and networks such as EUSTAR and EUSHNet can employ a common measure prospectively in a systematic way, such that the networks and patients in the countries involved can benefit from the consistency provided by a cross-culturally valid measure.

The objectives of this study therefore were to: (i) translate and adapt the SScQoL for use in Germany, France, Italy, Poland, Spain and Sweden; (ii) undertake a cross-cultural validation of the SScQoL for use in these countries; (iii) calibrate a common scale that is comparable across countries and (iv) ultimately incorporate the translated and validated version of the SScQoL into the EUSTAR MEDS database to create a common minimum dataset for PROMS in SSc research in Europe. This paper reports the results of the objectives (i) through (iii).

METHODS

Study design

This was a multicentre cross-sectional analytic study involving seven European countries; Germany, France, Italy, Poland, Spain, Sweden and the UK. The study involved two phases (i) cross-cultural adaptation and (ii) cross-cultural validation.

Cross-cultural adaptation phase

The English SScQoL was adapted into six languages using the well-established process of cross-cultural adaptation of self-report measures.[14] The aim of cross-cultural adaptation is to ensure conceptual equivalence between original and target versions of a questionnaire. This process involved the following stages for each of the collaborating centres: (i) the original (English) version was translated into the target language by two translators working independently; (ii) the translations were compared and any inconsistencies resolved; (iii) the translated tools were then translated back into English by a translator not involved in stage one; (iv) once satisfactory translations had evolved, all four versions were reviewed by an expert committee and any outstanding inconsistencies were resolved by discussion and (v) the adapted questionnaires were completed by 30 patients with SSc in each of the collaborating centres.

Cross-cultural validation phase

The aim of this stage was to ensure measurement equivalence of the SScQoL, to enable common measurement across the seven countries. During this phase, the translated questionnaires were completed by 100–270 patients (in each country) by either postal or site survey. Participants were native speakers of the target languages except in Sweden where seven participants were non-native but all had lived in Sweden for several years and had a good ability to speak and read Swedish. The data from the new SScQoL were then subjected to Rasch analysis, which involved testing the construct validity of each translated tool, internal consistency and the cross-cultural invariance of the tool across all the seven countries. Finally, the common measure was calibrated, which takes account of cultural differences and, if successful, provides for pooling and comparison of measurements across the various culturally adapted versions.

Patients

Each centre recruited a convenience sample of patients from rheumatology outpatient clinics and/or patient databases. The inclusion criteria were: (i) consultant diagnosis of SSc according to ARA/ACR 1980 criteria,[15] (ii) aged ≥18 years and (iii) willingness and ability to complete and return a questionnaire. The only exclusion criterion was an inability to understand or complete the written questionnaire. Participation in the study was voluntary and each of the collaborating centres followed ethical procedures applicable to their respective countries before recruiting patients. Local investigators in each collaborating country handled all patients' (interview) data collected during the cross-cultural adaptation phase. The data collected during the validation phase in each country was then sent to the University of Leeds for psychometric testing using Rasch models. Data transferred were limited to anonymised SScQoL data containing patient's age and gender information.

Data analysis

The validation data were analysed using RUMM2030 software (RUMM Laboratory, Perth, Western Australia). First, each country-specific dataset was tested for fit to the Rasch model. Fit to the Rasch model implies construct validity, reliability, unidimensionality and statistical sufficiency of the total score from the scale.[16–19] Model fit was determined by item–person interaction statistics which compare the difference between observed responses and values expected by the model (standardised residuals). The following statistics suggest fit to the model: (i) item–person interaction statistics, distributed as a Z statistic with a mean of 0 and SD of 1; (ii) item χ^2 statistic (comparing the difference between observed and expected values) with a non-significant probability—several χ^2 are computed for each item across groups, therefore Bonferroni adjustment is required to avoid type I errors due to multiple testing;[20] (iii) item–trait interaction statistic reported as a non-significant χ^2 probability, reflecting the invariance of the SScQoL to different levels of quality of life.

An estimate of internal consistency (reliability) was determined by person separation index (PSI), which represents the ability of the SScQoL to distinguish between people with different levels of reported quality of life. A value of 0.7 is required for group use.[19]

Although fit to Rasch model implies unidimensionality of the scale, further tests were carried out to confirm the assumption of local independence of items,[21] unidimensionality and differential item functioning. The Rasch model assumes that each item independently contributes to the underlying construct, no significant item–item residual correlations are expected therefore, after contribution to the construct is removed. Where significant item–item residual correlations were identified (through residual correlation matrices), these locally dependent items were grouped and treated as a unit, referred to as a 'testlet', which represent a subscale. Two investigators (MN and ACR) grouped the items by consensus into the following testlets: function, emotion, sleep, social and pain, which in turn map onto the International Classification

of Functioning, Disability and Health model. The testlets were treated as 'superitems' in the subsequent analyses.

Unidimensionality was confirmed using the principal component analysis and t-test-based method proposed by Smith.[22] Two sets of items hypothesised to represent low levels and high levels of quality of life were defined, based on the correlation between items and the first residual factor. An independent t-test was then used to compare the difference in these estimates for each person. Unidimensionality was confirmed if ≤5% of the t-tests were significant or if lower bound of a binomial 95% CI of the observed proportion overlapped 5%.[19 22]

Cross-cultural (measurement) equivalence was tested using the differential item functioning (DIF) analysis feature in-built into RUMM2030. This is based on a two-way analysis of variance (ANOVA) of residuals across each level of person factor (in this case, culture) and across different levels of trait (in this case, quality of life). Presence of uniform DIF was suggested if the P value of the main effects (culture) was significant. This test flags the presence of significant DIF in the pooled datasets (significant difference between two or more group means) but does not specify where the difference lies. The post hoc Tukey test which performs a pairwise comparison of means was used to explore DIF patterns and identify which country-specific dataset(s) exhibited the DIF. Once identified, the testlet affected by cross-cultural DIF was 'split' into two, to provide a culturally specific (emic) testlet for the country exhibiting the DIF and a culturally general (etic) testlet for the rest of the countries. Once the DIF-affected testlet was split, the pooled data were reanalysed to assess fit to the model. This method of post hoc DIF analysis is detailed elsewhere.[23–26]

When fit to the model was established, the raw SScQoL scores were mapped against the corresponding Rasch-transformed (logit-based) scores and were linearly transformed to calibrate an interval scale of the same range. This allows for transformation of raw scores to interval scaling.[27] The raw-to-linear score conversion table provided the adjustment for the cross-cultural difference via the split testlet.[23–25 27]

RESULTS

Cross-cultural adaptation

The adaptation of the SScQoL into European languages was largely seamless except for the German dataset in which patients had reported problems in providing strictly dichotomous 'yes/no' responses on the following 10 items: (Q4) my condition makes me angry; (Q9) my condition means I have disturbed sleep; (Q11) it has affected the health of people around me; (Q12) my hands do not work as well as they did; (Q13) it puts a strain on my personal relationships; (Q15) any sort of activity is difficult; (Q19) I cannot cope at all; (Q20) sleeping badly has affected me a lot; (Q25) I struggle to wash myself as I would like; (Q27) I feel helpless and (Q29) I miss being able to sort things out. In Sweden, patients reported problems with two items: with regard to (Q5) 'I get upset when I cannot do things' they preferred using 'disappointed' or 'sad' instead of 'upset' and for (Q10) 'it has affected me a lot socially', participants suggested to remove 'a lot'.

For the Spanish translation, in item Q27 (I feel helpless), the translators had difficulties in finding a word that captured the English meaning of 'helpless' ('impotencia' in Spanish). A consensus was reached among the translators that the Spanish word 'impotencia' which means 'impotence' in English had the closest meaning to 'helpless/powerlessness'. Since 'impotencia' also means sexual dysfunction, translators recommended that clarification should be provided to the patients when the questionnaire is issued to avoid confusion.

Table 1 Sample characteristics by country

	Sample	Gender		Age	
Country	N	M (%)	F (%)	Mean	SD
UK	121	15 (12.40)	106 (87.60)	57.09	12.073
France	115	18 (15.65)	97 (84.35)	59.05	13.226
Italy	131	16 (12.31)	114 (87.69)	57.96	15.031
Sweden	102	9 (8.74)	94 (91.26)	60.01	12.332
Germany	274	27 (9.90)	239 (87.20)	60.84	10.569
Poland	231	33 (14.29)	198 (85.71)	55.85	12.552
Spain	106	19 (17.92)	87 (82.08)	54.84	13.971
Pooled	1080	137 (12.69)	943 (87.31)	57.95	13.894

Validation

Patient characteristics

In total, 1080 patients were recruited and their age and gender distribution parameters are summarised in table 1.

Fit to the Rasch model

Table 2 presents item–person fit statistics reliability and unidimensionality of the SScQoL for individual countries' datasets. The initial analyses of the 29-item scales for each country (based on individual items, table 2A) suggest an initial lack of fit for the German, Italian, Polish and Swedish data (values representing a perfect fit to the model are given in the lowest row of table 2). Individual item fit statistics for each country are provided in the online supplementary table S1 . Assessment of the residual correlation matrix revealed significant local dependence (item–item residual correlation >0.3), which was largely responsible for the lack of fit.

The 29 items of the full scale were the mapped by consensus by the project leaders (MN and ACR) onto five domains corresponding to the components of the International Classification of Functioning, Disability and Health model: function (activity limitation), emotional (personal factors), sleep (personal factors), social (participation restrictions) and pain (impairment) (see table 3).

Using the 5-testlet model, the responses for six countries (France, Italy, Poland, Spain, Sweden and the UK) showed a good fit to the Rasch model confirming construct validity, reliability and unidimensionality in the country-specific data (see table 2B). The German data continued to exhibit significant deviations from the Rasch model (item residual mean −0.698, SD 2.795, item–trait interaction χ^2 P<0.001). The datasets for the six countries that had evidence of fit to the Rasch model (France, Italy, Poland, Spain, Sweden and the UK) were combined in a pooled analysis and the results suggested that each testlet had an acceptable fit to the Rasch model (see table 2B). However, the item–trait (testlet–trait) interaction χ^2 statistic for the pooled dataset continued to display significant deviation from the model expectations (χ^2=63.909, df=45, P=0.034) suggesting lack of invariance (presence of DIF) across different levels of quality of life.

Cross-cultural invariance

DIF analysis highlighted a significant cross-cultural bias in the social subscale (table 4). Post hoc Tukey analysis revealed that the DIF was displayed by the Italian dataset. The social subscale

Table 2 Item–person fit statistics, reliability and unidimensionality of the SScQoL (country-specific data)

	Item fit residual		Person fit residual		Item–trait χ^2 interaction		Person separation reliability		Unidimensionality test*
Country	Mean	SD	Mean	SD	Value (DF)	P value	PSI	N	Proportion of independent t-tests (binomial 95% CI)
A: Analysis of the SScQoL with individual 29 items									
UK	−0.232	1.043	−0.233	0.680	48.932 (29)	0.012	0.919	112	0.099 (0.060 to 0.138)
France	−0.291	1.048	−0.271	0.855	34.868 (29)	0.209	0.893	111	0.087 (0.047 to 0.127)
Germany	−0.704	2.034	−0.484	1.456	398 (116)	**<0.001**	0.881	263	0.063 (0.028 to 0.099)
Italy	−0.205	1.051	−0.285	0.756	88.662 (58)	**0.006**	0.890	125	0.053 (0.016 to 0.091)
Poland	−0.520	1.380	−0.342	0.950	187.494 (116)	**<0.001**	0.902	221	0.099 (0.060 to 0.138)
Spain	−0.166	0.751	−0.220	0.654	42.439 (29)	0.051	0.906	95	0.075 (0.033 to 0.116)
Sweden	−0.264	0.910	−0.273	0.728	60.886 (29)	**<0.001**	0.892	101	0.068 (0.026 to 0.110)
B: Analysis of the SScQoL as a 5-testlet scale									
UK	0.016	1.535	−0.288	0.848	4.992 (5)	0.417	0.896	109	0.050 (0.011 to 0.088)
France	0.020	1.380	−0.280	0.895	2.368 (5)	0.796	0.826	106	0.028 (-0.013 to 0.068)
Germany	−0.698	2.795	−0.285	0.873	61.952 (20)	**<0.001**	0.852	263	0.030 (-0.013 to 0.073)
Italy	−0.137	1.888	−0.315	0.831	4.281 (5)	0.510	0.818	125	0.038 (0.001 to 0.075)
Poland	−0.348	2.139	−0.308	0.966	22.450 (15)	0.096	0.853	221	0.043 (0.015 to 0.071)
Spain	−0.265	1.346	−0.221	0.711	11.430 (5)	0.043	0.846	92	0.018 (-0.023 to 0.060)
Sweden	−0.136	0.793	−0.253	0.819	25.665 (44)	0.988	0.813	95	0.030 (-0.013 to 0.073)
Perfect fit	0	1	0	1		>0.05	>0.70		≤0.05 or lower-bound 95%CI≤0.05

P value, χ^2 probability, where >0.05 (>0.01 for Bonferroni correction) suggest adequate fit to the model.
*Unidimensionality was deemed supported if ≤5% (0.05) of independent t-tests were significant or if lower bound of a binomial 95% CI of the observed proportion overlapped 5% (0.05).
DF, degree of freedom; PSI, Person Separation Index (internal consistency) reliability; SScQoL, Systemic Sclerosis Quality of Life Questionnaire.

was therefore 'split' such that there was a social-etic subscale which is culturally general (for five countries—France, Poland, Spain, Sweden and the UK) and a social-emic subscale which was culturally specific to Italy. This split improved the overall fit statistics of the pooled data (see the online supplementary table S2). The subsequent item–trait χ^2 statistic suggested adequate fit to the model (χ^2=65.580, df=54, P=0.140) and the reliability remained good (PSI=0.841).

Calibrating an interval scale

Following DIF analysis and the adjustment for cross-cultural DIF, the raw scale scores were transformed into logit-based (interval level) scores for the five testlets, with an adjusted social subscale for Italy (see table 5).

DISCUSSION

The original SScQoL was developed with patients to ensure it captures HRQoL aspects that are of interest to patients. Having satisfied the requirements of the Rasch model expectations, the tool has demonstrated validity, reliability and statistical sufficiency.[7] In this study, a new UK dataset was collected and the conclusions were consistent with those of the original development study. Additionally, we have employed a standardised method of questionnaire adaptation into the European languages (and associated cultures), ensuring that the tool maintains a conceptual equivalence between the original and the adapted (translated) versions. Furthermore, in this analysis, we have transposed the concept of SScQoL into five subscale measures, where each item contributes to the subscale and the overall dimension. Therefore, the subscale score provides the estimate of quality of life specific for that domain (function, emotional, sleep and so on) and the total score provides a sufficient statistic for overall health-related quality of life in SSc.

Table 3 Testlets formed by grouping items

Testlet	Number of items	Items
Function	6	1, 12, 14, 15, 22 and 25
Emotional	13	2, 3, 4, 5, 6, 7, 8, 17, 18, 19, 24, 27 and 29
Sleep	2	9 and 20
Social	6	10, 11, 13, 16, 21 and 23
Pain	2	26 and 28

Rasch analyses confirmed measurement equivalence between the English and all adapted versions except the German, where patients had found it difficult to complete some dichotomous items, indicating a preference for rating scales or having more options. This implies that for those patients, the dichotomous items, as presented, failed to capture the full range of their responses. As the SScQoL is a needs-based measure, failure of the tool to measure the full range of patient responses appears to have had an impact on the validity of the German version of the scale as reflected in the lack of fit to the model. Future work on this specific version will explore the wording and presentation of the root questions, as well as the potential for, and impact of making available trichotomous or higher level response options.

All translations, except the German version, demonstrated sufficient validity and fit to the Rasch model to support their use for single-country studies and within-country comparisons. The current analysis confirmed the unidimensionality and measurement equivalence in five of the six countries when evaluated across five domains. Good internal validity and reliability ensures that clinicians (and patients) can use the tool with confidence when evaluating HRQoL at person and at population levels. Using the SScQoL alongside other outcome measures ensures that quality of life is being taken into account when providing care.

Since comparable scales are now available for six of the seven languages/cultures, employing the logit-based transformed scores as summarised in table 4 enables accurate estimation of SScQoL as an interval level measure, as well as comparability across cultures.

Table 4 Results of DIF analysis in pooled data (France, Italy, Poland, Spain, Sweden and UK)

	Main effects: country (uniform DIF)				Interaction effects: class interval by country (non-uniform DIF)			
Testlet	Mean square	F-statistic	Degrees of freedom	P value*	Mean square	F-statistic	Degrees of freedom	P value*
Function	1.736	2.711	5	0.030	1.033	1.613	36	0.015
Emotional	1.777	3.297	5	0.011	0.574	1.065	36	0.371
Sleep	3.832	3.264	5	0.012	0.961	0.818	36	0.766
Social	9.603	15.839	5	**<0.001**	0.706	1.165	36	0.240
Pain	1.627	1.680	5	0.153	0.978	1.010	36	0.455

*Significant Bonferroni adjusted P value ≤0.003 suggest presence of uniform DIF.
DIF, differential item functioning.

This means that it is now possible to pool large datasets across countries and/or to develop collaborative projects using a common measure of SScQoL. It is recommended therefore that future studies report subscale scores routinely, as well as reporting single-scale scores, to facilitate comparison and data pooling across countries. As further work will be required to explore how the German SScQoL works in other samples, caution will be required until this work is complete when comparing between German scores and scores from other countries.

This study sets out to establish a common measure of QoL in SSc across seven countries. The study has two main limitations. First, the cross-cultural validation of the SScQoL in Germany did not work as expected. Two subsets of data were collected for this analysis and these datasets, both individually and in pooled

Table 5 Conversion table for raw-to-linear (Rasch transformed) scores with cross-cultural adjustment

Raw scores (yes=1, no=0)	Function (all)	Emotional (all)	Sleep (all)	Social (Italy)	Social (others)	Pain (all)	Total (Italy)	Total (others)
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1.0	1.3	1.6	1.0	0.8	1.6	1.0	2.9	3.6
2.0	2.4	2.8	2.0	1.4	2.6	2.0	4.8	6.0
3.0	3.1	3.7		1.8	3.3		6.0	7.5
4.0	3.8	4.5		2.4	3.9		6.9	8.6
5.0	4.8	5.2		3.7	4.7		7.7	9.6
6.0	6.0	5.9		6.0	6.0		8.3	10.3
7.0		6.6					8.9	11.0
8.0		7.3					9.4	11.6
9.0		8.0					9.9	12.2
10.0		8.9					10.3	12.8
11.0		9.9					10.8	13.3
12.0		11.2					11.2	13.8
13.0		13.0					11.6	14.3
14.0							12.1	14.8
15.0							12.5	15.3
16.0							12.9	15.8
17.0							13.4	16.3
18.0							13.9	16.8
19.0							14.3	17.3
20.0							14.8	17.8
21.0							15.3	18.4
22.0							15.9	19.0
23.0							16.5	19.7
24.0							17.1	20.4
25.0							18.0	21.3
26.0							19.0	22.4
27.0							20.6	23.8
28.0							23.6	25.9
29.0							29.0	29.0

The SScQoL has dichotomous yes/no responses, coded as 1 (yes) and 0 (no), yielding a scoring range 0–6 for the function subscale, 0–13 for the emotional subscale, and so on. The scores obtained from the patient are the raw scores and these must be converted to linear scores using the conversion chart. For example, if a patient has a raw score of 2 on the functional subscale, this will be transformed to 2.4, if the patient has a raw scores of 3 on the emotional subscale this will transformed to 3.7, and so on in the other subscales. The social subscale is split, with transformed scores for Italy and the rest of the countries. If a patient from Italy has a raw score of 4 on the social subscale, this will be transformed to 2.4, but a raw score of 4 from patients in other countries will be transformed to 3.9. Adding up all the transformed subscale scores gives the total SScQoL score which is a comparable estimate of the patient's quality of life (range 0–29), higher scores indicating a worse quality of life.
Others=France, Poland, Spain, Sweden and the UK.
SScQoL, Systemic Sclerosis Quality of Life Questionnaire.

form, showed lack of fit the model. Further work is required to explore different ways of formatting the items in such a way that a full range of patient responses will be better captured. Second, ethics committees in some countries permitted collecting only basic demographic details (age and gender) in addition to the SScQoL items, and this may have limited the factors or subgroups being tested for invariance. Third, being a cross-sectional study, this study did not assess the sensitivity to change of the adapted versions. Sensitivity to change was established for the original (English) version, and it is expected that the adapted versions will also demonstrate this. Further research will be required to determine the minimal clinically important difference to support measurement of the impact or of treatments on the quality of life in people with SSc. As result of the successful cross-cultural validation of the SScQoL into six different European countries, we recommend for this tool to be translated into more European languages and to be adopted as part of a core set of tools used in SSc observational and clinical trials studies. An implementation phase, working in combination with colleagues within the EUSTAR network and beyond, is required to move towards a more systematic approach to clinical data capture in SSc research.

CONCLUSION

The individual SScQoL items have translated well into five European languages and overall, the scale maintained its construct validity, working well as a five-subscale questionnaire. Using the logit-based transformed scores, measures of quality of life in SSc can be directly compared across five countries (France, Poland, Spain, Sweden and UK). Data from Italy are also comparable with the other five countries using a separate adjusted scale, which sufficiently recalibrates the scores in the social subscale, so as to allow a valid comparison across countries. While comparison between German scores and the other countries will need further testing, it is likely that this can be accomplished with some extra work and in the interim, this study has provided a common measure of quality of life in people with SSc across six European countries. Future work will be required to define the thresholds of health-related quality of life and clinically meaningful change in SSc and to further adapt the SScQoL into a wider range of languages and cultural settings. Different versions of the SScQoL can be obtained at: https://doi.org/10.5518/325.[28]

Author affiliations
[1]Academic Rheumatology Unit, University Hospitals Bristol NHS Foundation Trust, Bristol, UK
[2]Department of Nursing and Midwifery, University of the West of England, Bristol, UK
[3]Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds, UK
[4]NIHR Leeds Biomedical Research Centre, Leeds, UK
[5]Rheumatology A Department, Cochin Hospital, Paris Descartes University, Paris, France
[6]Department of Rheumatology and Clinical Immunology, Justus-Liebig University Giessen, Bad Nauheim, Germany
[7]Rheumatology Department, Hospital General Hospitalet-Moisès Broggi, Sant Joan Despí, Spain
[8]Department of Rheumatology, Lund University, Lund, Sweden
[9]Department of Experimental Medicine, University of Florence, Florence, Italy
[10]Department of Geriatric Medicine, Division of Rheumatology, Azienda Ospedaliero-Universitaria Careggi (AOUC), Florence, Italy
[11]Department of Rheumatology and Immunology, St. Josef Hospital, Wuppertal, Germany
[12]Department of Integrated Medical Care, Medical University of Bialystok, Bialystok, Poland
[13]Department of Foreign Languages, Medical University of Bialystok, Bialystok, Poland
[14]Department of Rheumatology and Internal Diseases, Medical University of Bialystok, Bialystok, Poland

Acknowledgements This was a collaborative work and the authors wish to thank the following people for their help in making the work possible: Dr Naomi Chapman (Reay) of the Royal College of Nursing, who is the early developer of the SScQoL for providing us with the permission to use the SScQoL; Mr Freddie Lewery from the School of Medicine, University of Leeds for helping with data input; Dr Mike Horton (director of the psychometric laboratory for health sciences, university to Leeds) for technical advice on psychometric testing. The authors also wish to thank the following people for their help in data collection in Poland: Marzena Olesińska (Warsaw), Marek Brzosko (Szczecin), Piotr Leszczyński (Poznań), Katarzyna Pawlak-Buś (Poznań), Bogdan Batko(Krakow), Piotr Wiland (Wroclaw), Maria Majdan (Lublin), Małgorzata Bykowska-Sochacka (Sopot), Wojciech Romanowski (Śrem), Aleksandra Zon-Giebel (Ustroń) and Sławomir Jeka (Bydgoszcz).

Contributors AR (professor of clinical biomechanics) is the principal investigator, designed the study, led the grant application, oversaw the project and undertook the statistical analyses, interpretation of the results and revised the study report for intellectual content. MN (senior lecturer in rheumatology nursing) codesigned the study with AR and BA-P, drafted the statistical analysis plan, contributed to the grant application, coordinated the project, undertook the statistical analyses, interpretation of the results and drafted the study report and revised it for intellectual content. BA-P (clinical post doctoral research fellow) was a member of the study team, a co-applicant on the study grant, contributed to the drafting of the manuscripts and revised it for intellectual content. YA (professor of rheumatology, rheumatologist) led the study team in France, contributed to the drafting of the study report and revised it for intellectual content. FDG (associate professor, head of scleroderma programme, rheumatologist) was a member of the study team, a co-applicant on the study grant, contributed to the drafting of the manuscripts and revised it for intellectual content. MF (resident physician and biostatistician) was a member of the study team in Germany, contributed to the translation of the SScQoL, data collection and drafting and revising the manuscripts for intellectual content. SG-D (rheumatology nurse specialist) led the study team in Spain, contributed to the drafting of the manuscripts and revised it for intellectual content. RH (associate professor of rheumatologist) was a member of the study team in Sweden, contributed to the drafting of the manuscripts and revised it for intellectual content. CK (physician assistant) was a member of the study team in Germany, contributed to the drafting of the manuscripts and revised it for intellectual content. MM-C (professor of rheumatology and medicine) led the study team in Italy, contributed to the drafting of the manuscripts and revised it for intellectual content. UM-L (professor of rheumatology) led the study team in Germany, contributed to the drafting of the study report and revised it for intellectual content. GS (associate professor, rheumatology occupational therapist) led the study team in Sweden, contributed to the drafting of the manuscripts and revised it for intellectual content. VT-S (consultant rheumatologist) was a member of the study team in Spain, contributed to the drafting of the manuscripts and revised it for intellectual content. TS (consultant rheumatologist) was a member of the study team in Germany, responsible for data collection in Germany. He contributed to the drafting of the manuscripts and revised it for intellectual content. MS (senior lecturer in nursing) led the study team, adaptation of the SScQoL and data collection in Poland and contributed to the drafting of the study report and revised it for intellectual content. JS (english instructor) was a member of the study team in Poland, responsible for the translation process of the SScQoL, contributed to the drafting of the study report and revised it for intellectual content. SS (professor of rheumatology) was a member of the study team in Poland, responsible for data collection for the validation of the SScQoL, contributed to the drafting of the study report and revised it for intellectual content. All authors read and approved the final version.

Funding MN and AR declare financial support (grant) for the submitted work from European League Against Rheumatism (grant reference HPR019).

Disclaimer The funder was not involved in the preparation of the study protocol, running of the study, and analysis or preparation of the report. There are no other relationships or activities that could appear to have influenced the submitted work.

Competing interests All authors have completed the unified competing interests form at http://www.icmje.org/coi_disclosure.pdf (available from the corresponding author) . The corresponding author had full access to all the study data and shares the final responsibility for publication with all coauthors.

Patient consent Detail has been removed from this case description/these case descriptions to ensure anonymity. The editors and reviewers have seen the detailed information available and are satisfied that the information backs up the case the authors are making.

Ethics approval This was a multicentre study conducted in seven countries (Germany, France, Italy, Poland, Spain, Sweden and UK) and was approved by local ethical committees or institutional review boards in each participating country.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Johnson SR, Fransen J, Khanna D, *et al*. Validation of potential classification criteria for systemic sclerosis. *Arthritis Care Res* 2012;64:358–67.
2 Abraham DJ, Varga J. Scleroderma: from cell and molecular mechanisms to disease models. *Trends Immunol* 2005;26:587–95.
3 Geyer M, Müller-Ladner U. The pathogenesis of systemic sclerosis revisited. *Clin Rev Allergy Immunol* 2011;40:92–103.
4 Strickland G, Pauling J, Cavill C, *et al*. Predictors of health-related quality of life and fatigue in systemic sclerosis: evaluation of the EuroQol-5D and FACIT-F assessment tools. *Clin Rheumatol* 2012;31:1215–22.
5 Gualtierotti R, Scalone L, Ingegnoli F, *et al*. Health related quality of life assessment in patients with systemic sclerosis. *Reumatismo* 2010;62:210–4.
6 Hudson M, Thombs BD, Steele R, *et al*. Health-related quality of life in systemic sclerosis: a systematic review. *Arthritis Rheum* 2009;61:1112–20.
7 Reay N. *The quality of life in patients with diffuse and limited systemic sclerosis [Monograph*: University of Leeds, 2008.
8 McKenna SP, Doward LC. The needannrheumdis-2017-212412s-based approach to quality of life assessment. *Value Health* 2004;7(s1):S1–3.
9 Maslow AH, Frager R, Fadiman J, *et al*; *Motivation and personality*. Harper & Row New York, 1970.
10 Tijhuis GJ, de Jong Z, Zwinderman AH, *et al*. The validity of the Rheumatoid Arthritis Quality of Life (RAQoL) questionnaire. *Rheumatology* 2001;40:1112–9.
11 Keenan AM, McKenna SP, Doward LC, *et al*. Development and validation of a needs-based quality of life instrument for osteoarthritis. *Arthritis Rheum* 2008;59:841–8.
12 Doward LC, Spoorenberg A, Cook SA, *et al*. Development of the ASQoL: a quality of life instrument specific to ankylosing spondylitis. *Ann Rheum Dis* 2003;62:20–6.
13 Allcock RJ, Forrest I, Corris PA, *et al*. A study of the prevalence of systemic sclerosis in northeast England. *Rheumatology* 2004;43:596–602.
14 Beaton DE, Bombardier C, Guillemin F, *et al*. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 2000;25:3186–91.
15 Masi AT. Subcommittee For Scleroderma Criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Preliminary criteria for the classification of systemic sclerosis (scleroderma). *Arthritis & Rheumatism* 1980;23:581–90.
16 Bond TG, Fox CM. *Applying the Rasch model: Fundamental measurement in the human sciences*. London: Lawrence Arlbaum Associates, 2001.
17 Rosenbaum PR. Criterion-related construct validity. *Psychometrika* 1989;54–625–33.
18 Andersen EB. Sufficient statistics and latent trait models. *Psychometrika* 1977;42:69–81.
19 Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Rheum* 2007;57:1358–62.
20 Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. *BMJ* 1995;310:170.
21 Lord FM, Novick MR, Birnbaum A. *Statistical theories of mental test scores*, 1968.
22 Smith EV. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *J Appl Meas* 2002;3:205–31.
23 Tennant A, Penta M, Tesio L, *et al*. Assessing and adjusting for cross-cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model: the PRO-ESOR project. *Med Care* 2004;42:I–37.
24 Lange R, Thalbourne MA, Houran J, *et al*. Depressive Response Sets due to gender and culture-based Differential Item Functioning. *Pers Individ Dif* 2002;33:937–54.
25 Hagquist C, Andrich D. Recent advances in analysis of differential item functioning in health research using the Rasch model. *Health Qual Life Outcomes* 2017;15:181.
26 Ndosi M, Bremander A, Hamnes B, *et al*. Validation of the educational needs assessment tool as a generic instrument for rheumatic diseases in seven European countries. *Ann Rheum Dis* 2014;73:2122–9.
27 Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. *Arch Phys Med Rehabil* 1989;70:857–60.
28 Redmond A, Tennant A, Horton M, *et al*. *Systemic Sclerosis Quality of Life (SScQoL) instrument: translation and cross cultural validation across seven European countries (UK, France, Italy, Spain, Sweden, Germany and Poland)*. University of Leeds, 2018. [Dataset].