Measuring the impact of caring for a spouse with Alzheimer's disease: Validation of the Alzheimer’s Patient Partners Life Impact Questionnaire (APPLIQue)

Peter Hagell [1], Matthew Rouse [2], Stephen P. McKenna [3,4]

[1] The PRO-CARE Group, Faculty of Health Science, Kristianstad University, SE-291 88 Kristianstad, Sweden; Tel: +46 (0)44 250 3956; E-mail: Peter.Hagell@hkr.se
[2] Galen Research Ltd., B1 Chorlton Mill, 3 Cambridge Street, Manchester M1 5BY, UK; Tel: +44 (0)161 226 4446; E-mail: mrouse505@gmail.com
[3] Galen Research Ltd., B1 Chorlton Mill, 3 Cambridge Street, Manchester M1 5BY, UK; Tel: +44 (0)161 226 4446; E-mail: smckenna@galen-research.com
[4] School of Health Sciences, University of Manchester, Manchester, UK

As accepted for publication in the *Journal of Applied Measurement* on April 5, 2018.

Reference to final publication:
Hagell P, Rouse M, McKenna SP. Measuring the impact of caring for a spouse with Alzheimer's disease: Validation of the Alzheimer’s Patient Partners Life Impact Questionnaire (APPLIQue). *Journal of Applied Measurement* 2018; 19(3): 271-282.

Corresponding author:
Peter Hagell, PhD
Faculty of Health Science
Kristianstad University
SE-291 88 Kristianstad
Sweden
Tel: +46 44 250 3956
E-mail: Peter.Hagell@hkr.se

ABSTRACT

Alzheimer's disease (AD) is the most common form of dementia, characterized by cognitive, psychiatric and behavioral symptoms and increasing dependency. Family members typically assume increasing caregiving responsibilities, with considerable quality of life (QoL) impact. This article describes the testing of a needs-based QoL questionnaire for AD family caregivers. Initial analyses according to Rasch measurement theory suggested that items applied to spousal rather than non-spousal caregivers. Following removal of non-spousal responders, a 25-item questionnaire was identified that exhibited acceptable model fit, a mean (*SD*) person location of 0.194 (1.42) logits, residual correlations ≤0.173 and absence of DIF by age, gender or administration. Reliability was 0.85. This new measure, the Alzheimer's Patient Partners Life Impact Questionnaire (APPLIQue), may fill an important gap in assessing the impact of AD on spousal caregivers and outcomes of interventions aimed at caregivers as well as persons with AD.

Key words: Alzheimer's disease, caregivers, quality of life, Rasch measurement theory

Introduction

Alzheimer's disease (AD) is the most common form of dementia, affecting about 6% of people aged 65 and above. It is characterized by cognitive, psychiatric and behavioral symptoms, with increasing dependency in daily life (Burns and Iliffe, 2009). As such, it does not only affect the person diagnosed with AD but also other family members, who typically need to take on increasing caregiving responsibilities. This may lead to considerable caregiver impact and burden that increases over time and as the illness becomes more severe (Conde-Sala, Turro-Garriga, Calvo-Perxas, Vilalta-Franch, Lopez-Pousa, and Garre-Olmo, 2014; Pena-Longobardo and Oliva-Moreno, 2015). For example, a study from Spain (Pena-Longobardo and Oliva-Moreno, 2015) suggested that almost half of informal AD caregivers experienced health-related problems (e.g., depression and fatigue), 90% had social and leisure-related consequences (e.g., decreased time for leisure and social activities), and 75% of those still working experienced difficulties related to their professional lives (e.g., decreased working time or early retirement, difficulties keeping up with work demands, and financial problems) due to their caregiving role.

Despite this impact on family caregivers, there is a lack of measures designed to capture the impact of AD on family caregivers. In addition to health-related quality of life (HRQoL) instruments (Andreakou, Papadopoulos, Panagiotakos, and Niakas, 2016), the most commonly used caregiver questionnaires assess caregiver burden. Of these, the Zarit Burden Interview (ZBI) is the most widely used (Zarit, Reever, and Bach-Peterson, 1980). While this and similar instruments can provide valuable information, they typically also have important shortcomings. For example, their development was not based on a specific model or theory of the variable they attempt to represent, their content has been derived from professionals rather than being based on experiences reported directly by caregivers themselves, and/or their developments have not been informed by the requirements of a clear measurement model, but rather by generally vague classical test theory criteria (Farcnik and Persyko, 2002; Schwartz and Darlak, 2012). Similar issues are also evident in more recent developments (Cole, Ito, Chen, Cheng, Bolognese, and Li-McLeod, 2014; Erder, Wilcox, Chen, O'Quinn, Setyawan, and Saxton, 2012).

Here we report on the development and measurement properties of a self-report questionnaire specific to AD family caregivers. The theoretical basis for the measure is the needs-based quality of life (QoL) model. This model postulates that life gains its quality from the ability and capacity of individuals to satisfy their human needs (Hunt and McKenna, 1992; McKenna, 2011; McKenna and Doward, 2004). Consequently, the questionnaire assesses a unidimensional construct, with QoL assumed to be higher when most needs are fulfilled and poorer when few needs are being satisfied. In addition, and in interaction with the theoretical needs-based QoL model, the development of the new questionnaire was guided by Rasch measurement theory (RMT), which also was used to test its measurement properties.

The objectives of this study were to

- test how well AD family caregivers' responses to the new questionnaire met the requirements of RMT,
- use RMT to guide refinements of the questionnaire, and
- test whether an a priori theory-based general item hierarchy could be confirmed empirically.

Methods

The development of the new questionnaire is described in detail elsewhere (McKenna, Rouse, and Hagell, 2018). Briefly, questionnaire content was derived from qualitative interviews with 40 AD family caregivers (29 spousal and 11 non-spousal caregivers) in the UK. Analysis of the interview transcripts identified items (statements) relevant to need fulfillment. An initial draft questionnaire was tested for relevance and content validity in face-to-face cognitive debriefing interviews with AD caregivers in five countries (UK, Germany, Spain, Italy and the United States). Removal of culturally and/or linguistically non-equivalent items resulted in a draft questionnaire with 30 dichotomously scored (0 = *not true*; 1 = *true*) items. This version was administered by a postal survey to 268 AD family caregivers (Table 1) on two occasions, two weeks apart. This provided 473 completed questionnaires. The study was approved by the respective research ethics committees, and all participants gave written informed consent.

[Table 1 here]

Following data collection, four researchers experienced with the needs-based model reviewed items in terms of their expected hierarchical representation based on the needs-based QoL model, where each item was assessed as representing "lower", "moderate" or "higher" QoL impact. Examples of items expected to represent lower QoL impact included *irritability* and *distress*. *Independence* and *self-care* issues were expected to represent moderate QoL impact, while life being *a living hell* and a lack of *support* were expected to represent higher QoL impact.

This process also identified some potentially problematic items. The item *nobody calls round* was not considered to be on the same measurement continuum as the other items (thus considered a potential source of multidimensionality) as it relates to external factors (the behaviour of third parties). Furthermore, depending on individual preferences, it was considered difficult to establish whether this statement is positive or negative. In addition, some potential redundancies (i.e., local response dependencies) were also identified (*physically drained* / *tired all the time*, and *self-care* / *appearance*). However, all items were maintained and submitted for analysis according to RMT (Andrich, 1988; Andrich, Sheridan, and Luo, 2013; Hobart and Cano, 2009; Rasch, 1960) using the RUMM2030 software (Professional Edition, version 5.4) (Andrich, Sheridan, and Luo, 1997-2015). Bonferroni adjustments for multiple null hypothesis testing were applied (alpha level of significance = 0.05) (Bland and Altman, 1995; Hagell and Westergren, 2016).

Analyses addressed targeting, reliability (as estimated using the person separation index, PSI) (Andrich, 1982), model fit (item fit residuals, item chi-square statistics and item characteristic curves, ICC), local response dependency (relative item fit residual correlations >0.25) (Christensen, Makransky, and Horton, 2017; Marais, 2013), and 2-way ANOVA-based analyses of item fit residuals regarding uniform and non-uniform differential item functioning (DIF) (Hagquist and Andrich, 2017) by time (T1 vs. T2), age (≤59 vs. ≥60 years old; as defined by the median age of 59.5 years), gender and caregiver status (spousal vs. non-spousal). DIF by time of assessment was checked at the outset of the analyses and absence of DIF by time was taken as support for merging data from the two time points, thereby gaining precision of estimates (Kyngdon, 2011). Analyses include graphical, descriptive and statistical methods, which were considered of equal primacy. Formal testing of the extent to which the *a priori* theory-based general item hierarchy was confirmed empirically was conducted by correlating the expected "lower", "medium" and "higher" QoL impact assessments (scored 1 to 3, respectively) with RMT derived item logit locations using polyserial correlation (as implemented in the “polychoric” package for Stata, version 14.1). Since the analyses were conducted interactively within the RMT framework as well as in interaction with theoretical considerations, further details regarding the analysis process, as well as interpretations are integrated in the Results section.

Results

Initial analyses (*n*=473) revealed no DIF by time, but 11 items displayed DIF by caregiver status. In seven cases DIF was uniform, in three it was non-uniform, and one item exhibited a combination of uniform and non-uniform DIF. In addition, eight items showed misfit, either in terms of large and statistically significant chi-square values and/or fit residual values outside the generally acceptable ±2.5 range. These observations were largely corroborated by the respective ICCs. The mean (*SD*) person location was 0.045 (1.51) logits, and reliability (PSI) was 0.88.

These observations suggested differences in meaning of certain items to spousal and non-spousal caregivers. Indeed, previous studies have suggested that the impact of caring for someone with AD is dependent on this variable (Barber and Pasley, 1995; Conde-Sala, Garre-Olmo, Turro-Garriga, Vilalta-Franch, and Lopez-Pousa, 2010; Reed, Belger, Dell'agnello, Wimo, Argimon, Bruno, Dodel, Haro, Jones, and Vellas, 2014). Furthermore, re-review of the qualitative interview data from the development of the questionnaire suggested that spousal and non-spousal caregivers identified different experiences of caregiving. A review of the items in the questionnaire also indicated that they mainly emerged from interviews with spousal caregivers and, therefore, that it primarily represented their experiences. Consequently, non-spousal caregivers were removed from the sample and the remaining analyses were based on data from spousal caregivers only (210 completed questionnaires from *n* = 116 spousal caregivers) (Table 1).

[Figure 1 here]

In accord with the initial analysis of all carers, spousal carer responses to the 30 items did not suggest any DIF by time or age, but one item showed non-uniform DIF by gender (Figure 1). Reliability remained at 0.88 and the mean (*SD*) person location was 0.106 (1.43) logits. Evaluation of the fit between data and the Rasch model revealed standardized fit residuals ranging between –2.37 and 2.40, suggesting generally acceptable fit for all items (Table 2). However, inspection of the chi-square values indicated that the item with DIF misfit, which was corroborated by inspection of the associated ICC (Figure 2A). As expected from the item review (see Methods), the pattern of indices suggests that the misfit relates to multidimensionality. While other fit statistics did not signal any obvious problems, examination of the inter-item fit residual correlation matrix revealed coefficients > 0.25 above the overall mean correlation (*r*, –0.032) for six item pairs: items 13/26 (*r*, 0.348), 17/22 (*r*, 0.344), 4/7 (*r*, 0.326), 11/25 (*r*, 0.275), 21/24 (*r*, 0.252), and items 23/24 (*r*, 0.244). These findings confirmed the expected dependencies (items 11/25 and 21/24) empirically. However, upon re-review the additional dependencies were generally considered conceptually logical, with the exception of items 23/24.

[Figure 2 here]

[Table 2 here]

Table 2 also displays the hierarchical item ordering according to the estimated logit locations from less (negative logits) to more (positive logits) QoL impact, together with the respective uncertainties (expressed as standard errors) of item locations. Inspection of Table 2 suggests that *a priori* expectations regarding the general hierarchical item ordering was supported empirically. The polyserial correlation between a priori expectations and empirical item locations was 0.64.

The questionnaire was then modified iteratively according to conceptual considerations in combination with results from empirical testing. First, item 17, which exhibited both DIF and misfit, was deleted. Again, no DIF by time was found and standardized item fit residuals ranged between –2.30 and 2.46, suggesting generally acceptable fit for all items. Examination of the inter-item fit residual correlation matrix revealed coefficients >0.25 above the overall mean correlation (*r*, –0.033) for four item pairs: items 13/26 (*r*, 0.336), 4/7 (*r*, 0.317), 11/25 (*r*, 0.266), and items 21/24 (*r*, 0.258). Thus, the spuriously high fit residual correlation between items 23/24 had now disappeared. Reliability was 0.87 and the mean (*SD*) person location was 0.143 (1.46) logits.

Next, item pairs with local response dependency were reviewed to guide item reduction. This process was governed by item content as well as location, aiming for as wide a range and even spread as possible along the latent quantitative continuum. Consequently, items 4, 24, 25 and 26 were discarded sequentially and interactively in favour of items 7, 21, 11 and 13, respectively. The resulting 25-item scale yielded a mean (*SD*) person location of 0.194 (1.42) logits (Figure 3A). Reliability of this 25-item scale was 0.85, implying that the scale can detect 3.5 statistically distinct strata of people (as compared to 3.8 strata for the original 30-item version). The best point of person measurement was at about 0 logits, where the standard error of person locations approximated 0.44. It is also noted that the relation between raw scores and their implied linear measures suggested that raw scores are approximately linear within a relatively wide raw score range, from about 3 to 22 out of the possible 0-25 range (Figure 3B).

[Figure 3 here]

There was no DIF by time, age or gender and local dependence was acceptable with the largest relative inter-item residual correlation being 0.212 (absolute correlation = 0.173). Other item statistics are reported in Table 3. Item locations spanned 3.87 logits and all standardized item fit residuals were within acceptable limits, ranging between –2.25 and 2.22. However, item 30 exhibited a relatively large Chi-square value with borderline statistical significance. Inspection of the corresponding ICC suggested that the misfit primarily originated from persons representing the two class intervals at the higher end scoring higher and lower, respectively, than expected (Figure 2B). For comparison purposes, Figure 2C represents the ICC of a well-fitting item. Given the marginal significance of its misfit (p=0.002, corresponding to a Bonferroni adjusted p-value of 0.05), reasonable fit residual and its conceptual appropriateness, at this stage the item was considered acceptable within the frame of reference of the scale as a whole. The hierarchical item ordering of the revised 25-item measure (Table 3) exhibited improved accordance with *a priori* expectations, with a polyserial correlation of 0.73.

[Table 3 here]

Discussion

This paper has described the development and psychometric testing of a new self-reported needs-based QoL measure specific for spousal AD family caregivers. This new measure (named the Alzheimer's Patient Partners Life Impact Questionnaire, APPLIQue) covers a reasonably wide range of measurement and exhibits adequate model fit, reliability and measurement invariance by time, gender and age, and supported internal construct validity.

Five items from the original pool of 30 items were discarded. In three of these cases, the issues identified through testing were expected. In the other two cases the identified measurement issues were also substantively supported upon item re-review. It is thus important to note that amendments were not primarily data driven. As cautioned against elsewhere (Andrich, 1988; Andrich et al., 2013; Hagell, 2014; Hobart and Cano, 2009), over reliance on statistical criteria may be detrimental due to, for example, capitalizing on chance and sampling effects rather than theory. Therefore, a sixth item that also displayed some signs of misfit was retained in the APPLIQue since the statistical evidence against this item was inconsistent and less convincing in view of theoretical considerations. Importantly, hierarchical item locations of the final item set showed good correspondence with *a priori* expectations, providing support for construct validity.

Reducing the questionnaire from 30 to 25 items did not alter its measurement properties substantially. Items covered a similar range and reliability was only marginally reduced, with both the 30- and 25-item versions being able to detect between 3 and 4 strata of people. However, the APPLIQue fails to represent people located at the highest and lowest ranges, which suggests room for improvement in terms of measurement precision at these locations. Nevertheless, the vast majority of the sample is located within the range of item locations, suggesting that the APPLIQue should be adequate for most purposes.

Our observations corroborate previous studies (Barber and Pasley, 1995; Conde-Sala et al., 2010; Reed et al., 2014), suggesting that the impact of caring for a family member with AD differs between spousal and non-spousal (i.e., adult children) caregivers. Consequently, the APPLIQue is a spousal caregiver measure rather than one for use across all family caregivers. Further developmental work on the APPLIQue may be directed towards the development of a complementary questionnaire for non-spousal caregivers or the identification of items that are invariant between, and of equal relevance to, both groups. Given the experiences from this study, we suggest that the former option appears more feasible. However, a combined strategy may also be envisioned, where common items serve as links between the two in order to render comparable measures of impact, similar to what previously has been accomplished with patient-reported outcome measures across diagnoses (Rouse, Twiss, and McKenna, 2016).

In a similar vein, it is interesting to note that previous research has found caregiver burden to be similar among family caregivers of people with, e.g., dementia, stroke and Parkinson's disease (PD) and that the needs of family caregivers appear common between caregivers of people with AD and PD (Habermann and Davis, 2005; Thommessen, Aarsland, Braekhus, Oksengaard, Engedal, and Laake, 2002). Furthermore, neurodegenerative conditions such as AD, PD and Huntington's disease are all associated with varying degrees and types of mobility problems, cognitive impairments, behavioural and psychological changes (Smith, 2011), which could imply similarities in terms of caregiver impact. This suggests the possibility of developing a common caregiver QoL measure across neurodegenerative disorders, where there is a good case for the APPLIQue to provide a starting point. However, such developments would need to be coupled with additional qualitative interviews with caregivers of people with other neurodegenerative disorders, to ensure content validity. Nevertheless, if a core set of common items were identified, they could be used to co-calibrate various disease-specific, as well as spousal and non-spousal carer-specific items, thereby moving towards a common caregiver impact measurement system.

In conclusion, the APPLIQue is the first person-centered QoL measure specific to AD spousal caregivers. As such, it may fill an important gap in assessing the impact of AD on spousal caregivers and as an outcome measure in interventional trials targeting caregivers as well as persons with AD. [1]

---

[1] The APPLIQue is available for use in the UK, Germany, Spain, Italy, the United States, Greece and Sweden, and may be obtained from www.galen-research.com.

Acknowledgements

The authors wish to thank the participants for their collaboration. The study was supported by Kristianstad University, Kristianstad, Sweden.

Conflict of interest

None.

References

Andreakou, M. I., Papadopoulos, A. A., Panagiotakos, D. B., and Niakas, D. (2016). Assessment of health-related quality of life for caregivers of Alzheimer's disease patients. *International Journal of Alzheimer's Disease, 2016*, Article ID 9213968. www.hindawi.com/journals/ijad/2016/9213968

Andrich, D. (1982). An index of person separation in latent trait theory, the traditional KR-20 index, and the Guttman scale response pattern. *Educational Research and Perspectives, 9*, 95-104.

Andrich, D. (1988). *Rasch models for measurement*. Beverly Hills, CA: Sage.

Andrich, D., Sheridan, B., and Luo, G. (2015). RUMM2030: Rasch Unidimensional Models for Measurement [computer software]. Perth, WA, Australia: RUMM Laboratory Pty Ltd. www.rummlab.com.au

Andrich, D., Sheridan, B., and Luo, G. (2013). *Interpreting RUMM2030*. Perth, WA, Australia: RUMM Laboratory Pty Ltd.

Barber, C. E., and Pasley, B. K. (1995). Family care of Alzheimer's patients: The role of gender and generational relationship on caregiver outcomes. *Journal of Applied Gerontology, 14*, 172-192.

Bland, J. M., and Altman, D. G. (1995). Multiple significance tests: The Bonferroni method. *British Medical Journal, 310*, 170.

Burns, A., and Iliffe, S. (2009). Alzheimer's disease. *British Medical Journal, 338*, b158.

Christensen, K. B., Makransky, G., and Horton, M. (2017). Critical values for Yen's Q3: Identification of local dependence in the Rasch model using residual correlations. *Applied Psychological Measurement, 41*, 178-194.

Cole, J. C., Ito, D., Chen, Y. J., Cheng, R., Bolognese, J., and Li-McLeod, J. (2014). Impact of Alzheimer's Disease on Caregiver Questionnaire: Internal consistency, convergent validity, and test-retest reliability of a new measure for assessing caregiver burden. *Health and Quality of Life Outcomes, 12*, 114.

Conde-Sala, J. L., Garre-Olmo, J., Turro-Garriga, O., Vilalta-Franch, J., and Lopez-Pousa, S. (2010). Differential features of burden between spouse and adult-child caregivers of patients with Alzheimer's disease: An exploratory comparative design. *International Journal of Nursing Studies, 47*, 1262-1273.

Conde-Sala, J. L., Turro-Garriga, O., Calvo-Perxas, L., Vilalta-Franch, J., Lopez-Pousa, S., and Garre-Olmo, J. (2014). Three-year trajectories of caregiver burden in Alzheimer's disease. *Journal of Alzheimer's Disease, 42*, 623-633.

Erder, M. H., Wilcox, T. K., Chen, W. H., O'Quinn, S., Setyawan, J., and Saxton, J. (2012). A new measure of caregiver burden in Alzheimer's disease: the caregiver-perceived burden questionnaire. *American Journal of Alzheimer's Disease and Other Dementias, 27*, 474-482.

Farcnik, K., and Persyko, M. S. (2002). Assessment, measures and approaches to easing caregiver burden in Alzheimer's disease. *Drugs and Aging, 19*, 203-215.

Habermann, B., and Davis, L. L. (2005). Caring for family with Alzheimer's disease and Parkinson's disease: needs, challenges and satisfaction. *Journal of Gerontological Nursing, 31*, 49-54.

Hagell, P. (2014). Testing rating scale unidimensionality using the principal component analysis (PCA)/*t*-test protocol with the Rasch model: The primacy of theory over statistics. *Open Journal of Statistics, 4, 456-465.*

Hagell, P., and Westergren, A. (2016). Sample size and statistical conclusions from tests of fit to the Rasch model according to the Rasch unidimensional measurement model (RUMM) program in health outcome measurement. *Journal of Applied Measurement, 17*, 416-431.

Hagquist, C., and Andrich, D. (2017). Recent advances in analysis of differential item functioning in health research using the Rasch model. *Health and Quality of Life Outcomes, 15*, 181.

Hobart, J., and Cano, S. (2009). Improving the evaluation of therapeutic interventions in multiple sclerosis: The role of new psychometric methods. *Health Technology Assessment, 13*, iii, ix-x, 1-177.

Hunt, S. M., and McKenna, S. P. (1992). The QLDS: a scale for the measurement of quality of life in depression. *Health Policy, 22*, 307-319.

Kyngdon, A. (2011). Is combining samples productive? A quick check via tests of DIF. *Rasch Measurement Transactions, 25*, 1324-1325.

Marais, I. (2013). Local dependence. In K. B. Christensen, S. Kreiner, and M. Mesbah (Eds.), *Rasch models in health* (pp. 111-130). Croydon, Surrey, UK: John Wiley.

McKenna, S. P. (2011). Measuring patient-reported outcomes: moving beyond misplaced common sense to hard science. *BMC Medicine, 9*, 86.

McKenna, S. P., and Doward, L. C. (2004). The needs-based approach to quality of life assessment. *Value in Health, 7 Suppl 1*, S1-3.

McKenna, S. P., Rouse, M., and Hagell, P. (2018). International development of the Alzheimer's Patient Partners Life Impact Questionnaire (APPLIQue). *Submitted for publication*.

Pena-Longobardo, L. M., and Oliva-Moreno, J. (2015). Caregiver burden in Alzheimer's disease patients in Spain. *Journal of Alzheimer's Disease, 43*, 1293-1302.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen, Denmark: Danmarks Paedagogiske Institut.

Reed, C., Belger, M., Dell'agnello, G., Wimo, A., Argimon, J. M., Bruno, G., Dodel, R., Haro, J. M., Jones, R. W., and Vellas, B. (2014). Caregiver burden in Alzheimer's disease: Differential associations in adult-child and spousal caregivers in the GERAS observational study. *Dementia and Geriatric Cognitive Disorders Extra, 4*, 51-64.

Rouse, M., Twiss, J., and McKenna, S. P. (2016). Co-calibrating quality-of-life scores from three pulmonary disorders: Implications for comparative-effectiveness research. *Journal of Medical Economics, 19*, 596-603.

Schwartz, S., and Darlak, L. (2012). *Selected caregiver assessment measures: A resource inventory for practitioners. 2nd Edition.* Family Caregiver Alliance and Benjamin Rose Institute on Aging. Retrieved from www.caregiver.org/sites/caregiver.org/files/pdfs/SelCGAssmtMeas_ResInv_FINAL_12.10.12.pdf

Smith, S. (2011). *The neuro care manual: A guide to neurology for nurses and family carers*. New Delhi, India: B. Jain Publishers.

Thommessen, B., Aarsland, D., Braekhus, A., Oksengaard, A. R., Engedal, K., and Laake, K. (2002). The psychosocial burden on spouses of the elderly with stroke, dementia and Parkinson's disease. *International Journal of Geriatric Psychiatry, 17*, 78-84.

Zarit, S. H., Reever, K. E., and Bach-Peterson, J. (1980). Relatives of the impaired elderly: correlates of feelings of burden. *Gerontologist, 20*, 649-655.

Table 1

Demographic characteristics of postal survey caregivers and patients[a]

	Spousal and non-spousal caregivers (*n*=268)	Spousal caregivers (*n*=116)
Caregivers		
Age (years), mean (*SD*)	60.7 (13.2)	70.3 (9.6)
Female sex, *n* (%)	207 (77)	88 (76)
Duration of caregiving (years), mean (*SD*)	5.7 (6.9)	6.6 (9)
Duration of caregiving (years), median (q1-q3)	4 (2.4-6)	4 (3-6)
Perceived health, *n* (%)		
Excellent/very good	50 (19)	13 (11)
Good/fair	183 (69)	86 (75)
Poor	31 (12)	15 (13)
Patients		
Confused, *n* (%)		
All the time	72 (27)	29 (25)
Most of the time	110 (42)	44 (38)
Sometimes	76 (29)	39 (34)
No	5 (2)	4 (3)
Can be left alone, *n* (%)		
As long as needed	60 (23)	24 (21)
A short time	97 (37)	45 (39)
No	106 (40)	46 (40)
Recognise caregiver, *n* (%)		
All the time	118 (44)	61 (53)
Most of the time	82 (31)	32 (28)
Sometimes	44 (17)	14 (12)
Never	21 (8)	9 (8)
Incontinent, *n* (%)		
All the time	74 (29)	34 (30)
During the day	42 (16)	16 (14)
At night	28 (11)	8 (7)
No	115 (44)	54 (48)
Wander around the house, *n* (%)	129 (49)	63 (55)

[a] Percentages may not sum up to 100 due to rounding effects.

Table 2

Rasch item and fit statistics of the initial 30-item questionnaire among spousal AD caregivers (n=210)

Items				Fit	Chi-2
No.	Topic	Location [a]	*SE*	residual [b]	values [c]
15	Distress	–1.944	0.21	–0.88	4.93
18	Always on mind	–1.571	0.19	1.27	9.20
8	Sharing	–1.374	0.19	–0.83	3.41
4	Emotional impact	–1.354	0.19	–0.04	8.99
6	Freedom	–1.348	0.19	–0.15	4.24
7	Wearing down	–1.098	0.18	–1.75	9.48
30	Guilt	–1.026	0.18	2.06	9.98
2	Irritability	–0.906	0.18	1.18	1.86
5	Communication	–0.569	0.17	0.69	5.29
12	Want to cry	–0.552	0.17	–0.25	1.93
9	Pressure	–0.433	0.17	–0.20	1.39
11	Physically drained	–0.327	0.17	–2.37	8.10
23	Household jobs	–0.31	0.17	1.37	5.46
20	Independence	–0.292	0.17	–1.44	4.07
16	Energy drain	–0.239	0.17	–1.76	9.45
25	Tired all the time	–0.238	0.17	–1.70	7.02
21	Self-care	–0.128	0.17	–0.98	5.06
19	Inertia	0.454	0.16	–0.96	5.38
28	Recognition	0.506	0.16	–1.03	0.88
3	Shopping	0.596	0.17	0.42	5.05
27	Want to run away	0.694	0.17	0.50	1.99
10	Loneliness	0.73	0.17	–1.39	3.51
29	Situation makes me ill	0.85	0.17	–1.15	5.73
17	Nobody calls round	0.852	0.17	2.17	**19.11**
1	Lack of time	1.023	0.17	1.89	6.64
26	Illness ruins life	1.247	0.18	–1.99	8.95
14	Sleep	1.289	0.18	2.40	5.87
24	Appearance	1.691	0.19	–0.84	3.14
22	Support	1.868	0.19	0.36	6.03
13	A living hell	1.908	0.20	–1.30	5.65

[a] Ordered by estimated location in linear log-odds units (logits), with mean item location set at 0, from lowest to highest quality of life impact.

[b] Fit residuals summarise the deviation of observed from expected responses. Deviations from the recommended range of –2.5 to +2.5 are bold.

[c] Bonferroni corrected statistically significant deviations across class intervals (suggesting item misfit) are bold.

Note: AD, Alzheimer's disease; *SE*, standard error

Table 3

Rasch item and fit statistics of the final 25-item questionnaire among spousal AD caregivers (n=210)

Items				Fit	Chi-2
No.	Topic	Location [a]	*SE*	residual [b]	values [c]
15	Distress	–1.876	0.21	–1.05	5.89
18	Always on mind	–1.488	0.19	0.91	5.24
8	Sharing	–1.271	0.19	–0.76	2.33
6	Freedom	–1.261	0.19	–0.03	6.29
7	Wearing down	–1.027	0.18	–1.70	7.15
30	Guilt	–0.933	0.18	2.03	**15.16**
2	Irritability	–0.838	0.18	0.81	0.57
5	Communication	–0.481	0.17	0.69	5.54
12	Want to cry	–0.459	0.17	–0.39	2.34
9	Pressure	–0.354	0.17	–0.52	1.44
11	Physically drained	–0.254	0.17	–2.25	7.68
23	Household jobs	–0.216	0.17	1.36	5.82
20	Independence	–0.188	0.17	–1.66	3.99
16	Energy drain	–0.141	0.17	–1.87	6.14
21	Self-care	–0.019	0.17	–0.84	2.49
19	Inertia	0.544	0.16	–0.85	3.67
28	Recognition	0.586	0.16	–1.01	0.87
3	Shopping	0.696	0.17	0.17	5.61
27	Want to run away	0.795	0.17	0.42	4.46
10	Loneliness	0.821	0.17	–1.32	4.16
29	Situation makes me ill	0.943	0.17	–1.12	4.43
1	Lack of time	1.109	0.17	1.79	8.30
14	Appearance	1.376	0.18	2.22	8.56
22	Support	1.940	0.19	0.52	5.49
13	A living hell	1.996	0.20	–1.19	4.49

[a] Ordered by estimated location in linear log-odds units (logits), with mean item location set at 0, from lowest to highest quality of life impact.

[b] Fit residuals summarise the deviation of observed from expected responses. Deviations from the recommended range of –2.5 to +2.5 are bold.

[c] Bonferroni corrected statistically significant deviations across class intervals (suggesting item misfit) are bold.

Note: AD, Alzheimer's disease; *SE*, standard error

Legends to Figures

Figure 1

Non-uniform Differential Item Functioning (DIF) by gender for item 17 (*nobody calls round*). The item characteristic curve (ICC; grey curve) represents the expected response category endorsement (*y*-axis) at various levels of quality of life impact (*x*-axis; positive values = higher quality of life impact). Superimposed plots represent the observed responses by male (o) and female (x) caregivers.

Figure 2

Item characteristic curves (ICCs) representing items 17 (*nobody calls round*; Panel A), 30 (*guilt*; Panel B) and 2 (*irritability*; Panel C) from the initial 30-item (Panel A) and final 25-item (Panels B and C) questionnaire. The grey curves represent expected item responses (*y*-axis) for each person location (*x*-axis) on the common logit metric (*x*-axis; positive values = higher quality of life impact). Black dots represent the observed responses from groups of people at similar locations (*x*-axis).

Figure 3

Distribution of locations of people (Panel A, upper graph) and items (Panel A, lower graph) on the common logit metric (*x*-axis; positive values = higher quality of life impact) of the final 25-item questionnaire. All locations are relative to the mean item threshold location, which is set at 0 logits. Superimposed on the upper person distribution graph is the information function curve (the inverse of measurement error; higher values = less error and more information in scores, i.e., better measurement precision). Maximum information (vertical line under the information function curve) corresponds to a location of about 0 logits (representing a score of about 13 on the raw 0-25 score range; Panel B).

Figure 1

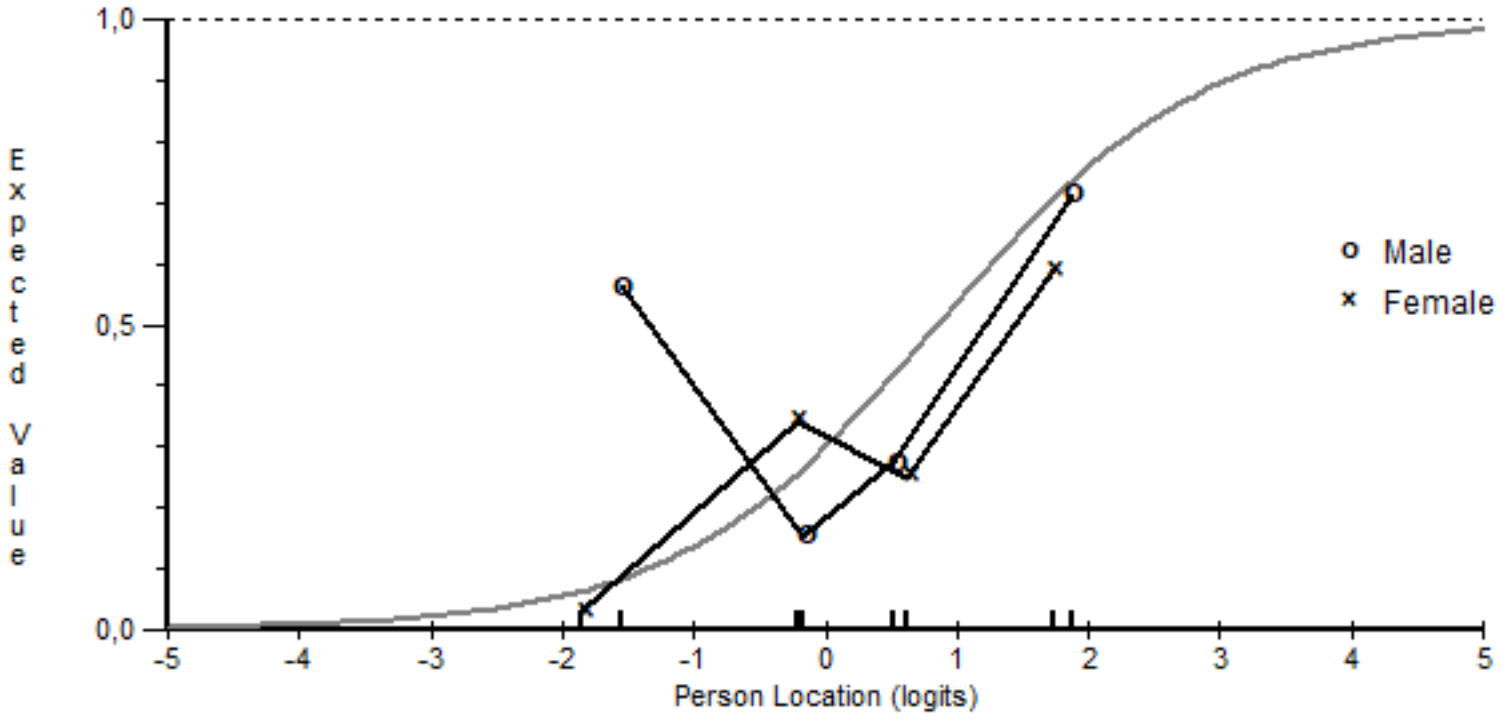

1,0
0,5
0,0
Expected Value
-5
-4
-3
-2
-1
0
1
2
3
4
5
Person Location (logits)
o Male
× Female

Figure 2

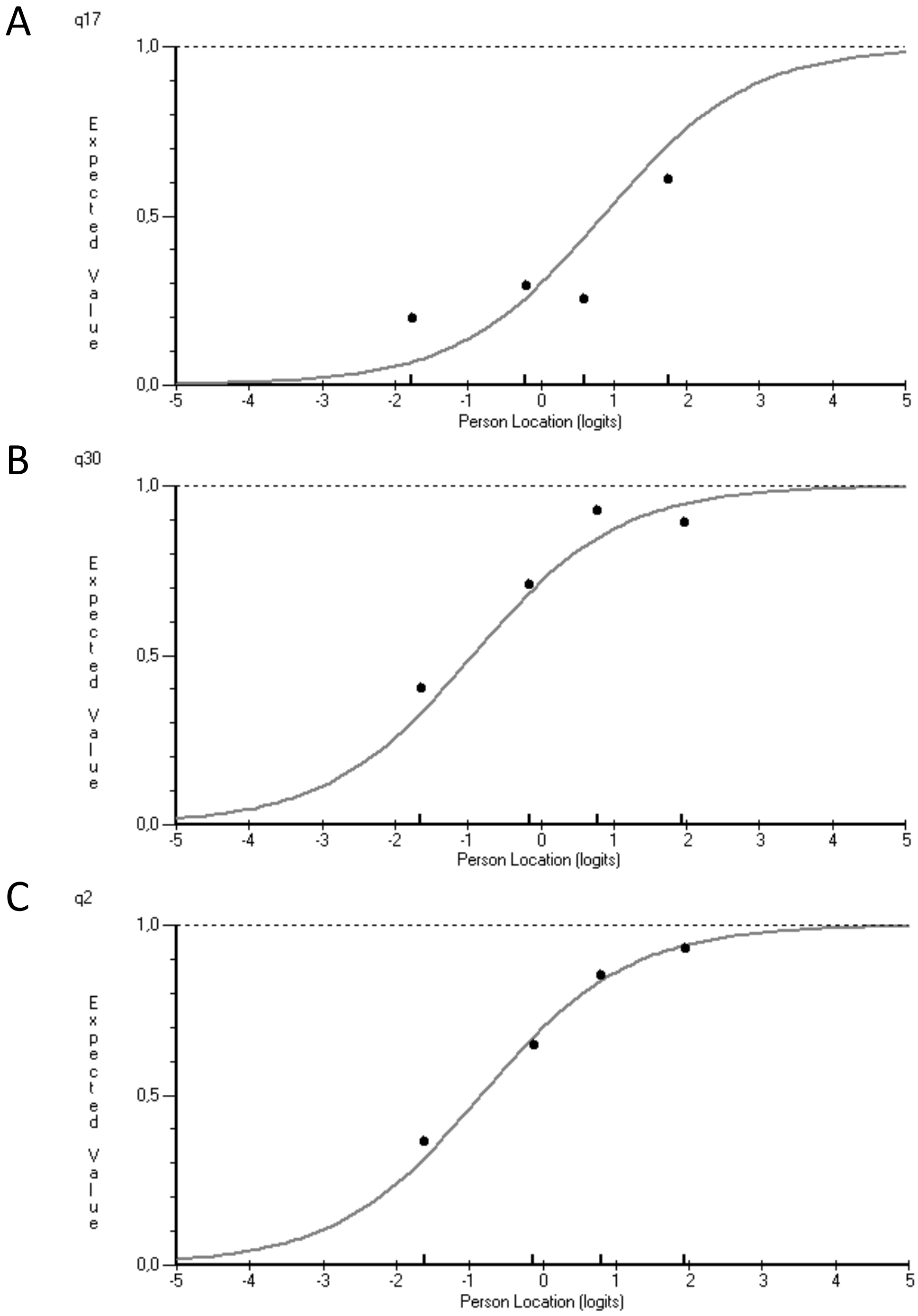
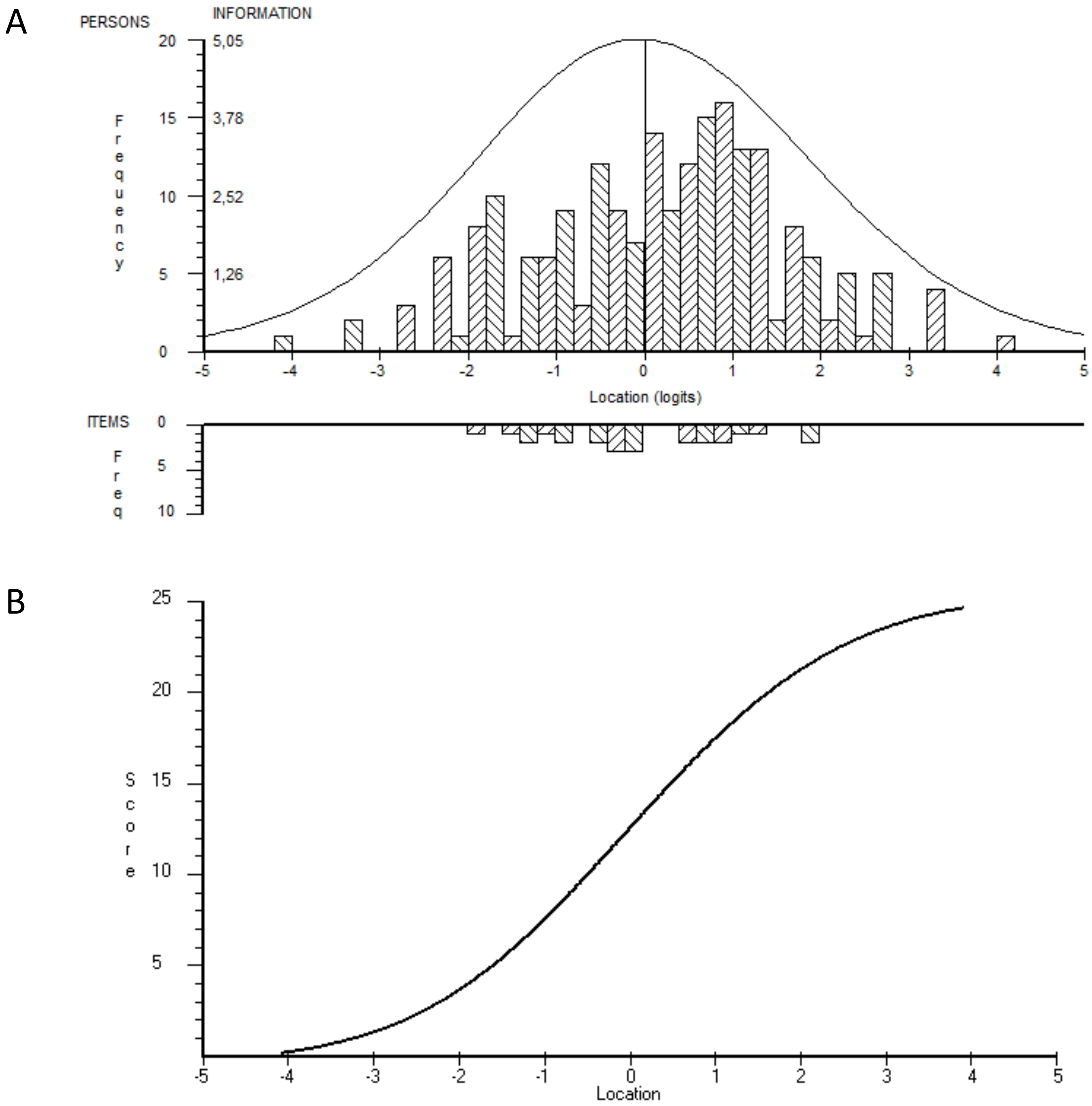

A
q17
Expected Value
1,0
0,5
0,0
-5
-4
-3
-2
-1
0
1
2
3
4
5
Person Location (logits)
B
q30
Expected Value
1,0
0,5
0,0
-5
-4
-3
-2
-1
0
1
2
3
4
5
Person Location (logits)
C
q2
Expected Value
1,0
0,5
0,0
-5
-4
-3
-2
-1
0
1
2
3
4
5
Person Location (logits)

Figure 3


A
PERSONS
INFORMATION
20
5,05
15
3,78
10
2,52
5
1,26
0
Frequency
-5
-4
-3
-2
-1
0
1
2
3
4
5
Location (logits)
ITEMS
0
5
10
Freq
B
25
20
15
10
5
Score
-5
-4
-3
-2
-1
0
1
2
3
4
5
Location