JOURNAL OF APPLIED MEASUREMENT, *4*(3), 249–257

Rasch Measurement in the Assessment of Amytrophic Lateral Sclerosis Patients

Josephine M. Norquist [1]

Ray Fitzpatrick [1]

Crispin Jenkinson [2,3]

[1] Department of Public Health, Institute of Health Sciences, University of Oxford

[2] Department of Public Health, Health Services Research Unit, University of Oxford

[3] Picker Institute Europe, Oxford

This paper examines the sensitivity to change over time of the Amyotrophic Lateral Sclerosis Assessment Questionnaire (ALSAQ-40). Individuals' health status change was assessed by means of the Rasch-based Reliable Change Index (RCI) for ALSAQ-40 questionnaires completed on two occasions, three months apart. In addition, at follow-up respondents indicated how much change they had experienced since baseline via dimension-specific self reported transition questions. 764 individuals returned questionnaires at baseline and follow-up. For all dimensions, of respondents defined by the RCI as worse, a majority rated themselves as worse. However, on two dimensions over 60% of the respondents who rated themselves as being worse were defined as unchanged by the RCI. As with effect size smaller RCI cut-off points might be needed for subjects with ALS. This study confirms that the ALSAQ-40 is a valid and responsive disease specific health related quality of life instrument for use in studies of patients with ALS or other motor neuron diseases.

Requests for reprints should be sent to Josephine M. Norquist, Department of Public Health, Institute of Health Sciences, University of Oxford, Old Road, Headington, Oxford, OX3 7LF (UK), e-mail: josephine.norquist@dphpc.ox.ac.uk.

Introduction

Amyotrophic lateral sclerosis (ALS) is a progressive neurological disease that attacks the nerve cells (neurons) that control voluntary muscles. This disease belongs to a group of neuro-degenerative disorders (e.g., muscular dystrophy, multiple sclerosis, and Parkinson's disease) that are characterised by the gradual degeneration and death of motor neurons. New treatments are emerging that may retard progression of the disease and prolong survival (Benzimon, et al., 1994). Therefore it is important to have adequate health related quality of life instruments to measure patients' health status and change in health status over time.

The 40 item Amyotrophic Lateral Sclerosis Assessment Questionnaire (ALSAQ-40) is a disease specific health related quality of life instrument for use in studies of patients with ALS or other motor neuron diseases (MND) (Jenkinson, et al., 1999a; Kiebert, et al., 2001; Jenkinson, et al., 1999b; Jenkinson, et al., 2000; Jenkinson, et al., 2001). The content of the measure was designed on the basis of patient self-report. The instrument covers five dimensions of health status that are affected by the disease: physical mobility, activities of daily living and independence, eating and drinking, communication, and emotional functioning. The questionnaire addresses experiences of importance to individuals with ALS/MND in such diverse areas as fear of falling when walking, difficulties cutting and eating food, difficulties participating in conversations, feelings of isolation, social embarrassment, as well as measuring feelings of fear and hopelessness about the future, that are all quite distinctively associated with ALS/MND disease. Dimension scores are coded on a scale of 0 (perfect health as assessed by the measure) to 100 (worse health as assessed by the measure).

The purpose of this paper is to determine whether the ALSAQ-40 is sensitive to change over time in individuals with ALS. In addition to examining the overall changes over time, sensitivity to change was assessed by examining the patterns of change scores observed in respondents' retrospective self-reporting of health status change via transition questions (Feinstein and Wells, 1977; Tugwell, et al., 1987; Fitzpatrick, et al., 1993a). Transition questions address the extent and direction of change since a previous assessment in a short and simple manner. These questions are often used to detect the minimal clinically important difference, which is the difference in score in the dimension of interest that patients perceive as beneficial (Juniper, et al., 1994). It has been argued that transition questions may be more sensitive to change than conventional change scores (Fitzpatrick, et al., 1993b). Here transition questions are used as benchmarks against which to assess evidence of change provided by the ALSAQ-40.

The sensitivity to change of the ALSAQ-40 was also examined by applying the Rasch model to the five dimensions of the instrument and by using the results to calculate the Reliable Change Index (RCI) (Jacobson and Truax, 1992). It has been argued that the Rasch model, because of improved measurement along the continuum of underlying constructs, may offer more accurate measurement of change over time in health status and health-related quality of life (McHorney, et al., 1997; Wolfe, 2001). Calculation of the RCI is a method used to assess the relevance of statistically significant changes. In this paper the ALSAQ-40 dimension specific RCIs are calculated using Rasch results and are used as a method to determine the instrument's ability to detect change in health status.

Methods

Subjects

A postal survey was conducted of all members of the Motor Neurone Disease Association for England, Wales and Northern Ireland. Questionnaires were completed on two occasions, three months apart. At baseline, the survey contained the ALSAQ-40, demographic questions, and, in addition, members were asked to provide their name and address if they were willing to take part in the follow up survey. At follow up, the survey was sent out to all respondents who had provided

their name and address at baseline. In this second survey the questionnaire contained the ALSAQ-40 and five transition questions. The transition questions asked respondents to judge change since the previous survey on each of the specific dimensions of the ALSAQ-40. For each of the transition questions respondents were asked if they were 'better', 'about the same', 'a little worse' or 'much worse' than three months ago.

Data Analyses

Before proceeding with the calculation of the dimension specific RCIs, the Rasch measurement approach was applied on all ALSAQ-40 items to confirm the structure of the questionnaire defined in a previous study through factor analysis (Jenkinson, et al., 2001) and to test the psychometric properties of the instrument. The Rasch Rating Scale model was applied throughout these analyses (Wright and Masters, 1982). The structure of the instrument was evaluated through a principal component analysis of the standardized residuals. The invariance of the item difficulties was established by plotting the baseline estimates against those obtained at follow-up including confidence bands based on the error of calibration (Wright and Masters, 1982).

Rasch analyses results were then used to assess respondents' change in health status through the calculation of the Reliable Change Index (RCI) (Jacobson and Truax, 1992). The RCI allows the assessment of the clinical relevance of statistically significant change (de Bruin, et al., 1997). Numerically the RCI represents the number of scale points needed on a given psychometric measure to determine if a change in score from pre- to post- treatment is due to real change or chance variation (Ferguson, et al., 2002). Jacobson and Truax suggested that when RCI values are greater than 1.96, it is unlikely ($p <$ 0.05) that the post-test score is not reflecting real change. Prieto, et al., applied a procedure equivalent to the one proposed by Jacobson and Truax to assess growth hormone deficiency in adult patients using the Rasch measurement approach to calculate the RCI statistic (Prieto, et al., 2001).

In this study to calculate the RCIs, person measures for each dimension of the ALSAQ-40 between baseline (Pb) and follow-up (Pf) were evaluated by standardizing their difference through its expected standard error,

$$\sqrt{SE_{Pb}^2 + SE_{Pb}^2}\,.$$

The standardized difference,

$$Z = P - P_f \sqrt{SE_{Pb}^2 + SE_{Pb}^2}\,,$$

has an expectation of zero and a variance of one. Conventionally, patients with $| z | < 2$ are defined as not having experienced any significant change in health status, patients with $z = -2$ as getting worse and patients with $z = 2$ as improved (Prieto, et al., 2001). To determine the appropriateness of the conventional cut-off points and to validate the transition questions, individuals were grouped according to whether they rated themselves as 'same', 'little worse' or 'worse' at follow-up. For each group effect size statistics and 95 % confidence intervals (C.I.) using both raw and Rasch measures were calculated and tabulated against the corresponding RCI values. Conventional cut-off points of 0.2 to 0.5 'small', 0.5 to 0.8 'moderate', and 0.8 or above 'large' were used for the effect size statistics (Cohen, 1988; Kazis, et al., 1989). Confidence intervals for effect sizes were obtained using the bootstrap algorithm (Efron and Tibshirani, 1993; Hastie, et al., 2001). A total of 1000 bootstrap samples with replacement were generated from each individual group. Effect size (ES) statistics were calculated for each resampling, which provided an estimate of the distribution for each ES. The 25th and 975th values of the ES distribution identified the 95% confidence interval.

Data were analysed using SPSS (Windows, version 8.0) for general descriptive statistics. For Rasch analyses the Winsteps program was used (Linacre and Wright, 2000). Confidence intervals for the ES estimates were obtained using STATA (Windows, version 7.0).

Results

Seven hundred and sixty four individuals returned both baseline and follow-up questionnaires. Of these 220 (28.8%) had been diagnosed with MND for a year or less while 470 (61.5%) had been diagnosed for 3 years or less. 492 (64.4%) were male, 657 (86.0%) lived with at least another adult, 35 (4.6%) lived in a nursing home, 604 (79.1%) were married or lived as married and 56 (7.3%) were currently in paid employment. The mean age at baseline for the entire sample was 64.4 (range 31.3-90.9, s.d. 11.4).

Table 1 summarises the raw score means (s.d.) for the two assessments and the mean differences between assessments for each ALSAQ-40 dimension. Except for the emotional functioning dimension, all other dimensions show a significant worsening at follow-up.

The frequencies of the responses to the dimension specific transition questions are reported in Table 2. Over 60% of the subjects responded as being 'a little worse' or 'much worse' when asked about changes in their physical ability and activity of daily living. For the 3 remaining dimensions over 40% responded as being 'a little worse' or 'much worse' compared to 3 months before the assessment.

To confirm the structure of the ALSAQ-40 questionnaire, all 40 items were combined into a single Rasch analysis followed by a principal component analysis of the standardized residuals. The results obtained from this analysis (available from the author) indicate that 5 distinct dimensions (physical mobility, activities of daily living/independence, eating and drinking, communication, and emotional functioning) can be obtained from the ALSAQ-40 questionnaire in agreement with previous analyses (Jenkinson, et al., 2001).

Table 1

Raw ALSAQ-40 dimension score means (s.d.). Patients with complete data (n=764)

ALSAQ-40 dimension	Mean (s.d.)		Mean diff [a]
	Baseline	Follow-up	
Physical Mobility	63.59 (29.47)	67.31 (28.75)	3.72 (17.34)*
ADL / Independence	67.15 (29.24)	70.94 (28.22)	3.79 (13.54)*
Eating and Drinking	40.42 (35.40)	45.23 (36.66)	4.54 (19.51)*
Communication	50.00 (39.38)	54.32 (39.43)	4.31 (14.55)*
Emotional Functioning	47.86 (24.64)	48.48 (26.18)	0.61 (15.51)

[a] Difference between follow-up and baseline scores.

* $p < 0.001$

Table 2

Transition question responses from patients with complete data

Compared to 3 months ago…	Response category				
	Better	About the same	A little worse	Much worse	Missing
how do you rate your physical ability?	8 (1.0%)	212 (27.7%)	282 (36.9%)	248 (32.5%)	14 (1.8%)
how do you rate your ability to do daily activities?	6 (0.8%)	267(34.9%)	225(29.5%)	247(32.3%)	19(2.5%)
how well do you manage to eat and drink?	11(1.4%)	372(48.7%)	189(24.7%)	174(22.8%)	18(2.4%)
how well can you communicate with others?	10(1.3%)	419(54.8%)	163(21.3%)	162(21.2%)	10(1.3%)
how would you rate your emotional state?	29(3.8%)	404(52.9%)	209(27.4%)	111(14.5%)	11(1.4%)

Table 3 lists item measures, standard errors, INFIT values, and separation and reliability indices for the Physical Mobility dimension at baseline and follow-up. Only one item (#7 'Pains in legs whilst walking') on this dimension shows misfit at both time points. The INFIT range of 0.6-1.4 suggested by Wright, et al., for 'rating scale' questionnaires was used to determine item fit (Wright, et al., 1994). The same analyses were performed for the remaining dimensions except for the eating and drinking dimension which is comprised of only 3 items and applying Rasch analysis was considered inappropriate. Each of the other dimensions had only one misfitting item at follow-up: #12 ('difficulty turning and moving in bed') for ADL/Independence, #30 ('felt self conscious about speech) for Communication and #33 ('felt embarrassed in social situations') for Emotional Functioning (data not shown). All three items had an INFIT value slightly above 1.4. For all dimensions item separation was > 10.00 except for the communication dimension (item sep = 5.02). Person separation was >3.00 at both baseline and follow-up across the four dimensions. All reliability indices were > 0.90. To determine item invariance over time baseline measures were plotted against the corresponding follow-up measures. Figure 1 presents the Physical Mobility dimension items with the 95% confidence bands. All items fall within the confidence bands. Same plots were created for the remaining dimensions and all items fall within the confidence bands indicating stability over time. For brevity only the Physical Mobility plot is shown.

The assessment of the relevance of statistically significant changes was obtained by using the Rasch results to calculate the RCI. The RCI values were then compared to the dimension specific transition question responses. Due to the low frequency of respondents rating themselves as

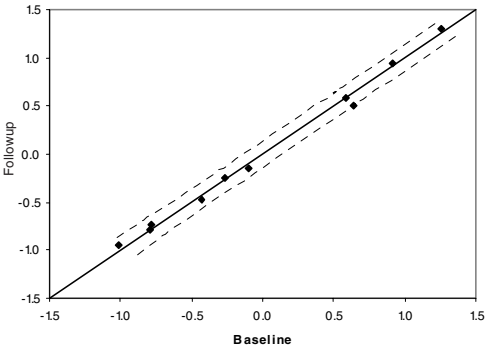

Figure 1. ALS Physical Mobility dimension item difficulties comparison

Table 3

Item measure, SE, and INFIT statistics ordered by baseline measures

Dimension / item no. / item description	Baseline			3 month follow-up		
	Measure	SE	INFIT MNSQ	Measure	SE	INFIT MNSQ
Physical Mobility						
2. Fallen whilst walking	1.26	0.05	1.00	1.31	0.05	0.96
7. Pains in legs whilst walking	0.91	0.05	1.83[a]	0.95	0.05	1.79[a]
4. Lost balance whilst walking	0.64	0.05	0.69	0.50	0.05	0.69
3. Stumbled or tripped whilst walking	0.59	0.05	0.76	0.58	0.05	0.75
9. Difficulty standing up	-0.10	0.05	0.96	-0.14	0.05	0.89
1. Difficulty walking short distances	-0.27	0.05	0.96	-0.25	0.05	1.03
10. Difficulty getting up out of chairs	-0.43	0.05	0.97	-0.47	0.05	1.03
6. Tired when walking	-0.78	0.05	1.03	-0.74	0.06	0.91
5. Concentrate when walking	-0.79	0.05	0.97	-0.79	0.06	0.99
8. Difficulty going up and down the stairs	-1.02	0.06	1.00	-0.94	0.06	1.10
	Separation		Reliability	Separation		Reliability
Item	14.59		0.99	14.28		0.99
Person	3.00		0.90	2.87		0.89

[a] INFIT outside reasonable range of 0.6–1.4

being 'better' after 3 months from the first assessment, this group of respondents was dropped from all comparisons. Table 4 cross-tabulates the sample grouped by RCI criteria and the transition question responses for each dimension. For all dimensions, of respondents defined by the statistical criteria of RCI as worse, a majority rated themselves as worse. Similarly for the stable respondents (defined by the statistical criteria) on two (Emotional Functioning and Communication) of the four dimensions a majority also rated themselves in this way. However, for the Physical Mobility and the ADL/Independence dimensions of respondents identified by RCI as stable, a majority rated themselves as 'a little worse' or 'worse'. This indicates that too many respondents are regarded as stable according to the RCI cut-off criteria.

Average RCI values, effect sizes and 95% CI were obtained for each transition response category to attempt to determine a more accurate cut-off point for subjects with ALS. Table 5 lists the RCI values and effect size values (calculated using both raw scores and Rasch measures) for each dimension divided into the three main

Table 4

Reliable Coefficient Index (RCI) summary by transition question per ALSAQ-40 dimension

Dimension	RCI criteria (z)	Transition question		
		Same	Little worse	Worse
Physical Mobility	Same (z > -2)	178 (30.8%)	228 (39.5%)	171 (29.6%)
	Worse (z ≤ -2)	8(10.8%)	25 (33.8%)	41(55.4%)
ADL/Independence	Same (z > -2)	276(39.6%)	191(30.8%)	184 (29.6%)
	Worse (z ≤ -2)	11(13.9%)	23(29.1%)	45(57.0%)
Emotional Functioning	Same (z > -2)	378(60.8%)	178(28.6%)	66(10.6%)
	Worse (z ≤ -2)	13(19.4%)	20(29.9%)	34(50.7%)
Communication	Same (z > -2)	385(60.5%)	132(20.8%)	119(18.7%)
	Worse (z ≤ -2)	14(18.9%)	28 (37.8%)	32(43.2%)

Table 5

Average RCI and Effect Size by ALSAQ-40 dimension and transition question response category

	RCI [95% C.I.]	Effect size* [95% C.I.]	Effect size** [95% C.I.]
Physical Mobility			
Same	-0.009 [-0.183; 0.164]	0.025 [-0.055; 0.099]	0.003 [-0.083; 0.072]
Little worse	0.288 [0.132; 0.445]	0.084 [0.022; 0.151]	0.124 [0.059; 0.189]
Worse	0.685 [0.485; 0.885]	0.310 [0.198; 0.424]	0.323 [0.227; 0.419]
ADL / Independence			
Same	0.078 [-0.219; 0.063]	0.012 [-0.061; 0.035]	-0.031 [-0.074; 0.011]
Little worse	0.543 [0.373; 0.713]	0.138 [0.061; 0.216]	0.213 [0.147; 0.283]
Worse	0.774 [0.599; 0.950]	0.281 [0.175; 0.385]	0.337 [0.260; 0.423]
Communication			
Same	0.124 [0.030; 0.218]	0.022 [-0.012; 0.052]	0.035 [0.006; 0.063]
Little worse	0.563 [0.336; 0.790]	0.144 [0.034; 0.240]	0.250 [0.161; 0.346]
Worse	0.782 [0.563; 0.999]	0.397 [0.268; 0.519]	0.389 [0.295; 0.477]
Emotional Functioning			
Same	-0.263 [-0.384; -0.141]	-0.114 [-0.181; -0.046]	-0.112 [-0.171; -0.052]
Little worse	0.245 [0.033; 0.450]	0.087 [-0.001; 0.188]	0.119 [0.013; 0.214]
Worse	0.981 [0.612; 1.350]	0.445 [0.242; 0.692]	0.456 [0.284; 0.646]

* Calculated dividing the mean difference by the s.d. at baseline using Rasch measures
** Calculated dividing the mean difference by the s.d. at baseline using Raw scores

transition response categories. In respondents rating themselves as 'worse' and with small but meaningful effect sizes in ALSAQ-40 (Jenkinson, et al., 2002), RCIs ranged from 0.68 to 0.98 according to dimension. Some statistics from Table 5 show overlapping confidence intervals (i.e. 'same' and 'little worse' for the physical mobility dimension, 'little worse' and 'worse' for the ADL dimension and the communication dimension) indicating a non-significant difference between the transition question responses. The emotional functioning dimension had all non-overlapping confidence interval.

Discussion

This paper provides further evidence that the ALSAQ-40 is comprised of five important and distinct dimensions of health-related quality of life. The results also show that the ALSAQ-40 instrument is sensitive to changes in health for individuals with ALS. Significant deterioration in health status was detected in just three months following the baseline assessment The pattern of change over time revealed by ALSAQ-40 indicates that, unlike other dimensions of health-related quality of life, emotional functioning did not deteriorate. This is consistent with other studies of patients with ALS using other instruments than ALSAQ-40, with emotional and psychological well-being being maintained whilst more physical dimensions deteriorate (Young, et al., 1995; Robbins, et al., 2001). Statistically significant differences in the amount of change were observed in the ALSAQ-40 dimensions in respondents reporting themselves as being the same compared to those rating themselves as being worse three months after baseline. An instrument able to distinguish between individuals who deteriorate from those who remain stable is very beneficial for future research in ALS patients.

The RCI method is a helpful method in the assessment of health status change over time. However, the results reported in Table 4 suggest that the conventional cut-off points of 2 suggested by Jacobson and Truax might be too stringent for subjects with ALS. On the Physical Mobility and the ADL/Independence dimensions over 60% of the respondents who rated themselves as being 'a little worse' or 'worse' were defined as unchanged by the RCI conventional cut-off point. Because ALS is normally associated with progressive deterioration, outcome measures are primarily needed that can distinguish subjects experiencing stability rather than deterioration since that will be the primary objective of interventions. Therefore, distinguishing ALS subjects who are stable from those who show even a little worsening over time is essential. Even though the conventional cut-off point may not be applicable in this context, the RCI concept can still be applied by empirically determining a more subjectively meaningful cut-off point (Jacobson and Truax, 1992). From the evidence of the current study RCI cut-off points between 0.5 and 1.00 may be more appropriate and consistent with respondents' judgements whether reflected in responses to transition questions or effect sizes of ALSAQ-40.

Conventionally effect size (ES) or standardized response mean (SRM) coefficients are used as a measure of an instrument's ability to detect change in health status. These coefficients are standardized by dividing the mean change in scores by the standard deviation of baseline scores (for the ES) or by the standard deviation of the changes (for the SRM). Calculation of these coefficients is performed on a single estimate of variability (baseline or change scores). Using these statistics may give an incomplete picture of an instrument's ability to detect clinical change at the individual patient level (Husted, et al., 2000). In contrast with traditional methods, the Rasch scoring method provides differential estimates of standard errors (SE) or measurement precision across the various levels of person measures. This quality is beneficial when the relevance of change in health is assessed through the RCI because this index can take into account the full range of SE. To date there is limited use of the RCI as a standardized and dimensionless representation of individual change in health status. Prieto, et al., (2001) used this index in the context of change in health in growth hormone deficiency patients. Their results lacked evidence of clinical relevance (or irrelevance) of change because no external indicator of change against which to make comparison was supplied. In the

current study responses to the dimension specific transition questions were used as a base of comparison between changes in health measured statistically by the RCI and subjectively by the respondents' own evaluation of change. Although some studies have suggested that retrospective methods of computing responsiveness yield little information about the ability of an instrument to detect treatment effects (Norman, et al., 1997; Guyatt, et al., 2002), it has also been argued that dimension specific transition questions are sensitive to changes experienced by patients (Fitzpatrick, et al., 1993a) and in some contexts more sensitive to change than conventional change scores (Fitzpatrick, et al., 1993b). In the present study transition questions were considered an appropriate method against which comparisons in changes in health can be made. Some transition response categories showed overlapping confidence interval indicating non-significant differences between these categories. However, no overlap occurred between 'same' and 'worse' indicating a significant difference between these response categories.

While the ALSAQ-40 was shown to be sensitive to changes over time in health-related quality of life, there were no meaningful differences obtained by Rasch compared with raw scoring of the instrument in this regard. Moreover while the RCI based on Rasch measures provides an additional and important approach to the evaluation of responsiveness of instruments, reliance on purely statistical criteria may result in missing patterns of change in health status that are important to patients.

References

Benzimon, J. T., Lacomblez, L., Meininger,V., and ALS/riluzole Study Group (1994). A controlled trial of riluzole in amyotrophic lateral sclerosis. *New England Journal of Medicine*, *330*, 585-591.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

de Bruin, A. F., Diederiks, J. P., de Witte, L. P., Stevens, F. C., and Philipsen, H. (1997). Assessing the responsiveness of a functional status measure: the Sickness Impact Profile versus the SIP68. *Journal of Clinical Epidemiology*, *50*, 529-540.

Efron, B. and Tibshirani, R. (1993). *An introduction to the bootstrap*. New York: Chapman and Hall.

Feinstein, A. and Wells, C. (1977). A new clinical taxonomy for rating change in functional activities of patients with angina. *American Heart Journal*, *93*, 172-182.

Ferguson, R. J., Robinson, A. B., and Splaine, M. (2002). Use of the Reliable Change Index to evaluate clinical significance in SF-36 outcomes. *Quality of Life Research*, *11*, 509-516.

Fitzpatrick, R., Ziebland, S., Jenkinson, C., and Mowat, A. (1993a). Transition questions to assess outcomes in rheumatoid arthritis. *British Journal of Rheumatology*, *32*, 807-811.

Fitzpatrick, R., Ziebland, S., Jenkinson, C., and Mowat, A. (1993b). A comparison of the sensitivity to change of several health status instruments in rheumatoid arthritis. *Journal of Rheumatology*, *20*, 429-436.

Guyatt, G., Norman, G., Juniper, E. F., and Griffith, L. E. (2002). A critical look at transition ratings. *Journal of Clinical Epidemiology*, *55*, 900-908.

Hastie, T., Tibshirani, R., and Friedman, J. H. (2001). *The elements of statistical learning*. New York: Springer.

Husted, J., Cook, R., Farewell, V. T., and Gladman, D. D. (2000). Methods for assessing responsiveness: a critical review and recommendations. *Journal of Clinical Epidemiology*, *53*, 459-468.

Jacobson, N. S., and Truax, P. (1992). *Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. Methodological issues and strategies in clinical research*. Washington, DC: American Psychological Association.

Jenkinson, C., Fitzpatrick, R., Brennan, C., Bromberg, M., and Swash, M. (1999a). De-

velopment and validation of a short measure of health status for individuals with amyotrophic lateral sclerosis/motor neurone disease: the ALSAQ-40. *Journal of Neurology*, *246* Suppl 3, III16-III21.

Jenkinson, C., Fitzpatrick, R., Brennan, C., and Swash, M. (1999b). Evidence for the validity and reliability of the ALS assessment questionnaire: The ALSAQ-40. *Amyotrophic Lateral Sclerosis and Other Motor Neuron Disorders*, *1*, 33-44.

Jenkinson, C., Fitzpatrick, R., Swash, M., and Levvy, G. (2001). *ALSAQ User Manual—Amyotrophic Lateral Sclerosis Assessment Questionnaire*. Oxford: Health Services Research Unit.

Jenkinson, C., Levvy, G., Fitzpatrick, R., and Garratt, A. (2000). The amyotrophic lateral sclerosis assessment questionnaire (ALSAQ-40): tests of data quality, score reliability and response rate in a survey of patients. *Journal of Neurological Sciences*, *180*, 94-100.

Jenkinson, C., Peto, V., Jones, G., and Fitzpatrick, R. (2003). Interpreting change scores on the Amyotrophic Lateral Sclerosis Assessment Questionnaire (ALSAQ-40). *Clinical Rehabilitation*, *71*, 380-385.

Juniper, E. F., Guyatt, G. H., Willan, A., and Griffith, L. E. (1994). Determining a minimal important change in a disease-specific Quality of Life Questionnaire. *Journal of Clinical Epidemiology*, *47*, 81-87.

Kazis, L. E., Anderson, J. J., and Meenan, R. F. (1989). Effect sizes for interpreting changes in health status. *Medical Care*, *27*, S178-S189.

Kiebert, G. W., Green, C., Murphy, C., Mitchell, J. D., O'Brien, M., Burrell, A., and Leighton, R. (2001). Patients' health-related quality of life and utilities associated with different stages of amyotrophic lateral sclerosis. *Journal of Neurological Sciences*, *191*, 87-93.

Linacre, J. M., and Wright, B. D. (2000). *A user's guide to WINSTEPS: Rasch model computer program.* Chicago: MESA Press.

McHorney, C. A., Haley, S. M., and Ware-JE, J. (1997). Evaluation of the MOS SF-36 Physical Functioning Scale (PF-10): II. Comparison of relative precision using Likert and Rasch scoring methods. *Journal of Clinical Epidemiology*, *50*, 451-461.

Norman, G. R., Stratford, P., and Regehr, G. (1997). Methodological problems in the retrospective computation of responsiveness to change: the lesson of Cronbach. *Journal of Clinical Epidemiology*, *50*, 869-879.

Prieto, L., Roset, M., and Badia, X. (2001). Rasch measurement in the Assessment of Growth Hormone Deficiency in adult patients. *Journal of Applied Measurement*, *2*, 48-64.

Robbins, R. A., Simmons, Z., Bremer, B. A., Walsh, S. M., and Fischer, S. (2001). Quality of life in ALS is maintained as physical function declines. *Neurology*, *56*, 442-444.

Tugwell, P., Bombardier, C., Buchanan, W. W., Goldsmith, C. H., Grace, E., and Hanna, B. (1987). The MACTAR Patient Preference Disability Questionnaire—An individualized functional priority approach for assessing improvement in physical disability in clinical trials in rheumatoid arthritis. *Journal of Rheumatology*, *14*, 446-451.

Wolfe, F. (2001). Which HAQ is best? A comparison of the HAQ, MHAQ and RA-HAQ, a difficult 8 item HAQ (DHAQ), and a rescored 20 item HAQ (HAQ20): analyses in 2,491 rheumatoid arthritis patients following leflunomide initiation. *Journal of Rheumatology*, *28*, 982-989.

Wright, B. D., Linacre, J. M., Gustafson, J.-E., and Martin-Lof, P. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, *8*, 370.

Wright, B. D., and Masters, G. N. (1982). *Rating scale analysis*. Chicago: MESA Press.

Young, C. A., Tedman, B. M., and Williams, I. R. (1995). Disease progression and perceptions of health in patients with motor neurone disease. *Journal of Neurological Sciences*, *129* Suppl, 50-53.