

Notice -- This material may be protected by copyright law  
(Title 17 US Code).

**\*19517776\*****Request # 19517776****Mail To:**

VA MEDICAL CENTER  
 CLINICAL INFORMATION MANAGEMENT-LIBRARY  
 Attn: Mary Hess  
 216 S. FOSTER DR.  
 BATON ROUGE, LA 70806

*Read*  
 MAR 17, 2006  
*Seat*

**DOCLINE: Journal Copy Epayment**

Title: Journal of clinical epidemiology.  
 Title Abbrev: J Clin Epidemiol  
 Citation: 2003 Feb;56(2):138-47  
 Article: Quality of life assessment in the community-dwelli  
 Author: Osborne R; Hawthorne G; Lew E; Gray L  
 .nlm Unique ID: 8801383 Verify: PubMed  
 PubMed UI: 12654408  
 ISSN: 0895-4356 (Print)  
 Publisher: Elsevier, New York :  
 Copyright: Copyright Compliance Guidelines  
 Authorization: MEH  
 Need By: N/A  
 Maximum Cost: **Free** ✓  
 Patron Name: Dr. Dumitrescu (mihnea@mac.com)  
 Referral Reason: Not owned (title)  
 Library Groups: FreeShare,VALNET,EFTS  
 Phone: 1.225.761-6850  
 Fax: 1.225.761-6805  
 Email: mary.hess@med.va.gov  
 Routing Reason: Routed to TXUTTY in Serial Routing - cell 2  
 Received: Mar 17, 2006 ( 02:42 PM EST )  
 Lender: UNIVERSITY OF TEXAS HEALTH CENTER AT TYLER/ TYLER/ TX  
 USA (TXUTTY)

This material may be protected by copyright law (TITLE 17,U.S. CODE)

**Bill to: LAUVNO**

VA MEDICAL CENTER  
 CLINICAL INFORMATION MANAGEMENT-LIBRARY  
 Attn: Mary Hess  
 216 S. FOSTER DR.  
 BATON ROUGE, LA 70806

# Quality of life assessment in the community-dwelling elderly: Validation of the Assessment of Quality of Life (AQoL) Instrument and comparison with the SF-36

Richard H. Osborne<sup>a,\*</sup>, Graeme Hawthorne<sup>b</sup>, Elizabeth A. Lew<sup>c</sup>, Len C. Gray<sup>d</sup>

<sup>a</sup>Centre for Rheumatic Diseases, Department of Medicine, The Royal Melbourne Hospital, and the Centre for Health Program Evaluation, School of Population Health, The University of Melbourne, Charles Connibere Building, Ground Floor, West Wing, Parkville 3040, Australia

<sup>b</sup>Australian Centre for Posttraumatic Mental Health, Department of Psychiatry, The University of Melbourne, Austin and Repatriation Medical Centre, Banksia Street, West Heidelberg, Victoria 3081, Australia

<sup>c</sup>Coordinated Healthcare, Northern Health, 1231 Plenty Road, Bundoora, Victoria 3083, Australia

<sup>d</sup>Department of Medicine, The University of Queensland, Department of Medicine, Princess Alexandra Hospital, Ipswich Rd., Woolloongabba, Queensland 4102, Australia

Received 6 August 2001; received in revised form 26 August 2002; accepted 2 October 2002

## Abstract

Measurement of Health-Related Quality of Life (HRQoL) of the elderly requires instruments with demonstrated sensitivity, reliability, and validity, particularly with the increasing proportion of older people entering the health care system. This article reports the psychometric properties of the 12-item Assessment of Quality of Life (AQoL) instrument in chronically ill community-dwelling elderly people with an 18-month follow-up. Comparator instruments included the SF-36 and the OARS. Construct validity of the AQoL was strong when examined via factor analysis and convergent and divergent validity against other scales. Receiver Operator Characteristic (ROC) curve analyses and relative efficiency estimates indicated the AQoL is sensitive, responsive, and had the strongest predicative validity for nursing home entry. It was also sensitive to economic prediction over the follow-up. Given these robust psychometric properties and the brevity of the scale, AQoL appears to be a suitable instrument for epidemiologic studies where HRQoL and utility data are required from elderly populations. © 2003 Elsevier Science Inc. All rights reserved.

**Keywords:** Elderly; Health-related quality of life; Questionnaires; Validation; Utility

## 1. Introduction

The health-related quality of life (HRQoL) of older people has become an important public health issue with the aging of populations in developed and developing countries. Many people are now living a longer life with less disability when compared with previous generations [1]. Although a variety of strategies have been developed to promote “healthy aging” [2] and many generic and disease-specific HRQoL instruments are available, few have been formally validated in populations of older people.

This situation is anomalous given the pressure on health care funding due to increased demand [3]. As the gap between funding availability and expectations widens it is important that programs are evaluated in terms of direct health status outcomes as well as economic evaluations that ex-

plore the efficiency (i.e., health gain for monetary value expended). This has led to increased demands for utility HRQoL instruments that have demonstrated responsiveness, reliability, and validity.

The *Assessment of Quality of Life* (AQoL) instrument [4] is a HRQoL questionnaire that has been scaled using multi-attribute utility theory and can therefore be used in cost-utility evaluations [5]. As with all generic utility instruments, the AQoL was designed to be used in the evaluation of a wide range of health interventions, from the medical and pharmacologic treatment of acute illness through to health promotion activities [4]. A principal difference between a generic utility instrument and a health status instrument, such as the SF-36 [6], is that the health states described in a utility instrument are weighted by preferences for being in the health state. The AQoL has been found to be as good or better than instruments commonly used in measuring the outcomes of people suffering a stroke [7], and it has been successfully used in studies of coordinated care [8], influenza [9], cochlear implants [10], population monitoring

\* Corresponding author. Tel.: +61-(03)-9342-8561; fax +61-(03)-9342-7632.

E-mail address: richardo@unimelb.edu.au (R.H. Osborne).

[11], and has been found to perform at least as well as other well-known utility instruments (e.g., Health Utilities Index-3, 15D, SF-6D, and EQ-5D) [12]. Although the AQoL has been designed to be an efficient measure within a variety of settings, astute researchers require evidence of efficient performance of new instruments within each setting.

The aim of the present study was to examine the psychometric properties of the AQoL and other instruments using data collected in a large sample of chronically ill elderly people participating in a randomized controlled trial of care coordination.

## 2. Methods

### 2.1. Participants

Participants were those participating in the North Eastern Coordinated Care Trial, a randomized controlled trial of care coordination versus usual care. The trial was one of nine trials commissioned by the Australian Commonwealth to explore an alternative form of service delivery, funding structure, and consequent health outcomes for people in the Australian community [13]. Eligibility criteria for this particular trial were designed to capture older people with chronic disease and/or dependencies who were living in the community. Dependencies were defined as being unable to perform at least two instrumental activities of daily living (IADL) or unable to perform at least one activity of daily living (ADL), or at risk of repeated hospitalization, or to have been hospitalized at least twice in the previous year. Between December 1997 to July 1998, 1,315 patients were referred to the trial. Most patients (87%) were referred through general practitioners (GPs) or other care organizations including community health centers (4%), acute hospitals (0.5%), and community services (1%). A further 7% were self-referred through media reports. Written consent was obtained at a face-to-face interview at patient's homes. Interpreters were arranged for non-English-speaking patients.

Of the original 1,315 referrals, 1056 (80%) participated in the trial. Thirty-two people withdrew prior to giving consent (including deaths and admissions to nursing homes). Seven potential participants consented, but withdrew consent prior to being randomized and were not included. Two hundred twenty clients did not proceed with the consent process, predominantly due to the recruitment cutoff date and/or incomplete documentation from their GP. Participants were randomized to receive either intervention (care coordination,  $n = 526$ ) or usual care ( $n = 530$ ), and completed self-report questionnaires at baseline and at six monthly intervals. At the third follow-up 827 assessments (78% of participants) were completed. The median (IQR) time in the trial was 18.5 (15.6–20.6) months. By trial end, 100 (9.5%) participants had entered a nursing home.

### 2.2. Chronic disease status and comorbidity

At the time of recruitment the subjects' GP recorded the primary chronic condition of each subject. The most com-

mon were: musculoskeletal disorders (274, 26%) including 158 with arthritis, circulatory system disorders (264, 25%) including 165 with a form of heart failure, respiratory system disorders (146, 14%) including 99 with chronic obstructive airways disease, mental disorders (136, 13%) including 64 with memory loss/dementia and 48 with anxiety, depression, or schizophrenia. GPs also recorded up to six specific comorbidities. The most common first comorbidity was circulatory disorder (333, including 140 with a form of heart failure and 112 with hypertension), followed by musculoskeletal disorder (252, including 130 with arthritis and 65 with degenerative disk disease) and 187 with a mental disorder (including 87 with depression, 44 with anxiety, and 44 with memory loss). For 25 subjects no comorbidities were recorded and for 987 subjects, 3 or more were recorded.

### 2.3. Measures

All instruments were administered at baseline and 6-month intervals except the OARS self-care scale, which was administered at baseline only.

The *Assessment of Quality of Life* (AQoL) instrument is a generic HRQoL utility instrument comprising five dimensions (Independent Living, Social Relationships, Physical Senses, Psychological Well-being, and Illness). Each scale provides weighted scores between 0.0 (scale worst health state) and 1.0 (scale best health state). The utility score ranges from  $-0.04$  (worst possible HRQoL state) to 0.00 (death equivalent HRQoL state) to 1.00 (full HRQoL), and is a weighted multiplicative model derived from the last four dimensions. The illness dimension, an indicator of healthcare resource consumption, is not used in the utility score. During construction the authors took care to (a) ensure wide coverage of the factors that impact on HRQoL, (b) meet standard psychometric requirements for reliable and valid measurement, (c) design the AQoL to be sensitive to a wide range of health states from very good to extremely poor health, and (d) construct an instrument that delivered psychometrically sound scores on the different dimensions of HRQoL [4,14]. A copy of the instrument can be obtained at <http://ariel.unimelb.edu.au/chpe/> (accessed August 2002).

The *SF-36* was used to measure the health status of participants [15,16]. The SF-36 is a widely used, extensively investigated, and validated instrument. It comprises eight health dimensions: Physical Functioning, Role Physical, Bodily Pain, General Health, Vitality, Social Function, Role Emotional, and Mental Health. Each of the eight dimensions is separately scored, using item weighting and additive scaling. Summed data are transformed onto a 0–100-point scale. These eight dimensions can be combined into two key health status measures; the Physical Component Summary (PCS index) and Mental Component Summary (MCS index). For computation of the PCS and MCS, each dimension score is weighted in a three-step process to produce a standardized T-score (where the population mean score is 50, SD = 10).

Australian norms are available and have been used in this report [17].

The *Older American Resources and Services* (OARS) instrument [18] was developed in the mid-1970s. It was designed to provide comprehensive assessments to reduce inappropriate institutionalization and to evaluate alternative care for the elderly. Like the SF-36, the OARS has been widely used. It comprises two parts: (a) the Multidimensional Functional Assessment Questionnaire (MFAQ), and (b) the Services Assessment Questionnaire. In the present study the Social Resources scale (nine items) and Self-care scale (15 items) from the MFAQ were administered.

#### 2.4. Data analysis

To assess the feasibility of the using the AQoL in an elderly population, we examined the frequency of floor effects (number of respondents with the lowest possible score) and ceiling effects (number of respondents with the highest possible score).

The convergent and divergent validity of the AQoL were assessed by comparison with the SF-36 and OARS. For convergent validity, if the AQoL and its subscales were valid measures of HRQoL, we would expect the strongest association to be with other scales measuring similar constructs (observed through a high correlation coefficient). These include the Independent Living scale of the AQoL and (a) the Social Resources and Self-care scales of the OARS and (b) the Physical Function scale of the SF-36. Conversely, for divergent validity, the weakest relationships should be observed between conceptually unrelated scales.

Construct validity was assessed using exploratory factor analysis (EFA) where the overall relatedness of all items was investigated via principal components analysis and, to determine whether the four hypothesized dimensions existed in the data, using varimax rotation.

Predictive validity is the ability of an instrument to predict future events or health states. We hypothesized that at baseline those people living in the community with the highest AQoL utility were less likely to enter a nursing home, hostel, or similar institution when compared with those with lower utility. In other words, those with the lowest initial HRQoL were more likely to decline in health such that they required institutional care. An instrument with predictive validity should be able to predict who enters and who does not enter institutional care. As a further test of predictive validity we also computed all health care costs over the length of the trial and matched these against AQoL baseline scores in the expectation that there would be a monotonic relationship between baseline utility score and future health care costs.

Sensitivity is the ability of an instrument to detect differences between groups. For this type of analysis two group differences were considered: the sensitivity of the instrument in being able to detect differences between those who stayed living in the community and those who entered a care institution (nursing home), and the effect of the care coordination intervention. The ability of the AQoL to detect

subgroup differences and treatment effects was compared with the other instruments using the relative efficiency (RE) statistic that is the ratio of the square of the *t*-statistic of the reference instrument (in our case, the AQoL) over the square of the *t*-statistic of the comparator instrument [19].

Responsiveness is the ability of a scale to detect change in an individual's score over time. In this study the participants whose health status declined such that they entered institutional care was used as the criterion of change in health. Change was assessed through comparison of baseline scores with the score prior to entering an institution. In a similar way to assessing sensitivity, the RE statistic was also used to assess the ability of the AQoL to detect such change within individuals.

Discriminative properties of the instruments were also compared with receiver operator characteristic (ROC) curves as proposed by Stucki et al. (1995) [20]. The ROC method provides a useful overview of the relationship between a measure and an external indicator of change [21]. For the current analysis, the data relating to entry to nursing home or remaining in the community (i.e. dichotomous external indicators) were used. The instrument that generates the largest area under the ROC curve is regarded as the most sensitive or responsive as an instrument with perfect discrimination has a area under the curve (AUC) of 1.0, whereas an instrument with no discriminatory power has an AUC = 0.50.

### 3. Results

The mean age of trial participants was 77 (SD = 9.7) years, 63% were female, 16% were born in countries other than Australia, 46% were married or living in de facto relationships, 42% were widowed, and 32.8% had private health insurance (see Table 1 for details).

The mean AQoL utility score was 0.33 (95%CI: 0.31–0.34). The Independent Living scale returned the lowest scale score (0.53, 0.51–0.54), followed by Psychological Well-being (0.75, 0.73–0.76) and Social Relationships (0.75, 0.73–

Table 1  
Demographic details of the study sample

		N	%
Age group	<60	61	5.8
	60 to 75	146	13.8
	70 to 75	175	16.6
	75 to 80	267	25.3
	80 to 85	235	22.3
	>85	172	16.3
Gender	Female	668	63.3
Language background	English	895	84.8
Marital Status	Never married	54	5.1
	Widowed/divorced/separated	510	48.3
	Married/de facto	491	46.5
	Missing	1	0.1
Health Insurance	Private	346	32.8
	Public only (Medicare)	708	67.2

0.77). Physical Senses contributed the least disutility (0.85, 0.84–0.86).

The mean (95% CI) score on the SF-36 PSC scale was quite low, 30 (29.5–30.6), whereas the MCS mean score was 46 (45.1–46.7). Given that the Australian norms for these scales are 50 (SD 10) the sample had substantially poor physical health (i.e., average physical health in the bottom 5% of the population), whereas the sample reported reasonable mental health status. The details are given in Table 2, along with the scores from the OARS scales.

Generally, the AQoL utility score demonstrated substantial spread across the entire scale range (Fig. 1). There were, however, some ceiling effects on two of the AQoL scales; for Social Relationships it was 22%, and 24% on the Physical Senses scale. Ceiling and floor effects were also observed on several of the SF-36 scales. For example, Role Physical and Role Emotion had between 27% and 55% floor and ceiling effects. Physical Function also demonstrated substantial floor effects (14%), whereas Bodily Pain and Social Function had substantial ceiling effects (17% and 31%, respectively). The OARS scales had extremely few floor or ceiling effects (all <6%). The details are given in Table 2.

### 3.1. Concurrent validity: associations with SF-36 and OARS

Conceptually related scales were more highly related than unrelated scales (see Table 3) suggesting convergent and divergent validity of the AQoL. The AQoL utility was moderately correlated with four of the eight SF-36 scales (Physical Function, Mental Health, General Health and Vi-

tal) and was least associated with Role Physical and Role Emotional. The AQoL utility score was also moderately correlated with the SF-36 MCS and PCS. The AQoL was, however, highly correlated with OARS self-care.

When AQoL scales were examined, the Independent Living scale was highly correlated with the SF-36 Physical Function and OARS self-care. This was expected, because this scale reflects the ability of a person to live independently, including aspects of activities of daily living. The Social Relationships scale, which reflects relationships with others, was only weakly associated with the SF-36 Social Function scale, although moderately related to the SF-36 Mental Health and Physical Function scales. It was also moderately correlated with the OARS Self-care, but weakly associated with the social resources. The Physical Senses scale, was weakly associated with OARS self-care. Finally, the Psychologic Well-being scale (covering sleep, pain, and general health) was highly correlated with the SF-36 Bodily Pain, Mental Health, General Health and Vitality scales, and also with the SF-36 MCS.

### 3.2. Construct validity

The factorial structure of the AQoL was examined using EFA. The results showed that in an unconstrained varimax rotation four factors were identified (i.e., four factors with eigenvalues greater than 1.0). Each factor represented one of the four original dimensions of the AQoL contributing to the utility score, each with three items. The average loading on the expected factors was 0.67 compared with the average loading of items on other factors was 0.14. These figures are consistent with the requirements for stable and reliable factor analysis [22].

Only four items crossloaded >0.30. Item 9 *Role in health and family* from the Social Relationships scale loaded on the Independent Living scale. Two other items loaded on Social Relationships scale, Item 12 *Communication with others* (from the Physical Senses scale) and Item 14 *Anxiety and depression* from the Psychological Well-being scale. Item 12 *Communication with others* also loaded weakly on the Psychological Well-being scale.

The internal consistency of the subscales was ranged from  $\alpha = 0.43$  for the Physical Senses scale, 0.53 for the Social Relationships scale, 0.52 for the Psychological Well-being scale and 0.76 for the Independent Living scale. The overall AQoL reliability was 0.73.

### 3.3. Sensitivity

To examine the sensitivity of the AQoL, we used the relative efficiency (RE) statistic that allows comparison of how well scales detect treatment effects. Table 4 shows the intervention group vs. the control group mean scale changes from baseline to the third follow-up. When the AQoL utility was referenced to “1.0” it was found to have a higher RE than each of the SF-36 scales except Bodily Pain, and was about the same as the MCS. The AQoL utility score was more efficient than the individual AQoL scales, except Independent living.

Table 2  
Baseline summary scores including floor and ceiling effects of scales

	Mean	SD	Floor <sup>a</sup>	Ceiling <sup>b</sup>
<b>AQoL</b>				
AQoL utility	0.33	0.25	0.1%	0.3%
Independent Living	0.53	0.30	5.3%	6.8%
Social relationships	0.75	0.27	0.4%	21.7%
Physical senses	0.85	0.16	0.0%	24.2%
Psychological wellbeing	0.75	0.21	0.9%	4.4%
<b>SF-36</b>				
Physical function	24.13	22.75	14.4%	1.0%
Role physical	41.57	43.19	42.4%	29.6%
Bodily pain	49.27	30.58	5.4%	17.0%
General health	41.22	23.79	4.1%	3.0%
Vitality	36.95	22.30	5.5%	2.0%
Social function	61.68	33.78	7.2%	30.5%
Role emotional	63.66	43.82	27.0%	55.0%
Mental health	63.48	23.09	9.0%	2.9%
PCS	30.03	9.47	NA	NA
MCS	45.9	12.68	NA	NA
<b>OARS</b>				
Self-care	2.83	1.34	5.6%	0.0%
Social resources	2.40	1.14	0.0%	0.0%

<sup>a</sup> Floor effects: percentage of subjects with lowest possible score on scale.

<sup>b</sup> Ceiling effects: percentage of subjects with highest possible score on scale. NA: the SF-36 scoring algorithm precludes ceiling and floor effects.

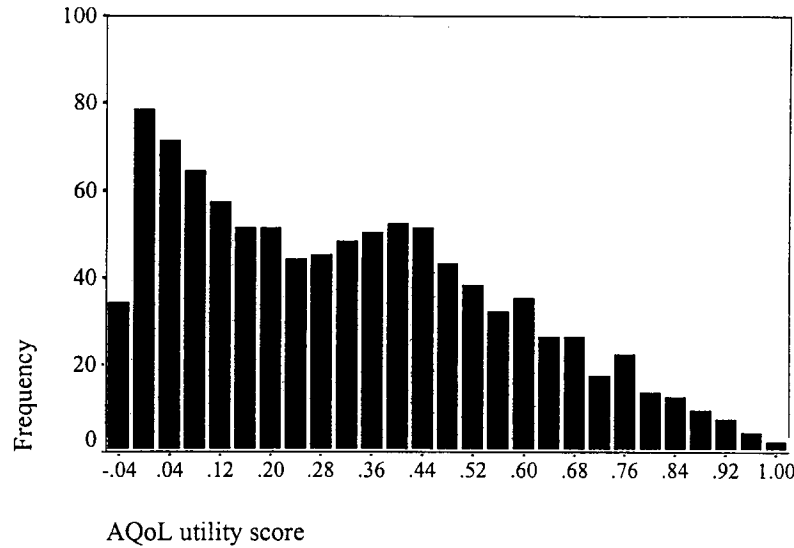


Fig. 1. Frequency distribution of AqoL utility score in chronically ill community-dwelling elderly.

The sensitivity of the various instruments was also tested by comparing the RE of the scales in detecting baseline differences between participants who remained living in the community vs. those who became institutionalized during the trial (Table 6). Relative to the AqoL (referenced to 1.0), the OARS Self-care scale was more sensitive. The Independent Living scale was more sensitive than the AqoL utility score, and the Social Relationships scale was equivalent. The SF-36 summary scales were less sensitive and the only subscale that was similar to the AqoL was Physical Function.

### 3.4. Responsiveness

Responsiveness is the ability of an instrument to detect change over time. A relatively large change in health status (and therefore dimensions of HRQoL) must occur in participants who leave their community dwelling to live in a nursing home. Table 6 shows that several scales detected changes and presents the RE for each scale compared with the AqoL utility score referenced to 1.0. The AqoL has the highest RE when compared with the other instruments.

Table 3  
Spearman's correlation coefficients between scales

	Independent living	Social relationships	Physical senses	Psychological well-being	AqoL utility
<b>AqoL</b>					
Social relationships	0.43				
Physical senses	0.24	0.24			
Psychological well-being	0.19	0.32	0.14		
AqoL utility	0.79	0.68	0.43	0.55	
<b>SF-36</b>					
Physical function	0.64	0.41	0.15	0.31	0.62
Role physical	0.09	0.10	0.04	0.25	0.19
Bodily pain	0.11	0.18	0.04	0.62	0.34
General health	0.30	0.35	0.09	0.47	0.43
Vitality	0.35	0.38	0.14	0.51	0.50
Social function	0.28	0.26	0.13	0.33	0.38
Role emotion	0.08	0.21	0.05	0.32	0.22
Mental health	0.26	0.41	0.16	0.53	0.45
PCS	0.30	0.17	0.03	0.39	0.37
MCS	0.23	0.39	0.14	0.48	0.41
<b>OARS</b>					
Self-care	-0.82	-0.40	-0.32	-0.11	-0.68
Social resources	0.03	-0.31	-0.06	-0.26	-0.20

Sample size for correlation between summary scales (AqoL, PCS/MCS)  $N = 1000$ , and between each subscale;  $N = 1017$  to  $1052$ . Using Cohen's convention [27], the magnitude of the correlations can be viewed as follows;  $>0.5$  large,  $0.5-0.3$  moderate,  $<0.3$  small.

Table 4  
Efficiency of scales to detect small treatment effects (i.e., sensitivity) of people randomized to receive care coordination (intervention) or usual care (control)

	Group	Mean <sup>a</sup> change	SD	t-test		Relative Efficiency <sup>b</sup>
				t-statistic	p	
<b>AQoL</b>						
AQoL utility	Intervention	0.02	0.21	1.45	0.15	1.00
	Control	-0.01	0.22			
Independent living	Intervention	-0.00	0.22	2.13	0.03	2.17
	Control	-0.04	0.22			
Social relationships	Intervention	-0.00	0.24	0.15	0.88	0.01
	Control	-0.00	0.29			
Physical senses	Intervention	0.02	0.12	-0.31	0.76	0.05
	Control	0.02	0.14			
Psychological well-being	Intervention	0.03	0.22	1.20	0.23	0.69
	Control	0.01	0.22			
<b>SF-36</b>						
Physical function	Intervention	-4.11	19.28	-1.02	0.31	0.50
	Control	-2.55	18.56			
Role physical	Intervention	4.38	55.31	0.21	0.84	0.02
	Control	3.47	53.13			
Bodily pain	Intervention	8.87	30.65	2.13	0.03	2.18
	Control	3.76	29.07			
General health	Intervention	-0.75	18.24	-0.03	0.98	0.00
	Control	-0.71	19.88			
Vitality	Intervention	-2.04	21.23	0.57	0.57	0.15
	Control	-3.07	23.65			
Social function	Intervention	8.67	38.58	0.29	0.77	0.04
	Control	7.75	40.34			
Role emotional	Intervention	15.49	52.41	0.71	0.48	0.24
	Control	12.45	53.68			
MCS	Intervention	3.80	20.29	1.41	0.16	0.95
	Control	1.37	22.52			
PCS	Intervention	-0.47	9.87	-0.01	0.99	0.00
	Control	-0.46	9.26			
<b>OARS</b>						
Social resources	Intervention	-0.22	1.08	-0.93	0.35	0.41
	Control	-0.13	1.13			

Listwise exclusion of missing variables (297 intervention and 324 control subjects).

<sup>a</sup> Mean change = baseline score less follow-up score (positive score indicates improvement in health status from baseline to follow-up).

<sup>b</sup> RE. Relative Efficiency, where the AQoL is referenced to 1.0. A value higher than 1.0 indicates that this scale is better (more efficient) than the AQoL in detecting differences and, conversely, a score lower than 1.0 indicates that the scale is not as good (less efficient) compared with the AQoL.

### 3.5. Predictive validity

We used AQoL baseline scores to predict future health care costs. To predict these costs we recoded baseline utility scores into quintiles and, using the trial's tracking of all health costs over the trial period, mapped these costs against the quintiles. The results are presented in Figure 2. This shows that for those in the highest quintile at baseline the average costs were AUD\$7,034 which monotonically rose to about AUD\$12,883 for those in the lowest quintile.

## 4. Discussion

This study examined several aspects of the psychometric performance of the AQoL in elderly community-dwelling adults. The practicality, validity, sensitivity, responsiveness, and predictive validity were explored and the results indicate that the AQoL possesses acceptable psychometric validity. When compared with other commonly used instru-

ments such as the SF-36 or the OARS, the AQoL appears to be a suitable instrument for assessing health-related outcomes in the chronically ill elderly.

For the AQoL utility score there was no evidence of either floor or ceiling effects, compared with five of the eight SF-36 scales. However, two of the four AQoL scales demonstrated ceiling effects. Almost a quarter of respondents scored the highest score in the Social Relationships and Physical Senses scales, that is, their HRQoL in these domains was not compromised, and probably reflects the inclusion criteria that focused on physical disabilities.

In our chronically ill sample, floor effects did not occur, thus indicating that the AQoL provides enough sufficiently "bad" health states for respondents to place themselves. This finding suggests that the AQoL provides the opportunity to detect change in individuals, as they are able to progress either from a "poor" HRQoL state through to a "good" HRQoL state, or to map progressive deterioration.

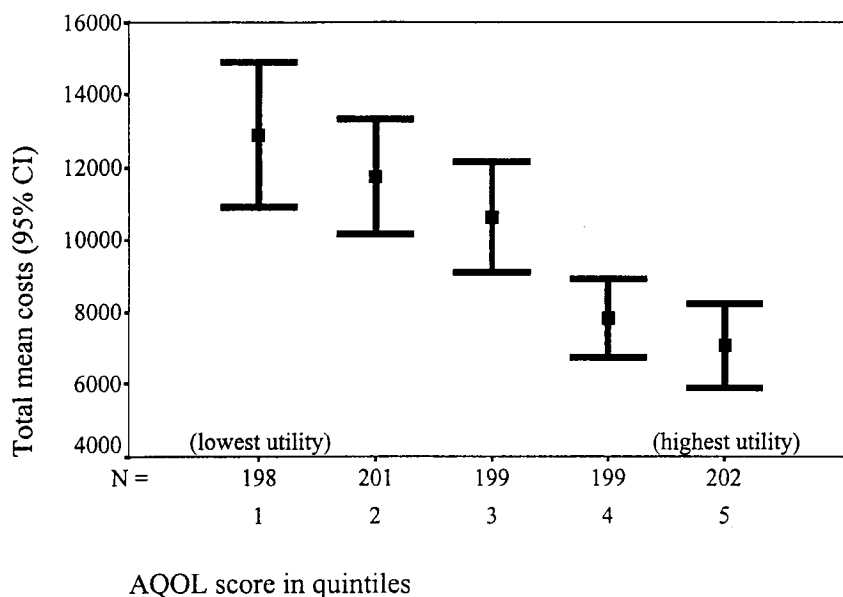


Fig. 2. Prediction of future health care costs from baseline AQoL utility scores.

The AQoL was designed to reflect overall HRQoL; therefore, substantial associations with various SF-36 scales should reflect this (i.e., concurrent validity). Several strong associations were noted (the highest was with the SF-36 Physical Function and Vitality) and moderate correlations with the summary indices. A high correlation between the AQoL and the OARS self-care scale was also found. The individual dimensions of the AQoL also have substantial concurrent validity. For example, the Independent Living scale is most similar to the SF-36 Physical Function. A remarkably high correlation was found with the 15-item OARS Self-care scale ( $r = 0.82$ ) suggesting that the three-item Independent Living scale is measuring almost precisely the same construct, but with only one-fifth of the items. The OARS was designed to assist in identifying factors that lead to inappropriate institutionalization [18]. It is conceivable that this scale is also identifying those aspects of life that strongly contribute to HRQoL in the elderly. The Psychological Well-being scale (with items that cover pain, general well-being, and sleep) is moderately associated with appropriate SF-36 subscales (Vitality, Bodily Pain, General Health) and with the MCS. The most similar comparator for the Social Relationships scale is the OARS Social resources scale; however, the correlation was weak ( $r = 0.22$ ), which may reflect the way the OARS scale is weighted such that about 50% of the scale is concerned with having someone to help, and a further 30% covers seeing people frequently enough to prevent feelings of loneliness. The AQoL Social Relationships scale covers concepts such as quality of and role in relationships as well as friendship/isolation. For the Physical Senses scale there were no direct comparator scales. Overall, these data indicate that the AQoL has reasonable concurrent validity.

The construct validity of the AQoL appears to be strong as the original four dimension factorial structure was reproduced. The factor analysis was unconstrained with only minor item crossloading. Given that the original construction sample of the AQoL was a mix of community and hospital adults of varying ages [4], consistent factor structures in disparate samples suggest that the conceptual structure (construct validity) is robust.

The Cronbach reliability estimate of the AQoL instrument ( $\alpha = 0.73$ ) was within the accepted range appropriate for a nondiagnostic generic instrument [23,24]. The reliability of some subscales were low, reflecting their broad content and the relatively few items in each ( $\alpha$  is a function of both the similarity of items and the number of items in a scale [23,24]).

The efficacy of each instrument in detecting group differences or longitudinal changes was compared. Three commonly used data structures were considered.

A comparison of two groups expected to have different HRQoL (i.e., the baseline characteristics of those who subsequently entered a nursing home versus those who remained in the community, Table 5). In this analysis the AQoL had relatively high sensitivity; the only scale with a higher RE or ROC was OARS self-care, which was specifically designed to evaluate unmet needs in the elderly. This is not a surprising result, as disease-specific instruments such as the OARS usually possess great sensitivity in the environment for which they are designed [25].

A comparison of those who received coordinated care vs. those who received usual care was also undertaken (Table 4). The differences in outcomes were known to be very small or nonexistent [13]. Only two subscales tended to demonstrate larger effects than the AQoL utility, namely the

Table 5

Efficiency of scales to detect baseline differences (sensitivity) between people who remain living in the community vs. those who subsequently enter an institution

	Group	Mean (SD) baseline score		t-test		RE <sup>a</sup>	ROC <sup>b</sup>
				t-statistic	P		
<b>AQoL</b>							
AQoL utility	Institution	0.26	(0.22)	−3.18	0.002	1.0	0.58**
	Community	0.34	(0.26)				
Independent living	Institution	0.46	(0.26)	−3.49	0.001	1.2	0.60**
	Community	0.55	(0.29)				
Social relationships	Institution	0.69	(0.31)	−3.07	0.002	0.9	0.56*
	Community	0.77	(0.26)				
Physical senses	Institution	0.83	(0.17)	−2.03	0.043	0.4	0.56*
	Community	0.86	(0.16)				
Psychological wellbeing	Institution	0.77	(0.21)	1.45	0.146	0.2	0.45*
	Community	0.75	(0.21)				
<b>SF-36</b>							
Physical function	Institution	19.09	(20.08)	−2.98	0.003	0.9	0.59**
	Community	25.35	(23.12)				
Role physical	Institution	41.61	(42.04)	0.57	0.571	0.0	0.48
	Community	39.39	(42.62)				
Bodily pain	Institution	53.84	(33.70)	2.32	0.020	0.5	0.45*
	Community	47.4	(29.50)				
General health	Institution	42.66	(25.12)	0.79	0.428	0.1	0.49
	Community	40.94	(23.28)				
Vitality	Institution	32.75	(20.40)	−2.24	0.025	0.5	0.56*
	Community	37.29	(22.25)				
Social function	Institution	59.67	(34.00)	−0.51	0.608	0.0	0.51
	Community	61.26	(33.57)				
Role emotional	Institution	64.23	(43.66)	0.61	0.542	0.0	0.49
	Community	61.75	(44.22)				
Mental health	Institution	60.91	(22.91)	−1.46	0.144	0.2	0.54
	Community	64.02	(23.07)				
PCS	Institution	30.26	(9.34)	0.6	0.552	0.0	0.52
	Community	29.74	(9.42)				
MCS	Institution	44.82	(12.74)	−0.95	0.342	0.1	0.48
	Community	45.94	(12.75)				
<b>OARS</b>							
Self-care	Institution	3.39	(1.25)	6.05	<0.001	3.6	0.67**
	Community	2.69	(1.27)				
Social resources	Institution	2.61	(1.24)	2.35	0.019	0.5	0.56*
	Community	2.36	(1.12)				

<sup>a</sup> RE; Relative efficiency, where the AQoL is referenced to 1.0. A value higher than 1.0 indicates that this scale is better (more efficient) than the AQoL in detecting differences and, conversely, a score lower than 1.0 indicates that the scale is not as good (less efficient) compared with the AQoL.

<sup>b</sup> Area under receiver operating characteristic (ROC) curves; \* P < .05, \*\* P < .01 indicate significance tests showing area under the curve statistically significantly greater than 0.5.

<sup>c</sup> Listwise exclusion of missing variables resulted in 137 institutionalized subjects and 848 subjects remaining in the community.

AQoL Independent Living scale and the SF-36 Bodily Pain scale (RE = 2.2 for both). The SF-36 MCS demonstrated similar efficiency to the AQoL.

The final data structure was longitudinal change (change from baseline to third follow-up for those who subsequently entered a nursing home). In this analysis the AQoL was the most responsive scale.

The predictive power of the AQoL also appears to be strong. The AQoL predicted health care costs following baseline. Those in the lowest quintile with the poorest HRQoL at baseline consumed 1.9 times the health care resources of those in the highest quintile.

The AQoL was designed as a utility measure for use in economic evaluations. One of the advantages of utility instruments is that they provide estimates on a life–death scale based on community preferences (where 0.00 reflects HRQoL states valued as being equivalent to death, and 1.00 HRQoL states valued as being equivalent to full health). Utility scales are designed to have equal interval properties reflecting a trade-off between HRQoL and life-length, that is of a utility score of 0.5 over 2 years is equivalent to a loss of 1 life year ( $0.5 \times 2$  years) [26]. Assuming this strong interval property for the AQoL, and using data from the present study, the intervention group received a 0.03 utility benefit from coordinated care when com-

Table 6

Efficiency of scales to detect changes from baseline to third follow-up (responsiveness) in those who subsequently entered an institution compared with those who stayed living in the community

Scale	Place of residence <sup>a</sup>	Mean (SD) change from baseline <sup>b</sup>		t-test		RE <sup>c</sup>	ROC area <sup>d</sup>
				t-statistic	P		
<b>AQoL</b>							
AQoL utility	Institution	-0.13	(0.22)	3.94	<0.000	1.0	0.67**
	Community	0.01	(0.21)				
Independent living	Institution	-0.12	(0.24)	2.98	0.003	0.6	0.60**
	Community	-0.02	(0.22)				
Social relationships	Institution	-0.10	(0.31)	2.30	0.022	0.3	0.60*
	Community	-0.00	(0.27)				
Physical senses	Institution	-0.02	(0.16)	1.96	0.050	0.2	0.58*
	Community	0.02	(0.13)				
Psychological wellbeing	Institution	-0.05	(0.23)	2.13	0.034	0.3	0.60*
	Community	0.02	(0.21)				
<b>SF-36</b>							
Physical function	Institution	-11.00	(20.95)	2.68	0.008	0.5	0.62*
	Community	-2.76	(18.67)				
Role physical	Institution	-3.13	(56.67)	0.85	0.396	0.0	0.55
	Community	4.39	(53.98)				
Bodily pain	Institution	1.33	(28.34)	1.07	0.287	0.1	0.57
	Community	6.54	(30.02)				
General health	Institution	-7.83	(22.20)	2.44	0.015	0.4	0.62*
	Community	-0.24	(18.79)				
Vitality	Institution	-11.04	(24.30)	2.47	0.014	0.4	0.59
	Community	-2.00	(22.49)				
Social function	Institution	-4.06	(48.07)	2.03	0.042	0.3	0.58
	Community	9.04	(38.72)				
Role emotional	Institution	10.00	(57.09)	0.48	0.631	0.0	0.51
	Community	14.17	(52.81)				
Mental health	Institution	0.12	(14.22)	1.36	0.175	0.1	0.54
	Community	2.97	(12.71)				
MCS	Institution	-0.13	(24.70)	0.81	0.420	0.0	0.53
	Community	2.71	(21.27)				
PCS	Institution	-3.44	(10.37)	2.04	0.041	0.3	0.59*
	Community	-0.26	(9.47)				
<b>OARS</b>							
Social resources	Institution	0.12	(0.95)	-1.72	0.086	0.2	0.59
	Community	-0.19	(1.12)				

<sup>a</sup> Subjects without baseline and third follow-up data excluded. Remaining data included 40 subjects subsequently entering a nursing home and 581 subjects remaining in the community.

<sup>b</sup> RE; Relative efficiency, where the AQoL is referenced to 1.0. A value higher than 1.0 indicates that this scale is better (more efficient) than the AQoL in detecting differences, conversely, and a score lower than 1.0 indicates that the scale is not as good (less efficient) compared with the AQoL.

<sup>c</sup> Area under receiver operating characteristic (ROC) curves; \*  $p < .05$ , \*\*  $p < .01$  indicate significance tests showing area under the curve statistically significantly greater than 0.5.

<sup>d</sup> Positive value implies increase in health from baseline.

pared with the control group. For a trial of 18 months with 500 people the gain in Quality Adjusted Life Years or QALY is  $0.03 \times 1.5 \times 500 = 22.5$  life years.

The findings from this study suggest that the AQoL is a responsive and sensitive instrument. At the psychometric level it is at least as good as the comparators used in this study. In addition, it was sensitive to simple tests of economic prediction.

Apart from reasons of measurement accuracy and validity, the findings from this validity study are important for a pragmatic reason. The sensitivity of instruments is critical for research and evaluation design as this affects the sample size. A study using a blunt (insensitive) instrument requires

a larger sample size to detect a statistically significant difference than a study using a precise (sensitive) instrument. This affects the power and cost of a study. The implication is that if small effects from an intervention are expected, responsive instruments are necessary to detect changes over time. Given the robust psychometric properties, the AQoL appears to be a suitable utility instrument for clinical and epidemiological studies in the elderly.

## References

- [1] Greengross S, Murphy E, Quam L, Rochon PA, Smith R. Aging: a subject that must be at the top of world agenda. *BMJ* 1997;315:1029–30.

- [2] Khaw KT. Healthy aging. *BMJ* 1997;315:1090–6.
- [3] Spillman BC, Lubitz J. The effect of longevity on spending for acute and long-term care. *N Engl J Med* 2000;342:1409–15.
- [4] Hawthorne G, Richardson J, Osborne R. The Assessment of Quality of Life (AQoL) Instrument: a psychometric measure of health related quality of life. *Qual Life Res* 1999;8:209–24.
- [5] Dolan P. The measurement of health-related quality of life for use in resource allocation decisions. In: Culyer AJ, Newhouse JP, editors. *Handbook of health economics*. Amsterdam: Elsevier; 2000.
- [6] Ware J, Kosinski M, Keller S. SF-36 Physical and Mental Health Summary Scales: a user's manual. Boston: The Health Institute, New England Medical Centre; 1994.
- [7] Sturm J, Osborne R, Dewey H, Donnan G, Macdonell R, Thrift A. Brief comprehensive quality of life assessment: the Assessment of Quality of Life (AQoL) instrument in the North East Melbourne Stroke Incidence Study (NEMESIS). *Stroke* 2002;33:2888–94.
- [8] Segal L, Day N, Day S, Dunt D, Piterman H, Robertson I, Hawthorne G. Evaluation of the Southern Health Care Network Coordinated Care Trial: summary report. Melbourne: Centre for Health Program Evaluation. Working Paper 122; 2000. <http://chpe.buseco.monash.edu.au/pubs/wp122.pdf> (Accessed August 2002).
- [9] Osborne RH, Hawthorne G, Papanicolaou M, Wegmueller Y. Measurement of rapid changes in health-related quality of life (HQoL) in people with influenza. *J Outcomes Res* 2000;4:15–30.
- [10] Hogan A, Hawthorne G, Kethel L, Giles E, White K, Stewart M, Plath B, Code C. Health-related quality-of-life outcomes from adult cochlear implantation: a cross-sectional study. *Cochlear Implants Int* 2001;2:115–28.
- [11] Epidemiology Department. South Australian Health Omnibus Survey. Adelaide: South Australian Health Commission; 1998.
- [12] Hawthorne G, Richardson J, Day NA. A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments. *Ann Med* 2001;33:358–70.
- [13] Commonwealth Department of Health and Aged Care. The Australian Coordinated Care Trials: final technical national evaluation report on the first round of trials. Canberra: Commonwealth of Australia; 2001 <http://www.health.gov.au/hsdd/primcare/acoorcar/pubs/acct/index.htm> (Accessed August 2002).
- [14] Hawthorne G, Richardson J, Day N, McNeil H. Using the "Assessment of Quality of Life" (AQoL) instrument. Melbourne: CHPE technical report 12; 2000 <http://chpe.buseco.monash.edu.au/pubs/tr12.pdf> (Accessed August 2002).
- [15] McCallum J. The SF-36 in an Australian sample: validating a new, generic health status measure. *Aust J Public Health* 1995;19:160–6.
- [16] Ware JE, Snow KK, Kosinski M, Gandek B. SF-36 health survey: manual and interpretation guide. Boston, MA: The Health Institute, New England Medical Centre; 1993.
- [17] Australian Bureau of Statistics. National Health Survey: Cat 4399.0; 1995.
- [18] Fillenbaum GG, Smyer MA. The development, validity, and reliability of the OARS multidimensional functional assessment questionnaire. *J Gerontol* 1981;36:428–34.
- [19] Fayers P, Machin D. Quality of life: assessment, analysis and interpretation. Chichester: Wiley; 2000.
- [20] Stucki G, Liang MH, Fossel AH, Katz JN. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *J Clin Epidemiol* 1995;48:1369–78.
- [21] Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol* 2000;53:459–68.
- [22] Rummel R. Applied factor analysis. Evanston: Northwestern University Press; 1970.
- [23] Cronbach L. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297–334.
- [24] Cortina J. What is coefficient alpha? Examination of theory and applications. *J Appl Psychol* 1993;78:98–104.
- [25] Vickrey BG, Hays RD, Genovese BJ, Myers LW, Ellison GW. Comparison of a generic to disease-targeted health-related quality-of-life measures for multiple sclerosis. *J Clin Epidemiol* 1997;50:557–69.
- [26] Richardson J. Cost utility analysis: what should be measured? *Soc Sci Med* 1994;39:7–21.
- [27] Cohen J. Statistical power analysis for the behavioural sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum; 1988.