Farmeconomia. Health economics and therapeutic pathways 2012; 13(2): 69-82

ORIGINAL RESEARCH

The effect of obesity upon Health Related Quality of Life (HRQoL). A comparison of the AQoL-8D and SF-36 instruments

Munir A Khan [1], Jeff Richardson [1], Paul O'Brien [2]

[1] Centre for Health Economics, Monash University, Clayton Vic, Australia

[2] Centre for Obesity Research & Education, Monash University, Clayton, VIC, Australia

ABSTRACT

OBJECTIVES: The objective of this study was to describe and measure the loss of health related quality of life (HRQoL) associated with obesity using two generic instruments. The first of these, the SF-36, is the most widely used and validated HRQoL instrument worldwide. However, it does not provide utility weights and cannot be used to measure quality adjusted life years (QALYs), an increasingly common unit for comparing the effect of health states in economic evaluation studies. The second, the AQoL-8D, is a multi-attribute utility (MAU) instrument which was developed to increase sensitivity of previous MAU instruments to psycho-social dimensions of a health state and to allow the calculation of QALYs. Since the two instruments differ, an important additional objective of the study was to determine the validity of the AQoL-8D as judged by the SF-36, and therefore the confidence which might be placed upon its use in the context of obesity.

METHODS: Data were obtained from patients waiting for bariatric surgery who had completed both the SF-36 and AQoL-8D quality of life instruments and a general questionnaire including height, weight, demographic and socio-economic information. For comparative purposes, scores were standardized using results from a representative sample of the general population. The content validity of the AQoL-8D was assessed by comparing it with the dimension scores from the SF-36 and the summary component (physical and mental) scores. Overall scores from the SF-36 and AQoL-8D instruments were regressed upon patient BMI and the results from the AQoL-8D used to estimate the effect of overweight and obesity upon utility and lost QALYs.

RESULTS: The comparison of the instruments indicated that the AQoL-8D has good convergent, concurrent and content validity. Using both instruments, obesity was significantly associated with lower scores for 14 of their 16 dimensions. AQoL-8D, in particular, identified a significant decrease in psychological and social health with happiness, self-worth, coping and mental health, all being poorer than in the control group. Regression results implied that a change in BMI from 30 to 50 decreases utility by 0.12 or by 13.8 percent of the average utility for a 25-35 year old.

CONCLUSION: AQoL-8D is a valid measure of utility in the context of obesity. Its inclusion of psycho-social effects significantly increases the measured adverse effects of obesity.

Keywords

Quality of life; Obesity; AQoL-8D; SF-36; Bariatric surgery

INTRODUCTION

The prevalence of obesity is rising sharply. By 2008, 3.71 million Australians or 17.5 percent of the population were estimated to be obese and three in five adults were either overweight or obese [1]. Rates were similar across Australian states but obesity was higher for males than females (18.5 versus 16.5 percent respectively). The trend is expected to continue and to follow the US trend where, in 2011, obesity varied from a low of 20.1 percent in Colorado to a high of 34.3 percent in West Virginia [2]. More generally, obesity rates in the western world have increased substantially over the past decade, and have affected most population groups, regardless of age, race, and income or education level. Evidence from OECD countries indicates that obesity tends to be more common among individuals in disadvantaged socio-economic groups, particularly among women [3].

Numerous studies have shown an association between obesity, defined as a BMI over 30, and the appearance or exacerbation of a variety of health problems including cardiovascular diseases, diabetes, sleep apnea, some forms of cancer, depression and a general impairment of the quality of life [4-6]. Obese individuals have an elevated risk of death compared with normal weight individuals or

Corresponding author
Prof. Jeff Richardson
Centre for Health Economics, Level 2, Building 75, Monash University, Clayton Vic 3800 Australia
Ph: +613 9905 0754;
Fax: +613 9905 8344
jeffrey.richardson@monash.edu

Disclosure
The authos have no financial competing interest to declare

those with a BMI between 18.5 and 25 [7]. As a consequence, weight related ill health and associated workforce absenteeism result in a high economic cost [8]. In the US 9 percent of health expenditures are associated with obesity and overweight [9] and for Canada and the USA the annual economic cost, including excess mortality and disability, has been estimated as $300 billion in 2009 [10]. A review of European studies suggests a total economic burden ranging from 0.09% of GDP in France in 1992 to 0.6% in Germany in 1990 and 0.6% in Greece in 2002 [11]. The latter results are dated and the cost has undoubtedly risen with increasing rates of obesity in the past one and a half decades. A recent UK study estimates the cost of obesity to the National Health Scheme and to the broader economy through lost productivity to be £4.2 and £15.8 billion respectively in 2007 [12].

In Australia Colagiuri, Lee et al. [13] have estimated expenditures of $21 billion or 1.7% of GDP arising from overweight and obesity. This includes health care related costs and transfer payments arising from government pensions and other payments but does not include premature mortality or productivity losses. Excluding transfers, but including productivity, career and health-care costs, is an estimated cost to the economy of $8.3 billion [1]. However, these data only relate to a partial estimate of the weight related burden of disease. They include a varying mix of costs, social transfers and the burden from premature mortality but omit the individual burden of a reduced quality of life (QoL). Reflecting measurement difficulties, there have been few attempts to quantify this in units which allow comparison with the burden of other health states.

Measuring health related quality of life (HRQoL)

Three types of questionnaire-based instrument may be used to measure the effects of a health state upon QoL. First, disease-specific instruments have been designed to focus upon health states which are relevant for a particular disease (for reviews, see Bowling [14] and McDowell [15]). For example the Impact of Weight on Quality of Life (IWQoL), was developed for use with individuals at all stages of obesity [16]. However, there is evidence to suggest that this instrument may be more relevant for morbidly obese individuals than for those with moderate obesity [17-19]. More generally, instruments in this first group are limited to the disease area for which they were designed and comparisons across disease groups are generally invalid.

This problem is overcome by use of a second, broad based generic, instrument such as the WHO Quality of Life (WHOQoL) scale [20], the Sickness Impact Profile [21] and the Nottingham Health Profile [22]. These were designed to include the major elements and dimensions of the quality of life implying that in principle they satisfactorily measure the majority of health states. This facilitates comparison and, additionally, increases the likelihood of detecting unexpected effects which may not be measured by a disease-specific instrument. The disadvantage of this second group of instruments is that the scores have no intrinsic meaning: they are index numbers which are difficult to integrate into an economic evaluation study which compares costs and benefits or seeks to measure the burden of disease.

For these tasks economists have created a third measurement tool, the multi attribute utility (MAU) instrument. This is similar to a generic instrument as it is based upon a generic "descriptive system", a set of items (questions and response categories) which describe the main dimensions and elements of the HRQoL. Their distinctive feature is that, additionally, they have a set of importance weights or a utility formula which reduces the responses to a single index of utility. This measures the strength of preference for a health state. Generic non-utility instruments may also be reduced to a single number but, by contrast with MAU instruments, this is obtained from unweighted responses; i.e. each response is given equal weighting and the final average score does not purport to have an independent meaning.

Utilities are measured on a (1.00-0.00) scale where 1.00 represents best health and 0.00 represents death. The preference number (utility) on this scale therefore indicates the strength of preference for quality versus quantity of life. This permits the calculation of quality adjusted life years (QALYs) as the product of life years and the index number (for example, 10 years with a utility index of 0.85 produces 8.5 QALYs). This is the unit of benefit used in cost utility analysis and, in principle, may be used to measure the HRQoL component of the burden of disease.

In practice both "generic" (i.e., unweighted generic) and MAU instruments differ in their construction and despite the claim that both are "generic" they contain different dimensions and items and produce different scores. In the two major comparative studies of MAU instruments to date, an average of 44% and 53% of the variation in utility in the instruments was not explained by other MAU instruments, despite purporting to measure the same quantity (utility) and all being ad-

ministered to the same individuals at the same point in time [23,24]. Consequently, good practice requires the validation of instruments in the different contexts in which they are to be used.

The objective of the present study was to describe the loss of HRQoL associated with obesity using both a generic and a MAU instrument. The generic instrument, the SF-36, is the most commonly used and validated of such instruments and has been widely used in the context of obesity [25]. The MAU instrument, the Assessment of Quality of Life (AQoL)-8D, produces utility scores which permit an estimate of the effect of obesity upon utility.

The AQoL-8D is a new instrument. Like the earlier AQoL instrument it was developed as a direct response to the problem of content invalidity of previous instruments, at least for some disease states. AQoL-8D is the largest MAU instrument to date, defining 2.4×10^{23} health states. This does not ensure content validity but, in contrast to non-AQoL MAU instruments, it was constructed using psychometric methods designed to achieve content validity [26]. Despite this, confidence in content validity must be evidence-based and testing of the AQoL-8D to date has been limited [27]. A second focus of the present paper is therefore the empirical evidence relating to the validity of the AQoL-8D in the context of obesity using the SF-36 and self-reported health to test the instrument.

The study employed data from both the general population and from obese patients on a waiting list for bariatric surgery. The data collection and their analysis are described in Section 2 which also describes and contrasts the two instruments. Results from the comparison of the instruments are presented in Section 3. In the discussion section the association between BMI and utility measured using the AQoL-8D instrument is compared with the results found in previous studies.

DATA, INSTRUMENTS AND ANALYSES

Study population

The study employed a sample of 690 individuals, of whom 196 were obese with a BMI above 30 ("patients"); and 494 were members of the general public ("population"). "Patients" were recruited from the Centre for Bariatric Surgery (CBS) in Melbourne. Because of limited capacity for the procedure, there are long queues for bariatric surgery in Australia. Patients over 18 years of age on the waiting list were approached by their clinicians and invited to participate in the research. Data were collected through self-completion questionnaires. The composite questionnaire included the 35 items of the AQoL-8D, the 36 items of the SF-36, and demographic questions including patients' height and weight and the existence of a comorbidity. Surveys were completed over the 21 month period to December 2009.

AQoL-8D data for the "public" were obtained during the interviews which were conducted for the construction of the instrument. These were selected to match the age-gender-education profile of the Australian population. Data were subsequently used to "standardize" results from patients by expressing their AQoL-8D scores as a percent of the population score. The SF-36 population norms for the Australian general public were obtained independently [28].

The instruments

The Short Form 36 (SF-36) was developed from the 245 items of the Rand Medical Outcomes Study [29]. From the original 40 physical and mental concepts, 8 were selected using psychometric procedures [30]. Reliability has been established on numerous occasions and in numerous countries [31,32]. Validity and sensitivity have likewise been tested in numerous contexts and countries (for a review see McDowell [15]). Use of the SF-36 is described in its manual and explained on the SF-36 website (http://www.sf-36.org).

The AQoL-8D is the fourth, and most comprehensive, of the AQoL instruments at the Centre for Health Economics, Monash University. It was developed to achieve increased sensitivity in the psycho-social dimensions which have been relatively neglected in previous MAU instruments, including earlier versions of the AQoL. Like the SF-36 its descriptive system was derived from a large number (250) of items and reduced to 8 dimensions using structural equation modeling (SEM) and data from a survey of 712 patients and members of the public [26]. A second survey of 620 patients and the public obtained utility (time tradeoff) scores for key parameters which permitted the creation of a utility scoring formula [33]. Internal and test-retest reliability coefficients are high [34]. Validation tests have been limited. Results are summarized in Richardson and Khan [27]. Use of the instrument is explained on the AQoL website [http://www.aqol.com.au/]. The algorithms for converting item responses into instrument scores are given in Hays et al. [35].

The dimensions and items of the two instruments are summarized in Table I. The eight

dimensions of SF-36 are grouped into two summary measures: the "physical component summary" (PCS) and "mental component summary" (MCS). The PCS consists of physical functioning (PF), role limit physical (RP), bodily pain (BP) and general health (GH). The MCS consists of vitality (VT), social functioning (SF), mental health (MH) and role limit emotional (RE). The AQoL-8D dimensions are also summarized in two measures: the "physical super dimension" (PSD) and the "mental super dimension" (MSD). The PSD includes the dimensions of independent living (IL), senses (Sen) and pain (Pain); and the MSD includes happiness (Hap), coping (Cop), relationships (Rel), self-worth (SW) and mental health (MH).

The SF-36 has a greater emphasis upon physical dimensions, with walking and the ability to work having 5 and 4 items respectively. Dexterity is included in the SF-36 but excluded from the AQoL-8D. However, the SF-36 includes no items relating directly to the physical senses or communication, whereas AQoL-8D has 3 and 1 items respectively relating to these elements. AQoL-8D has a relatively greater emphasis upon psychological and social dimensions with 5 and 4 items dedicated to the broad concept of depression/anxiety and social relationships respectively. AQoL-8D, but not SF-36, includes items for self-esteem and intimacy/sexual relationships. SF-36 has more items directly concerned with emotion related vitality. The differences between the instruments reflect the different psychometric analyses used in their development, referenced above.

Analysis

The comparison in Tabe I suggests a broad similarity between the content of the SF-36 and AQoL-8D but this is potentially misleading. The content of items depends upon their construction and specific wording and these differ. Omitted or superficially under-represented concepts may be detected by items from another dimension. The final scores for the two instruments also depend upon the combination formula, the way in which individual item responses are scored and combined.

Like other non-utility generic instruments the SF-36 combines unweighted item scores whose values depend only upon the response level. As noted, AQoL-8D combines items with an empirically derived formula, which gives values approximating utilities as derived by the time trade-off (TTO) technique [36]. While measurement units therefore differ, the two instrument scores would be expected to correlate if they were both sensitive to variation in the quality of life. This property is used in the following section to test the convergent validity of AQoL-8D – its positive correlation with independently validated measures. However, an overall correlation between the instruments is a weak test of content as it could occur when only a subset of the instrument dimensions correlated and the AQoL-8D was insensitive to other dimensions of the SF-36.

	Instrument dimensions and items	
	SF-36	**AQoL-8D**
Physical QoL	• Physical functioning (PF), 10 items: vigorous activities, moderate activities, lifting, climbing several flights of stairs, climbing one flight of stairs, bending, walking more than 1 km, walking ½ a km, walking 100 m, bathing • Role limit physical (RP), 4 Items: time spent on work, accomplished, limited to work, difficulty of performing work • Bodily pain (BP), 2 items: the degree of pain, interference with normal work due to pain • General health (GH), 6 items: general health, health rating, get sick a little easier, healthy as anybody, get worse, excellent health	• Independent living (IL), 4 items: household task, mobility outside the home, walking and self-care • Senses (Sen), 3 items: vision, hearing, communication • Pain (Pain), 3 items: experience of serious pain, the degree of pain, interference with usual activities caused by pain
Mental QoL	• Vitality (VT), 4 items: full of life, energy, fell worn out, feel tired • Social functioning (SF), 2 items: interference with normal activities, interference with social activities • Role limit emotional (RE), 3 items: time spent on work, accomplished less than you like, didn't work as carefully as usual • Mental health (MH), 5 items: nervous, felt down and nothing could cheer you up, felt calm, felt down, happiness	• Happiness (Hap), 4 items: contentment, enthusiasm, degree of feeling happiness, pleasure • Coping (Cop), 3 items: energy, being in control, coping with problems • Relationships (Rel), 7 items: relationship with family and friends, social isolation, social exclusion, intimate relationship, family and community role • Self-Worth (SW), 3 items: feeling like a burden, worthlessness, confidence • Mental health (MH), 8 items: feelings of depression, trouble sleeping, feelings of anger, self harm, feeling despair, worry, sadness, tranquility/agitation

Table I. *SF-36 and AQoL-8D Instrument dimensions and items*

A more exacting test of content validity was therefore carried out, namely a comparison of the overall AQoL-8D with each of the component dimension scores for the SF-36. A failure of the instrument to correlate with any of the component dimensions would suggest insensitivity of the AQoL-8D to that dimension. A test of concurrent validity – the ability of an instrument to discriminate between different groups – was also conducted, namely, a comparison of patients by their self-reported health. The test of concurrent validity was also applied to the comparator instrument, the SF-36.

Regression analyses were used to estimate the average incremental effect of BMI on QoL, i.e. the average change in the instrument scores with a change in BMI. The AQoL-8D score was used to estimate the effect of obesity on an individual's utility and the expected reduction in their QALYs.

Instrument	Mean (0-1) scale	Standard error	Range	Interquartile range (IQR)	Coefficient of relative variation (CRV)
SF-36					
Total	0.55	0.015	0.79	0.33	0.36
PCS*	0.40	0.007	0.54	0.14	0.22
MCS*	0.48	0.006	0.45	0.12	0.17
AQoL-8D					
Total	0.69	0.013	0.78	0.28	0.27
PSD*	0.63	0.011	0.65	0.23	0.25
MSD*	0.31	0.012	0.84	0.27	0.55

Table II. *Summary statistics for the two instruxments*. Population sample: 196 patients (male 23%, female 77%). Age group in years: (%) 18-24 (6.6); 25-34 (13.8); 35-44 (13.8); 45-54 (32.7); 55-64 (21.9); >65 (3.1). Socio Economic Status (SES): group 1: (lowest) 5.8%; group: 2 5.8%; group 3: 12.9%; group 4: 25.8%; group 5: (highest) 49.7%*

MCS = mental component summary; MSD = mental super dimension; PSD = physical super dimension

* Note that the summary scores cannot be compared with each other since they reflect the number of items and scale units which differ

RESULTS

A total of 196 obese patients completed the questionnaire. As it was administered by the hospital, it was perceived as a prerequisite for treatment and the response rate was 100%. Reflecting the fact that patients were eligible for bariatric surgery they were atypical of the general population. The mean age was 45.7. Three quarters were female (77%) and only 20 percent were aged less than 34. One half (49.7%) were from the highest socioeconomic group and only 5.8% from the lowest (see Table II). Figure 1 reports the patient frequency by BMI and age. The distribution is heavily skewed towards the right hand side: 61% were morbidly obese and 28% were moderately obese. Only 11 percent were slightly obese and none fell in the normal weight range. The mean BMI was 43.3 for men, 42.4 for women and 42.6 overall.

Prior to undertaking any statistical analysis, the internal consistency of both instruments was tested using Cronbach's Alpha [37]. The alpha coefficient of 0.870 for AQoL-8D and 0.874 for SF-36 respectively suggest good internal consistency for both instruments.

SF36 vs. AQoL-8D

SF-36 responses were transformed from a (0.00-100) to a (0.00-1.00) scale for comparison with the AQoL-8D. The resulting frequency distributions of the two instru-

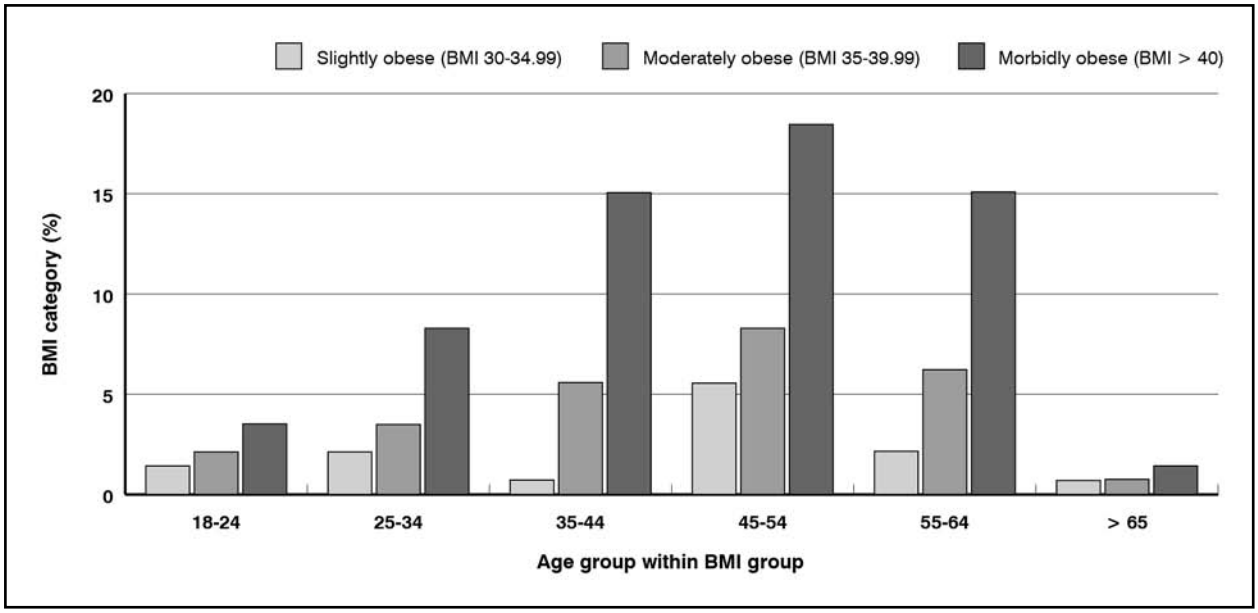

Figure 1. *Patient frequency by age and BMI group*

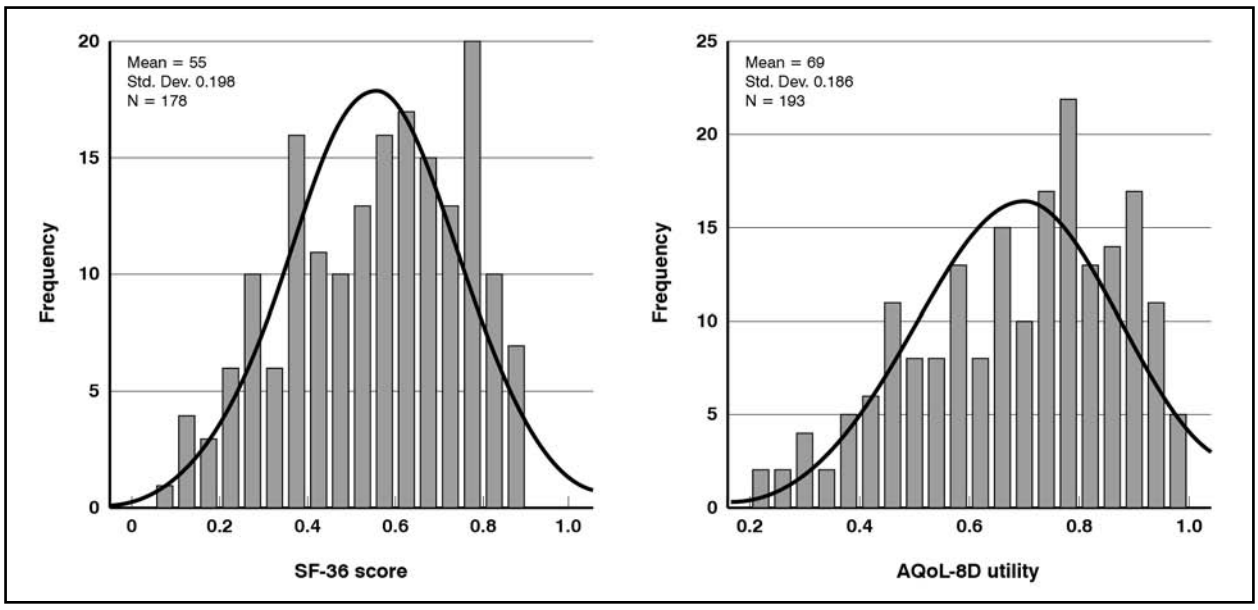

Figure 2. *Frequency distributions of patient QoL scores*

Attributes	SF-36									AQoL-8D								
	PCS			MCS			SF36			PSD			MSD			AQoL-8D		
	Mean	se	p	Mean	se	p	Mean	se	p	Mean	se	p	Mean	se	p	Mean	se	p
Gender																		
Male	0.43	0.013	0.01	0.49	0.01	0.126	0.62	0.03	0.02	0.67	0.024	0.02	0.35	0.026	0.096	0.74	0.03	0.049
Female	0.39	0.008		0.47	0.007		0.54	0.017		0.62	0.012		0.3	0.014		0.67	0.02	
Age																		
< 35	0.4	0.014	0.98	0.45	0.013	0.004	0.49	0.03	0.02	0.67	0.025	0.24	0.25	0.028	0.005	0.63	0.03	0.04
35-54	0.4	0.01		0.48	0.008		0.56	0.02		0.62	0.015		0.3	0.016		0.68	0.02	
> 55	0.4	0.017		0.51	0.012		0.61	0.03		0.62	0.022		0.37	0.024		0.74	0.03	

Table III. *Patient scores by age and gender*
MCS = mental component summary; MSD = mental super dimension; PCS = physical component summary; PSD = physical super dimension

ment scores for patients are shown in Figure 2. AQoL-8D exhibits a right hand skew reflecting the fact that it measures utilities (in the economist's sense). This is a reflection of people's reluctance to trade quality for quantity of life, as described earlier, in the vicinity of normal health.

AQoL-8D\ SF-36	AQoL-8D			SF-36	
	AQoL-8D	Physical	Mental	SF-36	Physical
AQoL-8D	1				
Physical (PSD)	0.728*	1			
Mental (MSD)	0.904*	0.506*	1		
SF-36	0.820*	0.634*	0.762*	1	
Physical (PCS)	0.533*	0.681*	0.406*	0.702*	1
Mental (MCS)	0.661*	0.284*	0.677*	0.665*	-0.021

Table IV. *Correlation between patient SF-36, AQoL-8D and component summary scores*
* Correlation is significant at the 1 % level (2-tailed)

Summary patient statistics are reported in Tables II and III. Measurement units and content of the SF-36 and AQoL-8D differ, which results in different numerical values for the sample mean (0.55 vs. 0.69) and different interquartile ranges, IQR (0.33 vs. 0.28). For comparative purposes the coefficients of relative variation (CRV) were calculated. These express the standard deviation as a percentage of the mean. From Table II the SF-36 displays more variation reflecting the skewed and truncated distribution of utility shown in Figure 2. AQoL-8D has greater variation in the physical and mental super dimension scores which do not measure utilities. From Table III both SF-36 and AQoL-8D identify significant differences in QoL by gender and age. Both physical summary measures (PCS, PSD) identify significant differences by gender but not by age. The two mental summary measures (MCS, MSD) reveal significant differences by age but not by gender. Neither instrument detected differences by the educational status of patients.

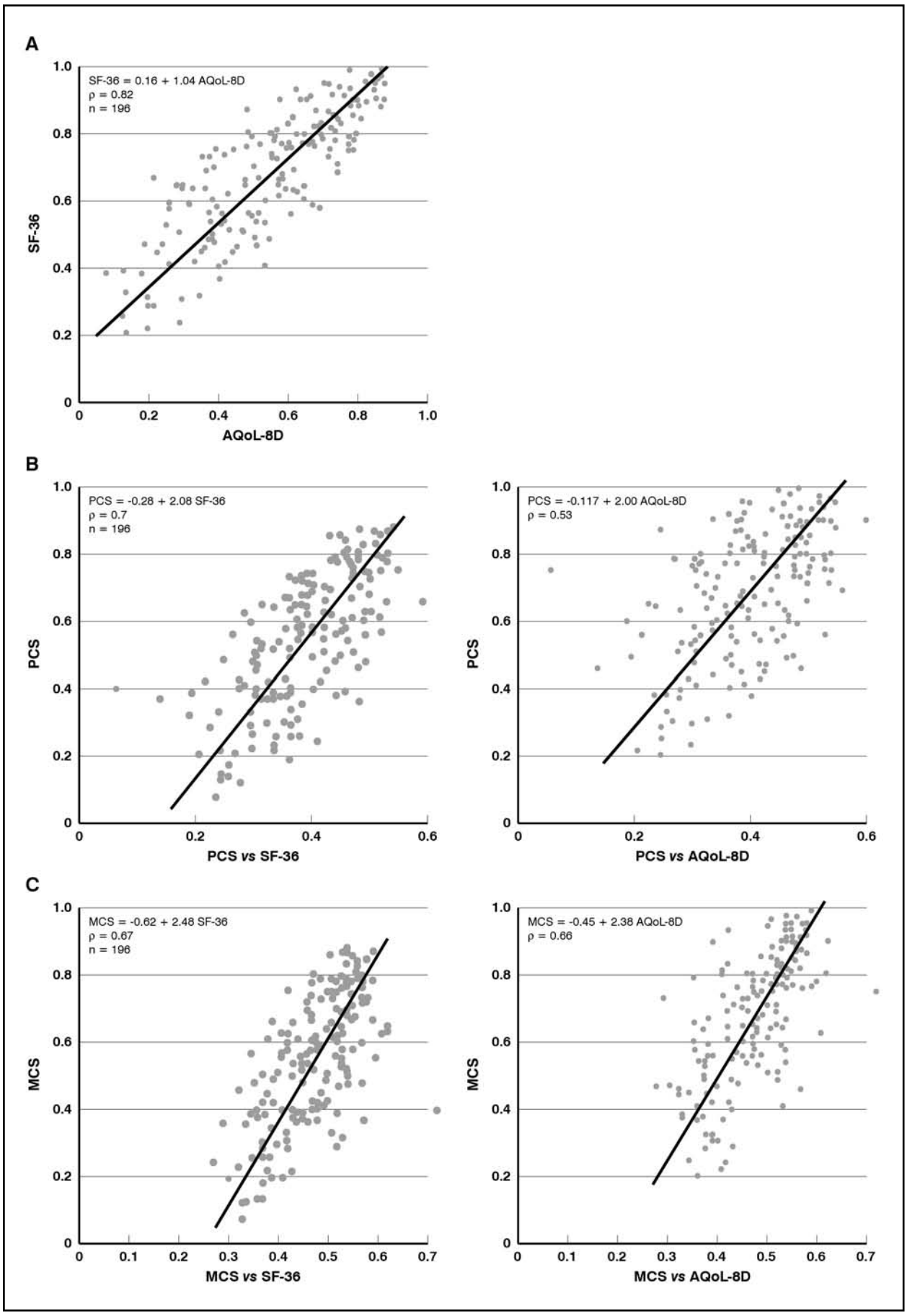

Figure 3. *(A). Geometric mean regression of SF-36 vs. AQoL-8D (Geometric Mean Regressions*): patient data. (B). Physical summary scores vs. SF-36 and AQoL-8D: Patient data. (C). Psycho-social component summary scores vs. SF-36 and AQoL-8D, patient data*

*Geometric mean regressions are derived from the geometric mean of parameters from the regression of each variable on the other. Results are independent of the choice of dependent and independent variable and are appropriate when both variables are subject to independent erro.

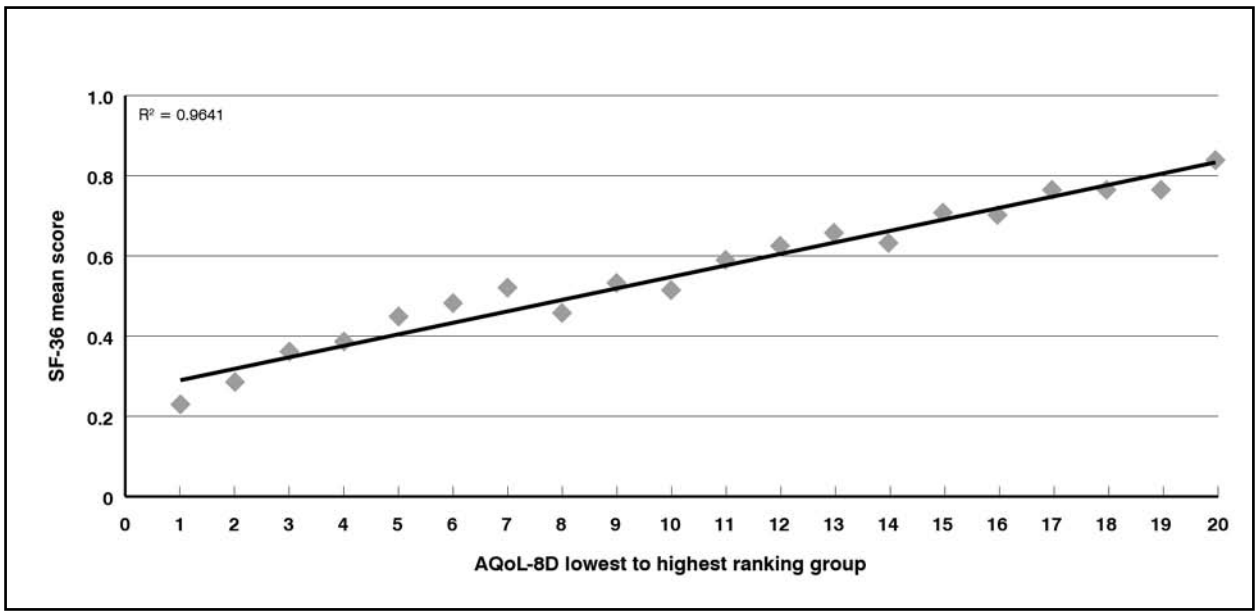

Figure 4. *Correlation between patient mean SF-36 scores and AQoL-8D utility scores*

Convergent validity

Table IV reports correlation coefficients between the instruments and their summary component scores. Figure 3 displays the linear relationship between the chief aggregates estimated with geometric mean regression. These permit error terms in both variables and provide results which are independent of the choice of dependent variable. The overall correlation of 0.82 between AQoL-8D and SF-36 indicates a very close relationship between them. By comparison, in the two multi instrument comparisons of utility instruments cited earlier, the highest correlation found by the Australian study between any two MAU instruments was 0.80 (15D and AQoL-4D) and in the US study, the highest correlations were 0.72 between SF-6D and HUI 3 and 0.82 between HUI 2 and HUI 3 [23,24]. Correlations between the two physical and two mental summary scores are lower (0.681 and 0.677 respectively) indicating that the instrument content differs in the two domains but the correlations still exceed the average correlation between utility instruments in the US study. The line of best fit between the SF-36 and AQoL-8D shown in Figure 3A has a slope coefficient of 1.04 which indicates that changes in the SF-36 provide a close approximation to changes in utility measured by the AQoL-8D.

The relationship between the overall physical and mental components of the two instruments shown in Figures 3B and 3C provide tests of convergent validity (the extent to which the component scales correlate in the expected way) and a test of the sensitivity of the overall instrument scores to the components. From Figure 3B the SF-36's physical component summary score, PCS, has a relatively low correlation with the overall AQoL-8D (0.53), but only a correlation of 0.7 with the overall SF-36 of which it is a part. That is, the correlation with AQoL-8D is 0.53/0.70 or 72% as large as the PCS correlation with the SF-36. From Figure 3C the SF-36 mental component summary score – MCS – has virtually the same correlation with the overall AQoL-8D (0.66) as it does with the overall SF-36 (0.67). The correlation between the AQoL-8D mental summary score – MSD – and the overall SF-36 (0.76) is higher than for any other pairwise comparison except for the MSD with AQoL-8D itself. Surprisingly, the SF-36 correlation of 0.76 with the MSD, the AQoL-8D summary score for men-

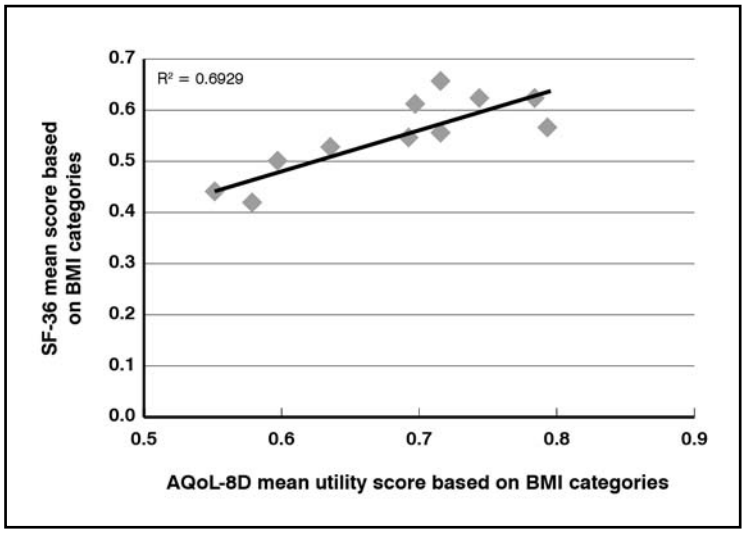

Figure 5. *Comparison of patient mean SF-36 and AQoL-8D scores with BMI*

tal health is greater than its correlation (0.66) with MCS one of its own components.

Evaluation studies typically use pooled or grouped results. To test the effect of this upon the overall correlation, data were pooled two ways. First, observations were ranked according to their AQoL-8D score and divided into 20 equal sized groups. The correlation between pooled SF-36 and AQoL-8D scores rises to 0.96 (Figure 4). Secondly, data were pooled by patient BMI. In this case the correlation was 0.69 (Figure 5).

	Physical dimensions	Mental dimensions	
Physical function	0.62*	Vitality	0.70*
Role limit	0.55*	Role limit (emotion)	0.63*
Bodily pain	0.58*	Mental health	0.77*
General health	0.60*	Social functioning	0.73*

Table V. *Content validity: correlation of patient AQoL-8D with SF-36 dimension scores*
* Correlation is significant at the 1 % level (2-tailed)l

Content validity

For the reasons outlined earlier, overall correlation is a relatively weak test of validity, and the more demanding test of content validity is the correlation between the AQoL-8D and the eight (independently validated) SF-36 dimension scores. Failure to correlate with any dimension would indicate insensitivity with respect to this dimension of health. Test results are reported in Table V. The coefficients vary from a low of 0.55 (role limit) to 0.77 (mental health). As these represent correlations between (non-utility) scores for particular dimensions of a health state and the total (utility) scores for complete health states, they are very high (see the average correlation of 0.66 and 0.73 between overall utilities in the two multi instrument comparative studies cited above).

Concurrent validity

The high correlations shown in Figures 3A to 3C imply that if the SF-36 has concurrent validity – the ability to discriminate between groups – then the AQoL-8D is likely to also display concurrent validity. A direct test of this is reported in Figure 6. Patients were divided into those reporting "very good" or "excellent", "good", "fair", "poor" or "very poor" health. Concurrent validity implies that the instruments should detect differences between those categories. Scores for both the SF-36, AQoL-8D and their physical and mental summary component scores are shown in Figure 6 as histograms with their 95% CI. The Figure displays the expected pattern. On this test the SF-36 has greater concurrent validity (F = 47.8, for the SF-36 vs. F = 34.8, for the AQoL-8D). This is attributable to greater discrimination by the PCS than the PSD. However the MSD has greater discriminatory power than MCS. Overall, however, both instruments perform well on this test.

Obesity and quality of life

Dimension effects

The effect of obesity upon HRQOL was analyzed by comparing the dimension and instrument scores of patients with the corresponding scores for the population obtained

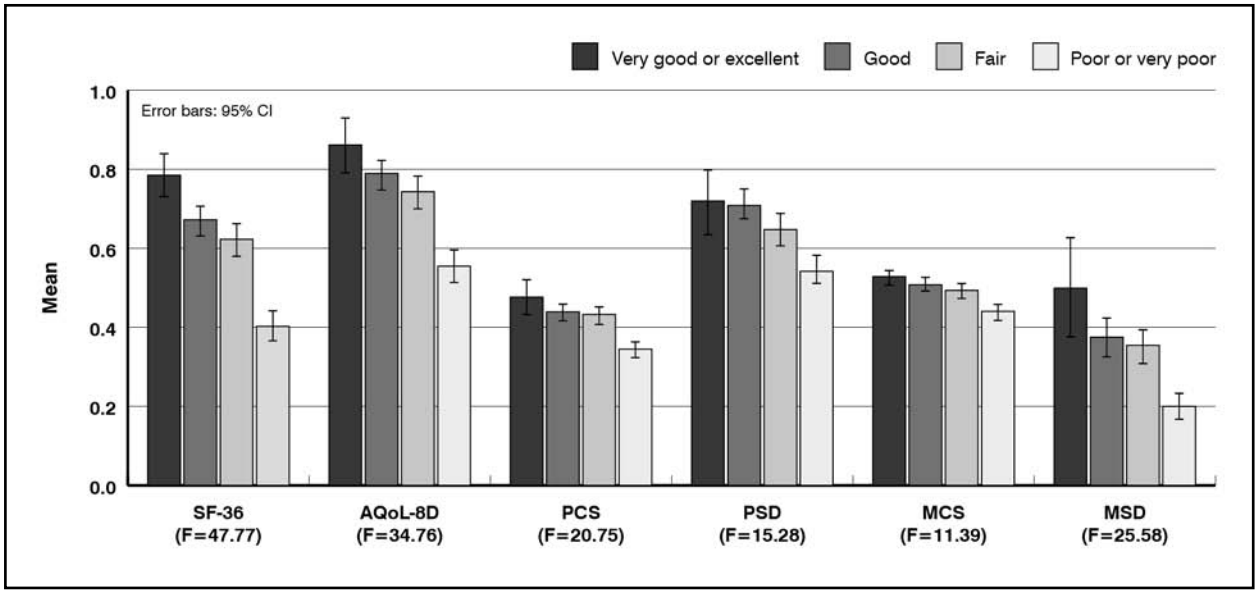

Figure 6. *Instrument scores by patient self-reported health*
MCS = mental component summary; MSD = mental super dimension; PCS = physical component summary; PSD = physical super dimension

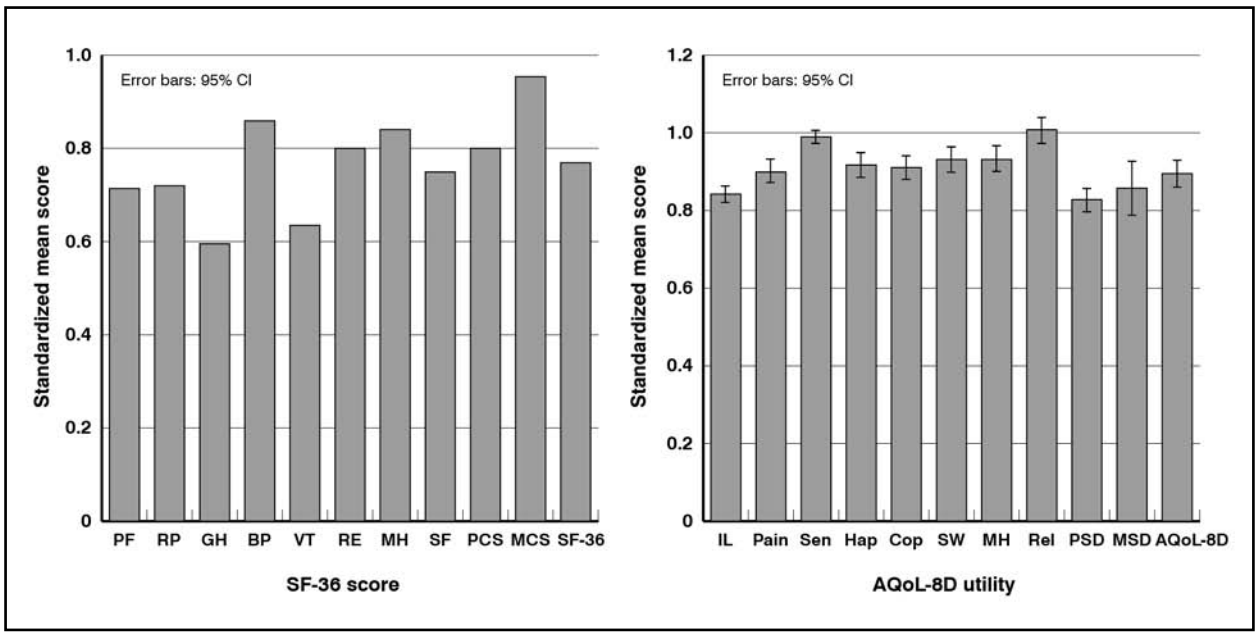

Figure 7. *Comparison of patient and population standardized scores by dimension. Instrument and dimensions scores are standardized by dividing patient scores by the corresponding average scores for the control (population) groups*
BP = bodily pain; Cop = coping; GH = general health; Hap = happiness; IL = independent living; MCS = mental component summary; MH = mental health; MSD = mental super dimension; PCS = physical component summary; PF = physical functioning; PSD = physical super dimension; Rel = relationships; RP = role limit physical; VT = vitality; RE = role limit emotional; Sen = senses; SF = social functioning; SW = self-worth

from the AQoL-8D and SF-36 norms data. Figure 7 presents the ratio of the scores for each of the 16 dimensions of the two instruments and for the physical and mental summary component scores. The ratio represents an index of the patient score where the population score is used as the norm. Consequently 1.00 would indicate a patient score equal to the population norm. A ratio less than 1.00 indicates a poorer outcome. With the exception of the AQoL-8D dimensions for senses and relationships, each of the patient dimension scores is significantly less than the population scores, i.e. the ratios are significantly less than unity. For the SF-36, the lowest relative dimension scores are for general health followed by vitality and physical functioning. For AQoL-8D the lowest relative scores are for independent living, pain and coping. Reflecting the emphasis in their construction, most of the decline in the SF-36 is attributable to the physical dimensions (PCS of 0.8 versus MCS of 0.95). In contrast, AQoL-8D physical and mental super-dimensions (PSD, MSD) scored 0.83 and 0.86 respectively.

BMI and QoL

The incremental change in instrument scores with an increase in BMI was analyzed using OLS linear regressions. These included BMI, patient age, gender and education as explanatory variables plus a single variable (illness) indicating whether a patient self-reported a "serious" illness. The analyses were carried out with the full dataset and separately for males and females. For comparative purposes both the SF-36 and AQoL-8D scores were used as dependent variables.

Both the SF-36 and AQoL-8D were regressed upon BMI and the other explanatory variables. Results are presented in Table VI. Regressions were very stable with little variation in the coefficients with the inclusion or exclusion of different variables. Patients with a self-reported serious illness had AQoL-8D scores which were, on average, 0.059 lower than for other patients. Females had lower scores. Patients below age 35 and between 35 and 54 had average AQoL-8D scores 0.116 and 0.074 lower than others. These unexpected results may reflect a greater effect of obesity upon the psychological wellbeing of younger people. Results for the educational dummy variables were insignificant. The coefficient on BMI was -0.006 in all iterations of the model; that is, an increase in BMI of 10 is predicted to decrease the AQoL-8D utility score by 0.06. The incremental effect shown in Table 5 implies that a change in BMI from 30 (slightly obese) to 50 (morbidly obese) decreases utility by 0.006 x 20 or by 0.12 which is 0.12/0.87 or 13.8 percent of the average utility score for a 25-34 year old. Since utility measures the preferred trade-off between quality and length of life this implies that an individual would accept a 13.8 percent reduction in their life expectancy to increase their QoL by the improvement predicted when BMI is reduced from 50 to 30.

Dependent Independent	AQoL-8D				SF-36			
	b	t	b	t	b	t	b	t
BMI	-0.006	-3.16***	-0.006	-3.39***	-0.005	-2.67	-0.005	-2.70***
Age								
• <35			-0.116	-2.66***			-0.044	-0.108^{ns}
• 35-54			-0.074	-2.14**			-0.058	1.51^{ns}
Education								
• Trade			0.053	1.43^{ns}			0.062	1.55^{ns}
• Degree			0.018	0.54^{ns}			0.041	1.11^{ns}
Illness			-0.059	-2.04**			-0.060	-1.88*
Female			-0.072	-2.12**			-0.085	-2.29**
Constant	0.941		1.083		0.79		0.85	
R^2	0.061		0.16		0.047		0.14	
F	10.00		3.94		7.12		3.22	

Table VI. *Regressions: all respondents, patient AQoL-8D, SF-36*
*** Significant at 1 percent level
** Significant at 5 percent level
* significant at 10 percent level

Results for the SF-36 parallel the AQoL-8D results to a surprising extent (considering that AQoL-8D but not SF-36 scores include the utility weights). Consistent with the linear relationship between AQoL-8D and SF-36 displayed in Figure 4, the coefficient on BMI is almost identical with the coefficient for AQoL-8D. Signs on other variables are uniformly consistent but, with the exception of the dummy variable for females, insignificant. The overall explanatory power of the SF-36 model is lower.

DISCUSSION

The dual objectives of this paper were, firstly, to compare results from AQoL-8D and SF-36 to determine whether the utilities derived from the former instrument were a valid representation of the health related quality of life as measured by the latter instrument and, secondly, to use results from both instruments to describe and quantify the effects of obesity on QoL. The first set of results is therefore a test of the validity of the AQoL-8D. The latter results describe the impact of obesity upon the dimensions of health related wellbeing and, using the AQoL-8D, the effect upon people's health related utility.

Taking account of the differences in questions and differences in scoring, the similarity in the results from the two instruments is striking. The frequency distributions of the overall scores differ reflecting the measurement of utility by AQoL-8D and unweighted scores by the SF-36. Nevertheless the linear relationship between them (Figure 4) reveals an incremental change in AQoL-8D utilities which is virtually identical to the change in SF-36 scores and the correlation between them is very high. The important caveat is that these results were derived from patients waiting for bariatric surgery and could not be reliably extrapolated to other population groups.

Tests conducted in the first half of the paper indicate that the AQoL-8D has good convergent and concurrent validity and, using the SF-36 as the criterion, good content validity. SF-36 scores varied more with self-reported health. However AQoL-8D was more closely related to BMI in the regression analyses. The AQoL-8D mental health super-dimension was more sensitive than its counterpart in the SF-36.

In the second half of the results, the instruments identified a significant deterioration in the quality of life in every dimension of both instruments with only two exceptions. AQoL-8D results identified a significant decrease in the psychological and social dimensions summarized by MSD. Happiness, self-worth, coping and mental health were all poorer and only social relationships were not affected.

BMI	Utility decrement from average utility for BMI = 20-25					
	HUI 3	EQ-5D	SF-6D	QWB	SF-6D	AQoL-8D*
25-29.9	-0.00	-0.013	-0.01	-0.014	-0.016	-0.0345
30.9-34.9	-0.02	-0.033	-0.06	-0.044	-0.030	-0.0645
> 35	-0.04	-0.073	-0.11		-0.052	-0.101

Table VII. *Loss of utility with BMI, various studies. Sources: HUI 3 [38]; EQ-5D [39]; SF-6D [40]; QWB [41]; SF-6D [8]*
* Present results

Regression analyses revealed very similar deterioration in the scores of both instruments with BMI. The results from the AQoL-8D regressions permit comparison with other studies which have employed MAU instruments. These are summarized in Table VII. The striking feature in this comparison is that the effect of BMI identified by AQoL-8D is approximately double the effect found in other studies. This is possibly attributable to the greater average BMI of patients in the present analyses, but may also be attributable to the wider scope of the AQoL-8D descriptive system and, in particular, its sensitivity to psycho-social dimensions. These are of lesser importance in the other instruments in the comparison.

Comorbidities

An important caveat in the interpretation of these (and other) results is that the causal path in the association between obesity and quality of life cannot be demonstrated by the statistical association. In Kortt and Clarke [8] variables were included for the major diseases associated with overweight and obesity. In the present study a cruder, single self-reported variable was included for the existence of "serious" illness. The inclusion of these variables does not preclude the possibility that the loss of utility was causally related to other illnesses and not to BMI *per se*. However, if the causal path was from BMI to any disease to the loss of QoL then conclusions for the importance of BMI with respect to QoL need little qualification as the pathway does not alter the magnitude of the loss. In principle, it is possible that illnesses contracted independently of BMI result in obesity and that the former, not the latter, is the reason for decreasing utility. While this hypothesis cannot be disproven, in the absence of supporting evidence, it is not compelling.

CONCLUSION

The two instruments used in this study had 16 dimensions describing the broad areas of health related mental and physical health. The "patients" included in the study revealed significantly lower scores on 14 of the 16 dimensions indicating that the effects of obesity are not confined to direct physical effects but include social and psychological problems.

An unusual feature of this study was the inclusion of an MAU instrument, the AQoL-8D, which included these psycho-social factors and permitted the quantification of the effects of obesity upon "utility" as measured in economic evaluation studies. To obtain credible results the new instrument required validation and this was achieved by comparison with the SF-36. Results indicated a closer relationship between this non-utility instrument and the AQoL-8D utility instrument than has been observed in the majority of studies comparing utility instruments. This justifies confidence in the results obtained by the AQoL-8D for economic evaluation.

Comparison of AQoL-8D scores with BMI suggest that, at least in the range of observations used in this study, there is a strong association between the two variables. The incremental effect of BMI is at least twice as great as obtained in previous studies, possibly as a result of the inclusion of a broader range of psycho-social variables in the instrument or high average BMI. If the relationship is causal, which is probable, then the loss of utility attributable to obesity implies that the health related burden of disease caused by the ongoing obesity epidemic is greater than previously believed which, in turn, increases the urgency for measures to combat it.

REFERENCES

1. Access Economics. The Growing Cost of Obesity in 2008: Three Years On, Report to Diabetes Australia, 2008. Available at: http://www.diabetesaustralia.com.au [last accessed May 2012]
2. Centers for Disease Control and Prevention. National Obesity Trends, 2011. Available at: http://www.cdc.gov/nchs/fastats/overwt.htm [last accessed May 2012]
3. Sassi F, M. Devaux, M. Cecchini, et al. The Obesity Epidemic: Analysis of Past and Projected Future Trends in Selected OECD Countries. Paris: OECD Health Working Paper, No. 45, 2009
4. Dennett SL, Boye K, Yurgin NR. The impact of body weight on patient utilities with or without Type 2 Diabetes: a review of the medical literature. *Value in Health* 2008; 11: 478-86; http://dx.doi.org/10.1111/j.1524-4733.2007.00260.x
5. National Task Force on the Prevention and Treatment of Obesity. Overweight, obesity and health risk. *Arch Intern Med* 2000; 160: 898-904
6. National Institute of Health. Clinical guidelines on the identification, evaluation and treatment of overweight and obesity in adults – the evidence report. *Obesity Research* 1998; Suppl 2: 51S-209S

7. Troiano R, Frongillo EJ, Sobal J, et al. The relationship between body weight and mortality: a quantitative analysis of combined information from existing studies. *Int J Obes Relat Metab Disord* 1996; 20: 63-75

8. Kortt MA, Clarke PM. Estimating utility values for health states of overweight and obese individuals using the SF36. *Quality of Life Research* 2005; 14: 2177-85

9. Feldstein EA, Trogdon JG, Cohen JW, et al. Annual medical spending attributable to obesity: payer-and service-specific estimates. *Health Affairs* 2009; 28: 822-31

10. Behan DF, Cox SH. Obesity and its relation to mortality and morbidity costs, December 2010, Committee on Life Insurance Research. Society of Actuaries, 2010. Available at: http://www.soa.org/research/research-projects/life-insurance/research-obesity-relation-mortality.aspx [last accessed May 2012]

11. Muller-Rienenschneider F, Reinhold T, Berghofer A, et al. Health-economic burden of obesity in Europe. *Eur J Epidemiol* 2008; 23: 499-509; http://dx.doi.org/10.1007/s10654-008-9239-1

12. Butland B, Jebb S, Kopelman P, et al. Foresight: Tackling Obesities: Future Choices - Project Report, 2nd edition. London, 2007. Available at: http://www.bis.gov.uk/foresight/our-work/projects/published-projects/tackling-obesities [last accessed May 2012]

13. Colagirui S, Lee CMY, Colagiuri Ret al. The cost of overweight and obesity in Australia. *Medical Journal Australia* 2010; 192: 260-4

14. Bowling A. Measuring health: a review of quality of life measurement scales. Maidenhead, Berkshire: Open University Press, 2005

15. McDowell I. Measuring health: a guide to rating scales and questionnaires. Oxford: Oxford University Press, 2006

16. Kolotkin R, Head S, Hamilton M, et al. Assessing impact of weight on quality of life. *Obesity Research* 1995; 3: 49-56

17. Duval K, Marceau P, Perusse L, et al. An overview of obesity-specific quality of life questionnaires. *Obesity Reviews* 2006; 7: 347-60; http://dx.doi.org/10.1111/j.1467-789X.2006.00244.x

18. McSweeny A, Creer T. Health-related quality of life assesment in medical care. *Dis Mon* 1995; 41: 1-71

19. Samsa G, Kolotkin R, Williams G, et al. Effect of moderate weight loss on health-related quality of life: an analysis of combined data from 4 randomized trials of sibutramine vs placebo. *Am J Manag Care* 2001; 7: 875-83

20. WHOQOL Group. Teh development of the World Health Organization quality of life assessment instrument (the WHOQOL). In: Quality of Life Assessment: International Perspectives. Edited by: Orley J, Kuyken W. Berlin: Springer-Verlag, 1994; pp. 41-57

21. Gilson B, Gilson J, Bergner M, et al. The Sickness Impact Profile: development of an outcome measure of health care. *Am J Public Health* 1975; 65: 1304-10

22. Hunt S, McEwen J, McKenn S. Measuring health status: a new tool for clinicians and epdemiologists. *J R Coll Gen Pract* 1985; 35: 185-8

23. Fryback DG, Palta M, Cherepanov D, et al. Comparison of 5 health related quality of life indexes using item response theory analysis. *Medical Decision Making* 2010; 30: 5-15; http://dx.doi.org/10.1177/0272989X09347016

24. 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

25. Tayyem R, Abdulmajid A, Atkinson J, et al. Analysis of health-related quality of life instruments measuring the impact of bariatric surgery. *Patient* 2011; 4: 73-87; http://dx.doi.org/10.2165/11584660-000000000-00000

26. Richardson J, Elsworth G, Iezzi A, et al. Increasing the Sensitivity of the AQoL Inventory for Evaluation of Interventions Affecting Mental Health, Research Paper 61. Melbourne: Centre for Health Economics, Monash University, 2011

27. Richardson J, Khan MA. Preliminary results for the Validation of the Assessment of Quality of Life AQoL-8D Instrument, Research Paper 47. Melbourne: Centre for Health Economics, Monash University, 2009

28. ABS. Austalian Bureau of Statistics, National Health Survey SF-36 Population Norms Australia ABS Catalogue No. 4399. Canberra, 1995. Available at: http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4399.01995?OpenDocument [last accessed May 2012]

29. Lohr K, Brook R, Kamberg C, et al. Use of medical care in the RAND Health Insurance Experiment. Diagnosis and service-specific anlayses in a randomized controlled trial. *Medical Care* 1986; 24: S1-S87

30. Ware J, Sherbourne D. The MOS 36-Item Short-Form Health Survey (SF-36). I Conceptual framework and item selection. *Medical Care* 1992; 30: 473-483

31. McHorney C, Ware J, Lu J-FR, et al. The MOS 36 item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Medical Care* 1994; 32: 40-66

32. Gandek B, Ware J, Aaronson N, et al. Tests of data quality, scaling assumptions, and reliability of the SF-36 in eleven countries: results from the IQLA project. *J Clin Epidemiol* 1998; 51: 1149-58

33. Richardson J, Sinha K, Iezzi A, et al. Modelling the Utility of Health States with the Assessment of Quality of Life (AQoL) 8D Instrument: Overview and Utility Scoring Algorithm, Research Paper 63. Melbourne: Centre for Health Economics, Monash University, 2011

34. Richardson J, Chen G, Iezzi A, et al. Transformations between the Assessment of Quality of Life (AQoL) Instruments, Research Paper 66. Melbourne: Centre for Health Economics, Monash University, 2011

35. Hays R, Sherbourne D, Mazel R. The RAND 36-item health survey 1.0. *Health Economics* 1993; 2: 217-227

36. Richardson J, Iezzi A. Psychometric validity and the AQoL-8D Multi Attribute Utility Instrument, Research Paper 71. Melbourne: Centre for Health Economics, Monash University; 2011

37. Cronbach L. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951; 16: 297-334

38. Trakas K, Oh P, Leiter L, et al. The health status of obese individuals in Canada. *Int J Obes Relat Metal Disord* 2001; 25: 662-8

39. Jia H, Lubetkin EI. The impact of obesity on health-related quality of life in the general adult US population. *J Public Health (Oxf)* 2005; 27: 156-64

40. Sach TH, Barton GR, Doherty M. The relationship between body mass index and health-related quality of life: comparing the EQ-5D, EuroQol VAS, and SF-6D. *Int J Obes* 2007; 31: 186-96; http://dx.doi.org/10.1038/sj.ijo.080336

41. Groessl RJ, Kaplan R, Barrett-Connon E, et al. Body Mass Index and quality of well-being in a community of older adults. *Am J Prev Med* 2004; 26: 126-9

