We used to develop the classification systems (DEMQOL-U and DEMQOL-Proxy-U) that have been valued to generate utility scales for use in the generation of quality-adjusted life-years (QALYs) [3].

Subsequently, we have focused on developing preference-based measures: constructing a health-state classification using Rasch analysis. Qual Life Res 2009;18:253–65.

Content Validity

Arons et al. criticize the content validity of the domain structure used in DEMQOL-U, in particular the omission of a physical domain (that impacts daily activities).

Although the conceptual framework that underlies DEMQOL and DEMQOL-Proxy included five domains (daily activities and looking after yourself, health and well-being, cognitive functioning, social relationships, and self-concept), these five domains were not fully supported by the factor analysis carried out for both the preliminary field test and the final field test during the original development process [4,5]. The final (item-reduced) version of DEMQOL did not fully represent daily activities (two items) or self-concept (0 items). We did not therefore intend to cover the original five conceptual domains in DEMQOL-U.

We instead carried out exploratory factor analysis by using a much larger data set (n = 644) and established a different five-factor structure for DEMQOL including 1) cognitive functioning, 2) positive emotion, 3) negative emotion, 4) social relationships, and 5) loneliness. We would argue that using a significantly larger data set in a population of patients with a definite diagnosis of dementia provides a stronger empirically based representation of the underlying factor structure of DEMQOL than was established by earlier analysis. Furthermore, in all the factor analysis carried out on DEMQOL [4,5], the well-being items have consistently split into positive and negative emotion, proving support for including them as separate dimensions of DEMQOL-U. Factors 1 to 3 assess mental well-being, and factors 4 and 5 assess social well-being, which, as suggested by Arons et al., are key domains to include in any measure of dementia-related health-related quality of life (HRQOL).

Qualitative work carried out by Smith et al. [4, 5] suggested that the daily activities items were more important when reported by a proxy than by people with dementia themselves.

Arons et al. have focused on DEMQOL-U, but the DEMQOL-Proxy-U classification system includes an appearance dimension, which is related to daily activities. The difference in the two classification systems serves to highlight one of the reasons why both DEMQOL and DEMQOL-Proxy should be used as complementary measures.

Although the physical aspect of HRQOL is likely to be an important area for many patients with dementia, Arons et al. point out that a dementia measure should emphasize mental and social components of HRQOL. A condition-specific measure should focus on those dimensions most related to the condition and the outcomes of its treatment. We argue that this is exactly what DEMQOL-U does. Therefore, the utility values generated by DEMQOL-U are valid for assessing key outcomes of any dementia-related intervention or treatment. It is also expected that the system will be used alongside other instruments that would assess co-occurring aspects of HRQOL. Problems with physical aspects of HRQOL arising from comorbidities will not impact on the size of the change in health-state value as a result of an intervention provided that there are no preference interactions with the mental and cognitive dimensions. Recent evidence from the valuation of a mental health measure (Clinical Outcomes in Routine Evaluation [CORE-6D]) supports the notion that the relationship between physical and mental health is additive [6,7].

Arons et al. also criticize the cognition and social relationships items included in DEMQOL-U because of the question stem used in the original instrument, which frames the question in terms of worrying about the concept (e.g. “How worried have you been about forgetting things that happened recently?”). DEMQOL includes an item directly asking about worry (“Have you felt worried”) that in the revalidation factor analysis fitted with the negative emotion factor. Because of the focus on worry, Arons et al. suggest the possibility of confounding. In the development of DEMQOL framing the questions in terms of worry about a particular concept was the stem that was most easily understood and so it was used for 15 of the 28 items. The original DEMQOL measures have been shown to be psychometrically valid and reliable by using this question stem, and therefore provide a valid basis for the development of a utility measure based on the standard question wording.

Analytical Strategies

Arons et al. also criticize the analytical strategies we used, which involves factor and Rasch analysis as part of a six-step process.
used to derive condition-specific health-state classification systems, and subsequently value them to generate utility scores [8].

The six-step process is used as a guide to developing condition-specific preference-based measures and has now been used widely across a range of conditions [9–13]. It does not prescribe the exact techniques for achieving each step. We used factor analysis to investigate the dimensional structure of the instrument and Rasch analysis to test the performance of the items and select the best quality item that performs well and is clinically meaningful. Other techniques could have been used, but in the context of the development of DEMQOL-U, we do not believe Arons et al.’s criticisms of the factor and Rasch analysis are valid. First, Arons et al. state that factor analysis seems unnecessary because the technique was used to develop the original DEMQOL. As described previously, however, the original factor analyses were carried out on small samples and proved inconclusive and difficult to interpret. Therefore, we would argue that it was essential to investigate the factor structure of DEMQOL by using a larger sample (n = 644 in Mulhern et al. [2]) to establish a more robust and interpretable domain structure on which to base the DEMQOL-U classification system. We have not used factor analysis to provide information regarding the weight or importance of the items, as all items in the dimensions established are tested by using Rasch, and the importance (or weight) of the selected items is established during the valuation study.

Arons et al. state that we may not have used the correct Rasch model because we apply it to Likert- scale data, which does not have the correct response structure. This criticism is leveled at DEMQOL-U [2] but also a number of studies developing condition-specific health-state preference-based measures using the six-step process as a guide [9–13]. As Arons et al. discuss, Rasch analysis requires a “cumulative” data structure, where if a person agrees with a certain level, he or she must also agree with the statements that precede this level. Arons et al., however, argue that HRQOL measures using Likert scales are “single peaked,” because respondents will agree with statements close to their own position on the scale and disagree with statements (or response categories) that are further away from their own position. Therefore, they argue that we should use a probabilistic unfolding model that takes into account this implied response process.

In response, we would argue that DEMQOL does fit the cumulative response structure required to use the standard polytomous Rasch model (which has been used to test a range of instruments assessing HRQOL [14]). This is because we are plotting a person’s location on the underlying scale measuring a particular HRQOL construct, where the severity of the construct increases cumulatively. For example, on the underlying scale, “felt cheerful a lot” is more cheerful than “felt cheerful not at all.” We are not judging a respondent’s opinion of a more abstract concept, which would require them to agree/disagree with the other levels leading to a peaked data structure (and use of an unfolding model). Therefore, we believe the standard polytomous model is appropriate to assess DEMQOL.

In summary we believe that DEMQOL-U is based on important and relevant components of HRQOL in dementia and provides a valid measurement of their utility. This is because the dimensions were derived from the further exploration of the factor structure of DEMQOL, and subsequent Rasch analysis selected the strongest item to represent each domain. Although further validity testing of the measures is required, we believe that the utility values produced may improve the measurement of quality-adjusted life-years in dementia.

Source of financial support: This project was funded by the National Institute of Health Research: Health Technology Assessment programme 07/73/01.

REFERENCES

[1] Arons AMM, Scholzel-Dorenbos CJM, Olde Rikkert MGM, Krabbe PFM. Improving the measurement of QALYs in dementia: some important considerations. Value Health 2012;15:XX–XX.