

An Exploratory Study: A Head-to-Head Comparison of the EQ-5D-5L and AQoL-8D for Long-Term Publicly Waitlisted Bariatric Surgery Patients Before and 3 Months After Bariatric Surgery

Julie A. Campbell¹ · Martin Hensher² · Amanda Neil¹ · Alison Venn¹ · Petr Otahal¹ · Stephen Wilkinson³ · Andrew J. Palmer¹

Published online: 13 November 2017

© The Author(s) 2017. This article is an open access publication

Abstract

Background Choice of a multi-attribute utility instrument (MAUI) that appropriately assesses an intervention's health-related quality-of-life (HRQoL) impacts is a vital part of healthcare resource allocation and clinical assessment.

Objective Our exploratory study compared the EuroQol (EQ)-5D-5L and Assessment of Quality of Life (AQoL)-8D MAUIs, which were used to assess the effect of bariatric surgery for a cohort of long-term publicly waitlisted, severely obese patients.

Methods The study was conducted at the Hobart Private Hospital (Tasmania, Australia). To compare the sensitivity and instrument content of the two MAUIs, we used dimensional comparisons by investigating the distribution of patient-reported responses (number/percentage) across the MAUIs' levels and dimensions; summary health-state

utility valuations (utilities); and individual/super-dimension scores (AQoL-8D) to investigate discriminatory power and HRQoL improvements preoperatively and 3 months postoperatively.

Results Participants' ($n = 23$) overall MAUI completion rate was 74%. Postoperative total weight loss was 9.9%. EQ-5D-5L utilities were relatively higher pre- and postoperatively than AQoL-8D utilities [mean standard deviation (SD) EQ-5D-5L 0.70 (0.25) to 0.80 (0.25); AQoL-8D 0.51 (0.24) to 0.61 (0.24)]. AQoL-8D Psychosocial super dimension was relatively low postoperatively [0.37 (0.25)], driving the instrument's lower utility. These results were supported by the dimensional comparisons that revealed an overall greater dispersion for the AQoL-8D. Nevertheless, there were clinical improvements in utilities for both instruments. AQoL-8D utilities were lower than population norms; not so the EQ-5D-5L utilities. The AQoL-8D dimensions of Happiness, Coping, and Self-worth improved the most.

Conclusions AQoL-8D more fully captured the impact of obesity and bariatric surgery on HRQoL (particularly psychosocial impacts) for long-term waitlisted bariatric surgery patients, even 3 months postoperatively. AQoL-8D preoperative utility revealed our population's HRQoL was lower than people with cancer or heart disease.

Electronic supplementary material The online version of this article (doi:10.1007/s41669-017-0060-1) contains supplementary material, which is available to authorized users.

✉ Andrew J. Palmer
Andrew.Palmer@utas.edu.au

¹ Menzies Institute for Medical Research, University of Tasmania, Medical Sciences 2 Building, 17 Liverpool Street, Hobart, TAS 7000, Australia

² Department of Health and Human Services, Level 2, 22 Elizabeth Street, Hobart, TAS 7000, Australia

³ Royal Hobart Hospital, 48 Liverpool Street, Hobart, TAS 7000, Australia

Key Points for Decision Makers

The Assessment of Quality of Life (AQoL)-8D may have superior discriminatory sensitivity compared to the EuroQol (EQ)-5D-5L for long-term waitlisted severely obese bariatric surgery patients.

There is potential for sub-optimal healthcare resource allocation if the selected multi-attribute utility instrument does not appropriately assess health-related quality-of-life (HRQoL) impacts for the bariatric surgery study population.

As an important and increasingly prevalent study population of bariatric surgery patients who inherently carry complex physical and psychosocial HRQoL needs, long-term waitlisted severely obese bariatric surgery patients showed improvements in HRQoL even 3 months postoperatively.

1 Introduction

Demand for publicly funded bariatric care in many countries is high; however, capacity is limited by healthcare funding decisions. Consequently, bariatric (metabolic, obesity or weight-loss) surgery waiting lists are long [1, 2]. Prolonged delays generally exist for people waitlisted for primary bariatric surgery in public health systems in many countries, including Australia [3–5].

Whilst it is acknowledged that these protracted multi-year wait times are detrimental to the bariatric surgery candidate's physical and psychosocial health [2, 6, 7], recent evidence has established that weight status is just one factor contributing to the complex health-related quality-of-life (HRQoL) needs of people who have received bariatric surgery [8, 9]. Nevertheless, there is a paucity of quantitative evidence regarding HRQoL impacts for long-term waitlisted bariatric surgery patients who have experienced multiyear wait times on public waiting lists and then undergo bariatric surgery [10, 11].

Multi-attribute utility instruments (MAUIs) are a HRQoL assessment tool designed to rapidly and conveniently assess and capture an individual's health-state utility values through application of pre-established formulae/weights to the array of responses obtained on the MAUI's questionnaire [9]. A MAUI is developed and defined with particular characteristics, including the number of questionnaire items; the depth and breadth of the descriptive/classification system; the number of health

states described; the number of individual and super dimensions (if there are super dimensions); and the algorithmic range.

For example, the number of health states described for the EuroQol (EQ)-5D-3L and 5L, Health Utilities Index (HUI) 3, 15D, Short-Form 6 Dimension (SF-6D), Quality of Well-Being (QWB) and Assessment of Quality of Life (AQoL)-8D MAUIs range from 243; 3125; 972,000; 3.1×10^{10} ; 18,000; 945; and 2.4×10^{23} , respectively [12]. Additionally, many MAUIs target physical health within their descriptive/classification systems. For example, for the EQ-5D-5L, one of its five dimensions relates to psychosocial health (Anxiety/Depression) and four out of five relate to physical health (Mobility, Self-care, Usual Activities and Pain) [13]. In contrast, for the AQoL-8D, three of the instrument's eight dimensions relate to physical health (Independent Living, Senses and Pain), and five of the eight dimensions relate to psychosocial health (Coping, Relationships, Self-worth, Happiness and Mental Health), and 25 of the 35 items (questions) inform the AQoL-8D's five psychosocial dimensions [14, 15]. The SF-6D describes six dimensions, namely Physical Functioning, Role Limitations, Social Functioning, Pain, Mental Health and Vitality [12, 16]. Both the AQoL-8D and SF-6D describe composite physical and psychosocial dimensions, namely the Physical and Psychosocial super dimensions (AQoL-8D), and the Physical and Mental Component Summaries (SF-6D) [14, 17].

A small number of MAUIs dominate the economic evaluation literature. These include the EQ-5D-3L (precursor to the EQ-5D-5L), HUI 3 and SF-6D. A review of 1663 studies between 2005 and 2010 found that these three instruments accounted for 63, 9.9, and 8.8% of the total, respectively [12]. Four other instruments in the review, the 15D, HUI 2, AQoL, and QWB, were used in 7, 4.6, 4.2, and 2.5% of the studies, respectively [18].

A recent cross-sectional study of patients who had received bariatric surgery in the private healthcare system many years previously [median [interquartile range (IQR)] 5 (3–8) years] found that the AQoL-8D and EQ-5D-5L instruments were not interchangeable for the study population [9]. Another recent study that investigated the 1-year health impacts for long-term waitlisted bariatric surgery patients (and complementary to this study using the same cohort of patients), suggested that the AQoL-8D preferentially captured HRQoL for the study population 1 year after surgery [11]. Importantly, this 1-year study did not directly compare the distributions of patient-reported responses across the depth and breadth of the MAUIs' dimensions of health (dimensional comparisons) [11]. As a single MAUI instrument, the AQoL-8D captures the vast majority of domains considered crucial for people who are considering, or who have undergone, bariatric surgery [9].

The choice of MAUI should be influenced by the sensitivity of the instrument to a patient group's health profile [9, 12]. If the choice of instrument does not appropriately capture and assess the individual's and study population's health profiles (particularly for complex physical and psychosocial HRQoL), vital healthcare information about a clinical intervention's health impact will be omitted from important resource allocation and planning decisions [9].

Utility valuations are key health economic metrics that are an input measure in the assessment of quality-adjusted life years (QALYs) [19]. Utility valuations measure the strength of preference for a particular health state and are represented as a number on a scale where 1.0 represents the best possible health state and 0.0 represents death. In principle, values less than zero are possible when a health state is worse than death [20]. Utility values assessed by MAUIs are not equivalent, with the difference between the descriptive/classification systems of the MAUIs the principal determinant [12]. Additionally, differences in descriptive/classification systems are estimated to explain an average of 66% of the difference between utilities obtained by MAUIs, and 81% of the difference between the utilities of the EQ-5D-5L and AQoL-8D [12].

MAUIs were not initially developed for clinical use; however, utility valuations can also be used to inform and/or predict clinical outcomes [21]. Clinicians have found that measuring utilities is of benefit to patient-clinical assessment, relationships, communication, and management [22]. Many MAUIs (including the EQ-5D-5L and 3L, AQoL-4D, SF-6D, 15D and HUI) report minimal clinically important differences or minimal important differences for their utility valuations [23–28]. A minimal clinically important difference is the smallest difference in score in the outcome of interest that patients perceive as beneficial and which would mandate a clinical change in the patient's management (both individually and collectively for a particular study population) [22, 23, 29, 30].

There is a paucity of evidence regarding short-term HRQoL impacts for people who have received bariatric surgery [31, 32]. A study published in 2007 provided 3-month (range 3–6 months) HRQoL impacts of bariatric surgery using the SF-36 [33]. A second study published in 2001 provided 1-, 3- and 6-month HRQoL impacts of bariatric surgery using the SF-36, bariatric analysis and reporting outcome system (BAROS) and Moorhead-Ardelt quality-of-life questionnaires [34]. Both studies found short-term improvements in the quality of life scores (however, these studies did not generate, nor investigate, utility valuations) after bariatric surgery.

Whilst it is acknowledged that integrating patient-reported outcomes (PROs) in clinical practice has the potential to enhance patient-centred care [35], PROs are not yet routinely collected in bariatric care. A recent

systematic review that identified and investigated prospective bariatric surgery studies that used validated PRO measures found that for PRO data to influence practice, well-designed and reported studies are required [36]. In turn, there is a potential for MAUIs to address this key gap regarding PROs in bariatric care subject to the particular MAUI's capacity to capture, assess and describe the relevant health states of the study population.

The main objective of this exploratory study was to directly compare the discriminatory power of two different MAUIs, namely the EQ-5D-5L and the AQoL-8D, which were used to assess the effect of bariatric surgery using a cohort of long-term publicly waitlisted, severely obese patients who underwent bariatric surgery as part of a government policy initiative to reduce waiting lists. As a secondary objective, we also aimed to investigate the role of the two MAUIs in the analysis of individual patient health states.

The EQ-5D suite of instruments dominates the clinical and economic literature, including that for bariatric surgery [14, 18]. Nevertheless, the AQoL-8D has been shown to have preferential psychometric properties compared to comparative MAUIs in study populations where the assessment of psychosocial health status is crucial, for example, intensive care unit (ICU) admissions (compared with SF-6D) [22] and people who had undergone bariatric surgery (compared with the EQ-5D-5L) [9, 11]. Additionally, a recent study that presented results from one of the broadest comparative surveys in terms of the range of diseases (arthritis, asthma, cancer, depression, diabetes, hearing loss and heart disease) and six MAUIs (EQ-5D-5L, SF-6D, HUI 3, 15D, QWB and AQoL-8D), and countries (Australia, the USA, the UK, Canada, Norway, and Germany) found that the AQoL-8D is the most sensitive instrument for measuring mental health [37]. This study also found that the pain component of the EQ-5D-5L has a greater impact than it does in any other instrument, and that the EQ-5D-5L is the most sensitive instrument for measuring pain [37].

Our exploratory study also investigated the relative magnitudes of the global utility valuations [12], clinical improvements of the utility valuations for both instruments, and also the impacts on individual domains of health through the AQoL-8D's individual and super-dimension scores.

In parallel with our previously published study that investigated the 1-year health impacts in long-term waitlisted patients [11], this current study aimed to investigate the distribution of the patient-reported responses of the two MAUIs for this population of public healthcare long-waiting bariatric surgery patients who inherently carry complex physical and psychosocial HRQoL needs.

2 Methods

2.1 Study Design

2.1.1 Recruitment of Participants

Recruitment of our study participants is described in detail in our previously published study [11]. In summary, a Tasmanian government policy decision was made in 2014 to allocate additional and targeted public funds to provide morbidly obese, long-term waitlisted patients with bariatric surgery in 2015 [38]. All participants underwent laparoscopic adjustable gastric band (LAGB) surgery by the same surgeon in the Hobart Private Hospital. Laparoscopic banding was carried out using Apollo APS or APL bands, with adjustment ports attached to the left anterior rectus sheath [39]. Postoperative fluid diets were maintained for 3 weeks, with subsequent transition to normal foods, accompanied by instruction on eating technique and exercise.

All data were de-identified. Ethics approval was granted by the University of Tasmania's Health and Medical Human Research Ethics Committee (HMHREC) before our study's recruitment of participants.

2.1.2 The Multi-attribute Utility Instruments and Questionnaire Completion

The selection and attributes of the EQ-D-5L and AQoL-8D MAUIs used in this study have previously been described in detail [11]. Another earlier study comparing the EQ-5D-5L and the AQoL-8D MAUIs for people who had undergone LAGB surgery many years previously provided a detailed summary of the divergent characteristics of the two purposively selected MAUIs [9, 11]. In summary, the two markedly different MAUIs were selected on the following basis: the EQ-5D-5L is the internationally prevalent instrument in economic evaluation (including the economic evaluation of bariatric surgery) [40]; four of the five instrument's health domains/classifications (and items) focus on physical HRQoL; and it takes less than 1 min to complete the EQ-5D-5L's questionnaire [13]. The EQ-5D-5L also contains a visual analogue scale (EQ-VAS) [22]. In contrast, the AQoL-8D's classification system is supported by psychometric principles and testing, and 25 of the instrument's 35 items capture and assess five (from eight) psychosocial domains of health, and three physical domains of health. The AQoL-8D describes billions of health states and takes 5 min to complete [14, 15, 41].

Participants were asked to self-complete both instruments' questionnaires before their bariatric surgery at the pre-admission preoperative clinics and at 3 months

postoperatively. Postoperative questionnaires were mailed out for self-completion with an explanatory cover letter and reply-paid envelope enclosed. We evaluated EQ-5D-5L and AQoL-8D questionnaire completion by assessing the overall proportion of participants who completed the questionnaire(s) at the study's two time points for whom an individual utility value could be generated.

2.2 Data Analysis

Participants with patient-reported HRQoL assessments for one or both instruments, for at least one time point where the MAUI algorithm (either instrument) could generate the instrument's utility valuations or scores were included in the analyses.

Descriptive baseline socio-demographic, clinical data, utility valuations and dimensional scores were presented as mean [standard deviation (SD)] and/or median (IQR) for continuous variables and frequency (%) for categorical variables. Body mass index (BMI) was calculated as weight (kg)/[height (m²)] and classified as obese (BMI 30–34.9 kg/m²), severely obese (BMI 35–39.9 kg/m²), morbidly obese (BMI 40–49.9 kg/m²), and super obese (BMI ≥50 kg/m²) [42].

2.2.1 Discriminant Sensitivity: Dimensional Comparisons (Both Instruments) and Dimensional Scores (AQoL-8D)

The relative discriminatory power of the instruments was investigated using two methodologies.

First, we calculated the distribution of participant responses across the levels and dimensions (the depth and breadth) of both instruments. This was achieved by collating the participant-reported response for each item and then calculating the percentage distribution of responses for each dimension [9, 16]. To illustrate, for the EQ-5D-5L individual dimension of Anxiety/Depression, the numbers of participants who gave each response level (1, 2, 3, 4 or 5) were converted to a percentage of the total number of participants in order to derive a 'five-level frequency distribution'. Detailed calculations for each item and dimension are provided in Appendix 1 [see the electronic supplementary material (ESM)]. Additionally, schematic representations of the dimensional comparisons were expressed as a percentage by calculating the average percentage before and after surgery. For example, the schematic representation of the physical dimensions of both instruments compared the average score of Mobility, Self-care, Usual Activities and Pain for the EQ-5D-5L and Independent Living, Sense and Pain for the AQoL-8D for each level before and after surgery.

Second, impacts on the individual domains of physical and psychosocial HRQoL were investigated through the AQoL-8D's summary scores for the eight individual dimensions and two super dimensions. The EQ-5D-5L generates a single utility valuation for an individual; however, it does not generate individual or summary scores for each and every one of its five separate dimensions.

2.2.2 Analyses of Summary Utility Valuations and EQ-VAS Scores

Utility valuations were generated for the EQ-5D-5L using the most recent UK value based on directly elicited preferences, the valuation ranging from -0.281 to 1.0 utility points [43, 44]. All five questions require a valid response to generate a utility score. EQ-5D population norms are sourced from UK data because there are no available Australian population norms [45]. For the AQoL-8D, we used the current version of the scoring algorithm incorporating Australian weights published on the AQoL group's website (<http://www.aqol.com.au>) (valuation range $+0.09$ to 1.0 utility points). For the AQoL-8D's scoring algorithm, an overall utility valuation can be generated with ten missing values scattered over all dimensions. Australian population norms were sourced from recently published valuations [41]. Individual and super-dimensional scores are also generated with the AQoL-8D's scoring algorithm.

A minimal clinically important difference (or minimal important difference) is the smallest difference in score in the outcome of interest which patients perceive as beneficial and which would mandate a change in the patient's management [23, 29, 30]. A recently reported composite minimal important difference for the EQ-5D-5L for chronic health conditions was reported as 0.04 utility points [46]. There is no established minimal important difference for the AQoL-8D; however, a minimal important difference for the AQoL-4D has previously been reported as 0.06 utility points, with a 95% confidence interval of 0.03 – 0.08 utility points [24]. This study conservatively adopted the upper bound of 0.08 utility points as the proxy minimal important difference for comparison of the pre- and postoperative AQoL-8D utility valuations. The established minimal important difference for the EQ-VAS is 10 points [47]. It has been suggested that with the expanded use of HRQoL endpoints (for example, analyses of utility valuations and scores within vastly different MAUI classification systems), the interpretation of HRQoL in the context of minimal important differences is imperative [23]. In turn, our study has included the interpretation of minimal important differences in its comparison of the EQ-5D-5L and AQoL-8D MAUIs.

Statistical analyses were undertaken using IBM SPSS (version 22) or R (version 3.0.2).

3 Results

3.1 Participants' Characteristics and Questionnaire Completion

Twenty-three participants were recruited to the study. For these participants, mean (SD) age was 50 (10) years, 43% were males, and mean (SD) and median (IQR) time on the public waiting list for bariatric surgery was 6.5 (2.0) and 6.3 (5.0–7.8) years, respectively.

Table 1 provides pre- and postoperative results for BMI, percentage total weight lost and percentage excess weight lost. Before surgery 39% of participants were classified as super obese (BMI ≥ 50 kg/m²) and 57% were classified as morbidly obese (BMI 40–49.9 kg/m²). After surgery, there was a 26% reduction in the super-obesity category. Similarly, after surgery, the morbidly obese category was reduced by 17%.

In regard to questionnaire completion, there was a 74% completion rate of questionnaires overall [Tables 2, 3 and 4; Appendix 1 (see the ESM)].

3.2 Sensitivity: Dimensional Comparisons

The relative discriminatory power of the instruments was investigated using the dimensional comparisons outlined in Sect. 2.2.1.

Table 2 (supported by Appendix 1 in the ESM) presents the 'before' and 'after' surgery distribution of participant responses for both MAUIs' 13 individual dimensions/domains of health across levels 1–5 (EQ-5D-5L) and levels 1 through to 4, 5 or 6 (AQoL-8D). Figure 1a–c also provide a schematic representation of the comparative distribution of the participants' responses across levels 1–6 for all dimensions (Fig. 1a), and for the physical dimensions of health for both instruments (EQ 5D-5L: Mobility, Self-care, Usual Activities and Pain; AQoL-8D: Independent Living, Senses and Pain) (Fig. 1b), and the psychosocial dimensions of health for both instruments (EQ-5D-5L: Anxiety/Depression; AQoL-8D: Coping, Mental Health, Relationships, Self-worth, Happiness) (Fig. 1c).

None of the participants responded to level 6 for the AQoL-8D items that provided for a level 6 response [namely Independent Living (one item), Senses (two items: vision and hearing), Mental Health (one item) and Relationships (one item)] (Table 2 and Appendix 1). Table 2 and Fig. 1a–c (supported by Appendix 1) revealed a more even dispersion of participant responses for the AQoL-8D than the EQ-5D-5L both pre- and postoperatively. The AQoL-8D more clearly distinguished between health states that are close to full health for the study population (Table 2, Fig. 1a–c, Appendix 1).

Table 1 Number of participants ($n = 23$) in obesity categories before and after surgery

	Before surgery	After surgery*	Change
BMI			
Mean (SD)	49.3 (9.4)	43.5 (7.2)	- 5.8
Median (IQR)	45.5 (41.6–55.4)	43.2 (38.7–49.6)	- 2.3
BMI (n, %)			
BMI ≥ 30 –34.9 kg/m ² (class I)	(1, 4%)	(2, 9%)	(+ 1, + 6%)
BMI ≥ 35 –39.9 kg/m ² (class II)	0	(7, 30%)	(+ 7, + 33%)
BMI ≥ 40 –49.9 kg/m ² (class III)	(13, 57%)	(9, 39%)	(- 4, - 17%)
BMI ≥ 50 kg/m ² **	(9, 39%)	(3, 13%)	(- 6, - 26%)
Weight (kg)			
Mean (SD)	139.7 (31.4)	125.9 (26.9)	- 13.8
Median (IQR)	134.0 (118.8–161.5)	124.5 (106.9–142.2)	- 8.1
% Total weight lost			
Mean (SD)	NA	9.9 (6.2)	NA
Median (IQR)	NA	11.0 (3.7–15.0)	NA
% Excess weight lost			
Mean (SD)	NA	21.5 (13.1)	NA
Median (IQR)	NA	24.7 (12.6–28.2)	NA

BMI body mass index, IQR interquartile range, NA not applicable, SD standard deviation

* $N = 21$; 2 participants' 3-month weight not available

**Super obese (≥ 50 kg/m²)

More specifically, postoperatively participants recorded 80% (76/95) of responses for the EQ-5D-5L at level 1 (perfect health: I have no problems) and level 2 (I have slight problems), the highest recorded response at level 1 being 74% for Self-care (decreased from 81% before surgery) (Table 2; Appendix 1). These results highlight the EQ-5D-5L's inability to distinguish between health states close to full/perfect health (utility score 1.0). Additionally, for the EQ-5D-5L's only psychosocial dimension of health (Anxiety/Depression), participants did not record responses at level 4 (I am severely anxious or depressed), nor level 5 (I am extremely anxious or depressed), indicating that the EQ-5D-5L's only psychosocial dimension is relatively limited. Before surgery, only 6% of participants recorded both levels 4 and 5 for Anxiety/Depression (Table 2, Appendix 1, and Fig. 1c). Participants recorded responses at level 4 (16%) for one of the EQ-5D-5L's individual dimensions (Pain) after surgery (Table 2; Appendix 1).

In contrast, participants' postoperative responses to the AQoL-8D questionnaire were less concentrated in the upper levels (i.e. more evenly dispersed across the levels), with only 58% (365/630) of responses recorded at levels 1 and 2 (Table 2, Fig. 1a, and Appendix 1), the highest recorded response at level 1 being 41% for Senses.

Participants also recorded responses at level 4 for all the AQoL-8D's individual dimensions, and participants also recorded responses at level 5 for both Pain and Mental Health. Additionally, the lowest percentage of participants scored at level 1 for the AQoL-8D's individual dimensions

of Happiness (15%), Coping (19%) and Mental Health (26%) (Table 2; Appendix 1). Nevertheless, Happiness and Coping substantially improved from before surgery to 3 months after surgery and approached population norms (Table 3), and this result is also revealed with the improvement of participants' preoperative scores at level 1 in Happiness (from 3% to 15%) and Coping (from 11% to 19%) (Table 2; Appendix 1).

The individual dimension that had the most similar distribution for both instruments across levels 1–5 was Pain/Discomfort for the EQ-5D-5L (level 1: 26%, level 2: 32%, level 3: 26%, level 4: 16% and level 5: 0%) and Pain for the AQoL-8D (level 1: 35%, level 2: 19%, level 3: 31%, level 4: 13% and level 5: 2%) (Table 2; Appendix 1). Three of the 35 AQoL-8D items contribute to the dimension of Pain. These items capture and assess how often the respondent suffers for the first Pain item 'serious pain', for the second Pain item the severity of 'pain or discomfort', and for the third Pain item of how often pain interferes with usual activities. The EQ-5D-5L individual dimension of Pain/Discomfort assesses the level of severity of pain/discomfort (no pain/discomfort, slight, moderate, severe, extreme).

3.3 Sensitivity: Comparison of Changes in Utility Valuations

Table 4 provides summary statistics for the changes in both instruments' utility valuations preoperatively to 3 months

Table 2 Dimensional comparisons of response rates (%) for all individual dimensions of the EQ-5D-5L and AQL-8D before surgery and 3 months after bariatric surgery

EQ-5D-5L		AQL-8D										
		Physical super dimension					Psychosocial super dimension					
Mobility	Self-care	Usual Activities	Pain/Discomfort	Anxiety/Depression	Independent Living (4 items)	Pain (3 items)	Senses (3 items)	Coping (3 items)	Mental Health (8 items)	Happiness (4 items)	Relationships (7 items)	Self-worth (3 items)
Before surgery (<i>n</i> = 16)												
Level 1	38%	81%	31%	25%	Level 1	27%	29%**	31%	11%	19%**	3%	24%**
Level 2	25%	6%	13%	38%	Level 2	17%	20%	40%	36%	23%	33%	20%
Level 3	31%	13%	25%	25%	Level 3	30%	27%	29%	29%	36%	42%	33%
Level 4	6%	0	31%	6%	Level 4	18%	18%	0	20%	15%	17%	24%
Level 5	0	0	0	6%	Level 5	8%	7%	0	4%	8%	5%	4%
Level 6	NA	NA	NA	NA	Level 6	0	NA	0	NA	0	NA	NA
After surgery (<i>n</i> = 19); % (% point change) [†]												
Level 1	53% (+15)	74% (-7)	26% (-5)	47% (+22)	Level 1	26% (-1)	35% (+6)	41% (+10)	19% (+8)	26% (+7)	15% (+12)	30% (+9)
Level 2	26% (+1)	21% (+15)	32% (+19)	37% (-1)	Level 2	31% (+14)	19% (-1)	30% (-10)	44% (+8)	24% (+1)	38% (+5)	26% (+6)
Level 3	21% (-10)	5% (-8)	26% (+1)	16% (-9)	Level 3	29% (-1)	31% (+4)	26% (-3)	24% (-5)	44% (+8)	39% (-3)	31% (-2)
Level 4	0 (-6)	0 (0)	16% (-15)	0 (-6)	Level 4	14% (-4)	13% (-5)	4% (+4)*	13% (-7)	4% (-11)*	8% (-9)	10% (+1)
Level 5	0 (0)	0 (0)	0 (0)	0 (-6)	Level 5	0 (-8)	2% (-5)	0 (0)	0 (0)	1% (-7)	0 (-5)	0 (-4)
Level 6	NA	NA	NA	NA	Level 6	0 (0)	NA	0 (0)	NA	0	NA	0 (0)

Detailed calculations supporting Table 2 are contained in Appendix 1 (see the electronic supplementary material)

AQL-8D items: AQL-8D Independent Living: 4 AQL-8D items: household tasks levels 1–5, getting around levels 1–6, mobility levels 1–6, self-care levels 1–5; AQL-8D Pain: 3 AQL-8D items: frequency of pain levels 1–4, degree of pain levels 1–4, pain interference levels 1–5; AQL-8D Senses: 3 AQL-8D items: vision levels 1–6, hearing levels 1–6, communication levels 1–4; AQL-8D Coping: 3 AQL-8D items: energy levels 1–5, being in control levels 1–5, coping with problems levels 1–5; AQL-8D Mental Health: 8 AQL-8D items: feelings of depression levels 1–6, trouble sleeping levels 1–5, feelings of anger levels 1–5, self-harm levels 1–5, feeling despair levels 1–5, worry levels 1–5, sadness levels 1–5, tranquility/agitation levels 1–5; AQL-8D Happiness: 4 AQL-8D items: contentment levels 1–5, enthusiasm levels 1–5, degree of feeling happiness levels 1–5, pleasure levels 1–5; AQL-8D Relationships: 7 AQL-8D items: relationship with family and friends levels 1–6 and levels 1–5, social isolation levels 1–5, social exclusion levels 1–5, intimate relationship levels 1–5, family role levels 1–5, community role levels 1–4; and AQL-8D Self-worth: 3 AQL-8D items: feeling like a burden levels 1–5, worthlessness levels 1–5, confidence levels 1–5

AQL-8D Assessment of Quality of Life, NA not applicable

*Columns add to 99 or 101 per cent due to rounding

[†]Figure in brackets after surgery reflects the percentage point change (i.e. 'after surgery' minus 'before surgery')

Table 3 Comparison of AQoL-8D individual and super-dimension scores before surgery and 3 months after surgery (total participants $n = 23$), and Australian population norms for total population and 45–54-year age group

	Before bariatric surgery ($n = 15$)			After bariatric surgery ($n = 18$)			Improvement in mean score preoperatively to 3 months postoperatively		Australian population norms	
	Mean (SD)	Min	Max	Mean (SD)	Min	Max	Change	45–54 years		Total
								Mean (SD)	Mean (SD)	
Dimensions of physical health										
Independent Living	0.69 (0.22)	0.39	1.00	0.75 (0.19)	0.41	1.00	+ 0.06	0.93 (0.12)	0.94 (0.11)	
Senses	0.81 (0.13)	0.56	1.00	0.83 (0.13)	0.59	1.00	+ 0.02	0.88 (0.10)	0.91 (0.10)	
Pain	0.56 (0.34)	0.16	1.00	0.62 (0.32)	0.21	1.00	+ 0.06	0.84 (0.21)	0.86 (0.19)	
Dimensions of psychosocial health										
Happiness	0.65 (0.16)	0.32	0.85	0.75 (0.15)	0.51	1.00	+ 0.10	0.77 (0.16)	0.80 (0.15)	
Coping	0.67 (0.15)	0.39	0.96	0.76 (0.15)	0.51	1.00	+ 0.09	0.80 (0.16)	0.83 (0.15)	
Relationships	0.62 (0.16)	0.47	1.00	0.67 (0.18)	0.47	1.00	+ 0.05	0.78 (0.16)	0.79 (0.16)	
Self-worth	0.65 (0.21)	0.35	1.00	0.76 (0.18)	0.39	1.00	+ 0.11	0.84 (0.16)	0.85 (0.15)	
Mental Health	0.54 (0.12)	0.28	0.73	0.60 (0.15)	0.36	0.96	+ 0.06	0.67 (0.17)	0.69 (0.17)	
Super-dimensions										
Physical super dimension	0.51 (0.29)	0.18	0.97	0.56 (0.27)	0.22	1.00	+ 0.05	0.79 (0.20)	0.83 (0.18)	
Psychosocial super dimension	0.25 (0.15)	0.08	0.49	0.37 (0.25)	0.12	0.97	+ 0.12	0.47 (0.24)	0.50 (0.24)	
Utility value for AQoL-8D	0.51 (0.24)	0.20	0.83	0.61 (0.24)	0.29	1.00	+ 0.10	0.77 (0.20)	0.80 (0.19)	

Max maximum, *Min* minimum, *SD* standard deviation

Table 4 Summary statistics for EQ-5D-5L and AQoL-8D at baseline (before surgery), difference between the two measures at baseline, and changes in the participants' scores over the 3 months of follow-up (total participants $n = 23$)

	EQ-5D-5L (baseline) ($n = 16$)	EQ-5D-5L (after surgery) ($n = 19$)	EQ-5D-5L change (after surgery– baseline)	AQoL-8D (baseline) ($n = 15$)	AQoL-8D (after surgery) ($n = 18$)	AQoL-8D change (after surgery– baseline)	Difference in baseline/after surgery scores: (EQ-5D-5L– AQoL-8D)	EQ-VAS (baseline) ($n = 16$)	EQ-VAS (after surgery) ($n = 19$)	EQ-VAS change (baseline to 3 months)
Mean	0.70	0.80	0.10	0.51	0.61	0.10	0.19/0.19	57	67	10
SD	0.25	0.25	0	0.24	0.24	0	0.01/0.10	25	24	(1)
Median	0.73	0.84	0.11	0.51	0.58	0.07	0.22/0.29	65	65	0
IQR	0.54–0.91	0.59–0.86	NA	0.29–0.78	0.43–0.78	NA	NA	34–73	48–90	NA
Minimum	0.24	0.46	0.22	0.20	0.29	0.09	0.04/0.17	15	27	12
Maximum	1.00	1.00	0	0.83	1.00	0.17	0.17/0	95	99	4

IQR interquartile range, NA not applicable, SD standard deviation, VAS visual analogue scale

postoperatively. The EQ-5D-5L revealed relatively higher summary utility valuations than the AQoL-8D both before and after surgery. Specifically, the order of magnitude of the EQ-5D-5L's mean utility valuations were 0.19 utility points greater than the mean AQoL-8D utility valuations preoperatively and 3 months postoperatively. The AQoL-8D particularly showed low summary utility valuations before surgery [EQ-5D-5L 0.70 (0.25); AQoL-8D 0.51 (0.24)].

Three months after surgery, the summary utility valuations revealed clinical improvements for both instruments. Nonetheless, the AQoL-8D showed substantially lower postoperative summary utility valuations than the EQ-5D-5L. More specifically, the EQ-5D-5L utility value increased by 0.10 points from mean (SD) 0.70 (0.25) to 0.80 (0.25). Similarly, the AQoL-8D utility value increased by 0.10 points from 0.51 (0.24) to 0.61 (0.24) (Table 4).

After surgery, the EQ-5D-5L utility valuations approached comparable population norms, but not so the AQoL-8D's utility valuations. The UK general population mean for the EQ-5D-5L is 0.86 [45], and for the AQoL-8D the general Australian population norm is 0.80 (0.19), and for the 45–54-year age group, it is 0.77 (0.20) [41] (Table 4).

Table 4 also provides mean (SD) pre- and postoperative EQ-VAS scores of 57 (25) to 67 (24) points, the difference equalling the established EQ-VAS minimal important difference of 10 points.

3.4 AQoL-8D Individual/Super-Dimension Scores

Table 3 provides the AQoL-8D's individual and super-dimension scores before surgery and 3 months after surgery, and the Australian population norms at the individual dimensional level for the general population and the 45–54-year age group. Additionally, Fig. 2a, b provide a schematic representation of the individual and super-dimension scores compared with the general Australian population norm. The EQ-5D-5L does not generate individual or super-dimension scores.

Improvements were observed for all eight individual dimension scores and the two super-dimension scores even 3 months after surgery. Three months after surgery, the Physical super dimension improved 0.05 points to mean (SD) 0.56 (0.27) points and the Psychosocial super-dimension score improved 0.12 points to 0.37 (0.25) points. Of the eight individual dimensional scores, Self-worth and Happiness improved the most 3 months after surgery by revealing gains of 0.11 points (Self-worth) and 0.10 points (Happiness). The postoperative scores for Happiness 0.75 (0.15) and Coping 0.76 (0.15) also approached both the 45–54-year age group and general population norms. Happiness was only 0.02 points less than the 45–54-year age group population norm and Coping was only 0.04

points less than the 45–54-year age group population norm. Other individual dimensional scores that improved by ≥ 0.05 points after surgery were Coping (0.09 points), Mental Health (0.06 points) and Relationships (0.05 points), which contribute to the Psychosocial super dimension. With regard to the Physical super dimension, Independent Living and Pain both improved 0.06 points and Senses showed a smaller improvement of 0.02 points (Table 3).

As mentioned previously, the cohort's HRQoL before surgery was substantially lower in comparison to population norms (Table 3; Fig. 2a, b). Individual dimensional scores improved 3 months postoperatively, but did not substantially approach Australian population norms, with the exception of two dimensions: Happiness and Coping (Table 3; Fig. 2a). The Psychosocial and Physical super dimensions' scores, while improved, were still substantially lower than the Australian general population norm at -0.13 and -0.27 points, respectively. The Physical super-dimension score was driven by the Pain dimension scoring 0.24 points less than the general population norm. Independent Living and Relationships also revealed large differences, scoring -0.19 and -0.13 points from the general population norm. Similarly, Mental Health/Self-worth and Senses also revealed scores of 0.09/0.09 and 0.08 less than their Australian general population norm equivalents, respectively. In contrast, the individual dimensions of Happiness and Coping approached both the general and 45–54-year age group population norms (Table 3; Fig. 2a, b).

4 Discussion

Our study is important because it is the first study to investigate the relative discriminatory power using dimensional comparisons of all 13 individual dimensions of the EQ-5D-5L and AqoL-8D for patients who endured multiyear wait times in a public health system and then underwent bariatric surgery.

As an important and emerging subgroup of bariatric surgery patients, our cohort also delivered an important and novel opportunity to provide clinicians with a better understanding of the 3-month postoperative impact of bariatric surgery on long-term waitlisted patients' complex physical and psychosocial domains of health.

4.1 A Head-to-Head Comparison of the EQ-5D-5L and AqoL-8D Revisited

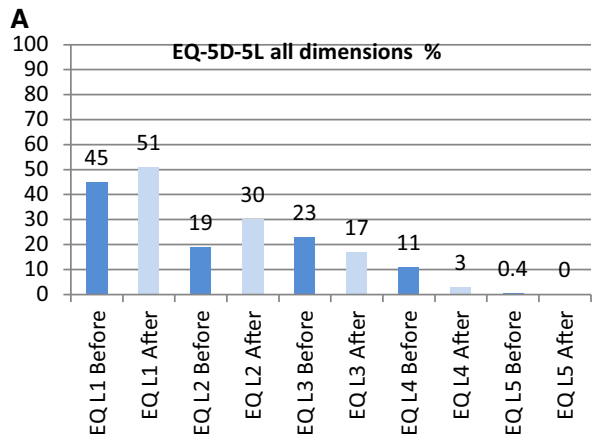
In support of our findings from our previously published cross-sectional head-to-head comparison of privately treated patients who received bariatric surgery many years

Fig. 1 a Distribution of participants' responses (%) for levels (L) 1–5 for all dimensions of EQ-5D-5L and AqoL-8D before surgery and 3 months after surgery. **b** Distribution of participants' responses (%) for Levels (L) 1–5 for the combined physical dimensions of EQ-5D-5L (Usual Activities, Self-care, Mobility, Pain) and AqoL-8D (Independent Living, Senses, Pain) before surgery and 3 months after surgery. **c** Distribution of participants' responses (%) for Levels (L) 1–5 for the combined psychosocial dimensions of EQ-5D-5L (Anxiety/Depression) and AqoL-8D (Coping, Mental Health, Happiness, Relationships, Self-worth) before surgery and 3 months after surgery

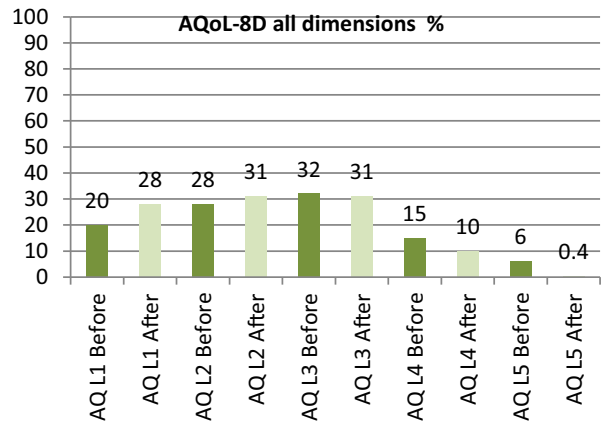
previously [9], this current longitudinal study revealed that the AqoL-8D preferentially captured and assessed the physical and psychosocial HRQoL for our cohort of long-term waitlisted patients who subsequently underwent bariatric surgery, even 3 months after their surgery.

Amongst other direct comparisons of the discriminatory power of the two instruments, our earlier head-to-head study's comparison of the patient-reported distribution of the levels of response compared three (total six) individual comparable dimensions of both instruments (EQ-5D-5L: Anxiety/Depression, Self-care, Pain/Discomfort; AqoL-8D: Mental Health, Independent Living, Pain) [9]. In contrast, this current paper's head-to-head comparison conducted a longitudinal investigation for a study population of long-term publicly waitlisted bariatric surgery patients who underwent bariatric surgery as a targeted government policy decision to reduce waiting lists. Compared with our earlier study's examination of six individual dimensions, we investigated the patient-reported distributions of responses for the dimensional comparisons of all 13 individual dimensions of health for both instruments. Consequently, this study included an additional four (of the five) psychosocial domains of health for the AqoL-8D's classification system.

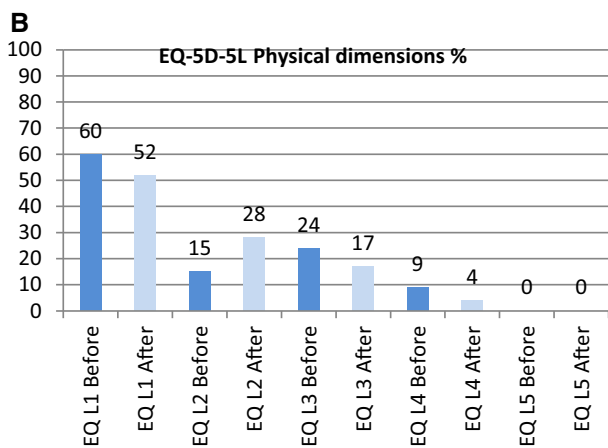
This current study particularly highlighted the depth and breadth of the AqoL-8D's classification system as compared to the EQ-5D-5L. Table 2 and Appendix 1 (see the ESM), coupled with schematic representations (Fig. 1a–c) of the dimensional comparisons, revealed that the AqoL-8D assessed and captured HRQoL across the broad classification system and through the levels (1 to 4–6) (there were no reported responses for level 6 for the AqoL-8D) given the relative dispersion of participants' responses away from perfect health. This is particularly highlighted with many of the responses for the EQ-5D-5L at level 1 (perfect health/ceiling effect) and level 2, compared to the AqoL-8D only recording just over half of the responses at levels 1 and 2. These findings support the superior discriminant sensitivity of the AqoL-8D across the individual dimensions of physical and psychosocial health for the study population and as assessed in our previously published work [9].



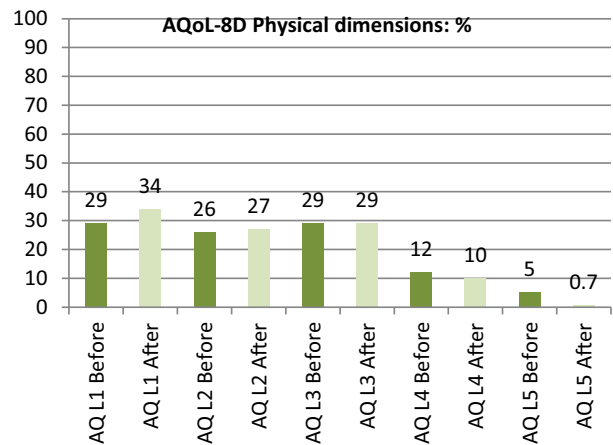
EQ-5D-5L five individual dimensions: Usual Activities, Self-care, Mobility, Pain and Anxiety/Depression.



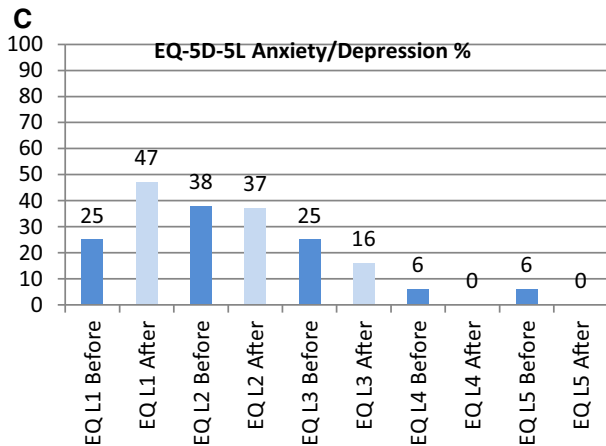
AQoL-8D eight individual dimensions: Independent Living, Senses, Pain, Happiness, Coping, Relationships, Self-worth, Mental Health.



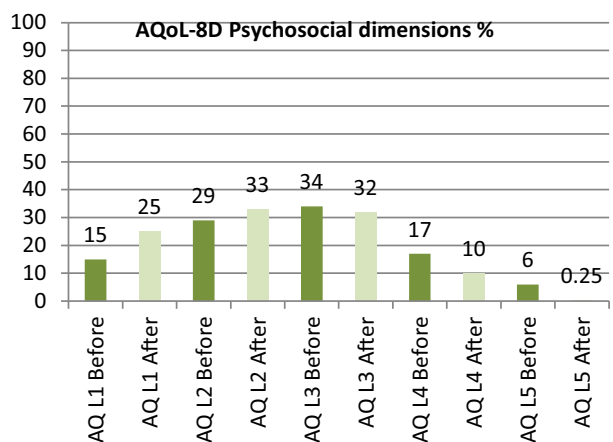
EQ-5D-5L four physical dimensions: Usual Activities, Self-care, Mobility, and Pain.



AQoL-8D three physical dimensions: Independent Living, Senses, Pain.

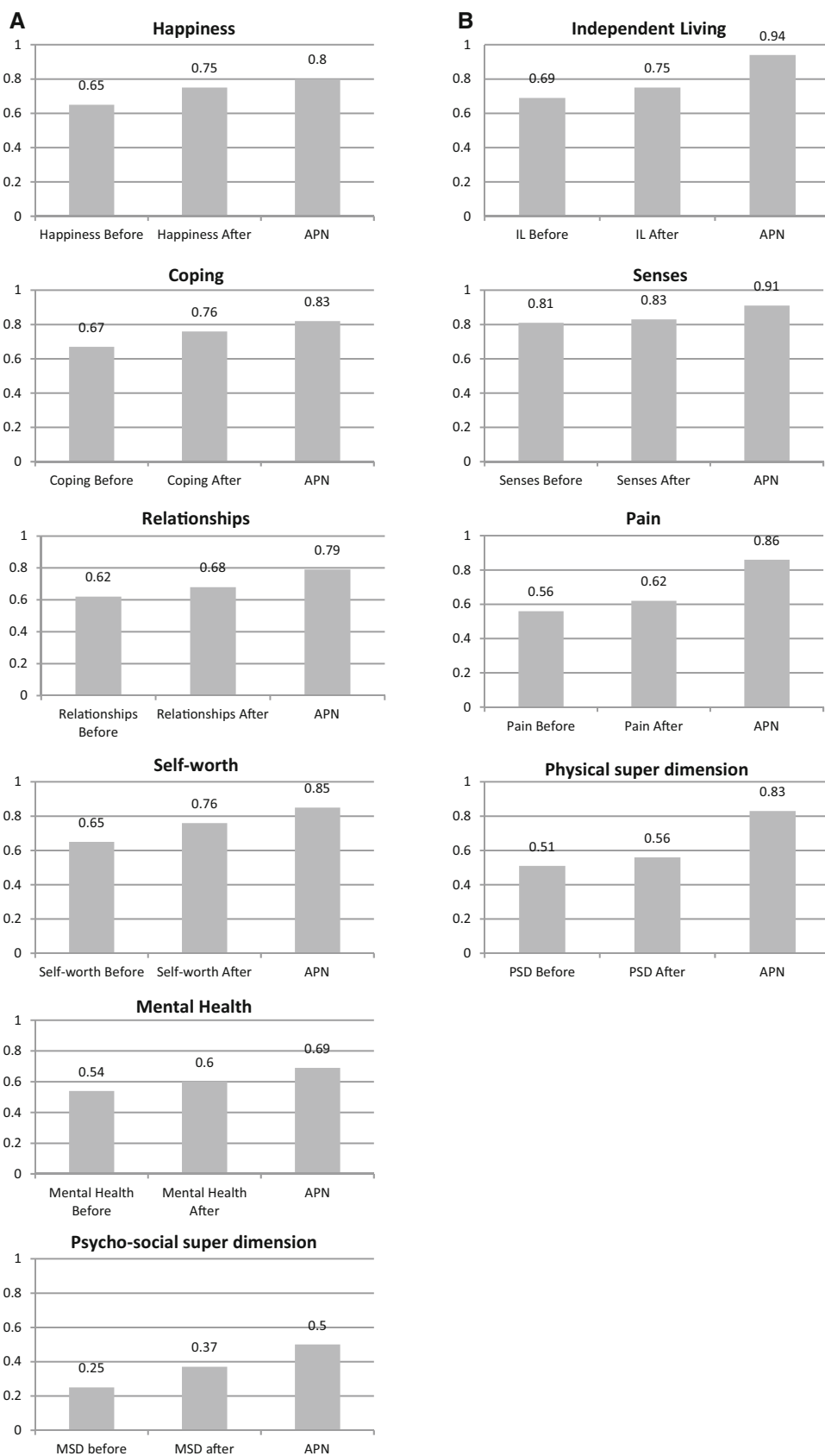


EQ-5D-5L one psychosocial dimension: Anxiety/Depression.



AQoL-8D five psycho-social dimensions: Happiness, Coping, Relationships, Self-worth, Mental Health.

Fig. 2 a Comparison of before surgery and 3 months after bariatric surgery AQL-8D scores and Australian Population norms (APN) for the individual psychosocial dimensions (Happiness, Coping, Relationships, Self-worth, Mental Health) and the Psychosocial super dimension. **b** Comparison of before surgery and 3 months after bariatric surgery AQL-8D scores and Australian Population norms (APN) for the individual physical dimensions (Independent Living, Senses, Pain) and the Physical super dimension



This study's dimensional comparisons also found the individual dimension that revealed the most similar distribution for both instruments was Pain/Discomfort (EQ-5D-5L) and Pain (AQoL-8D). Therefore, our study's results suggest that both instruments were sensitive to the individual health domain of pain for the study population. Nevertheless, the AQoL-8D provided evidence of change in other domains of health that could be affected by pain, such as sleep, which impacts the Mental Health dimension.

Another key finding of our current study was that the pre- and postoperative summary utility valuations for the EQ-5D-5L were substantially higher (and indeed approached general population norms after surgery) than the summary utility valuations of the AQoL-8D. The AQoL-8D's relatively low preoperative and 3-month postoperative summary utility valuation revealed two important findings: first, the instrument's superior discriminant sensitivity relative to the EQ-5D-5L for the study population due to the AQoL-8D's ability to preferentially capture domains of health that are relevant for the study population; and second, the substantially lower (particularly preoperative) HRQoL for the long-term publicly waitlisted bariatric surgery patients. These findings also accord with evidence that suggests in practice all MAUIs which purport to measure utility give numerical values that differ significantly [12, 41].

4.2 Utility Valuations

Another key finding of our current study was that change in global utility valuations from before to 3 months after bariatric surgery exceeded the established minimal important differences for both instruments, and for the EQ-VAS. The instruments' summary utility valuations highlighted these long-term waitlisted bariatric surgery patients' considerably diminished physical and psychosocial health status before surgery, and the postoperative summary utility valuations revealed a clinical short-term improvement within the 3-month timeframe. Nevertheless, as discussed previously, compared to the EQ-5D-5L, the AQoL-8D revealed substantially lower pre- and postoperative utility valuations that did not approach population norms.

In particular, this study highlighted the substantially diminished preoperative AQoL-8D utility valuation for our study population. To provide a comparative perspective of the severity of our study population's diminished health state, a recent investigation that used data from a multinational (Australia, Canada, Germany, Norway, the UK and USA) cross-sectional survey found that for composite study populations of people with cancer or heart disease, the AQoL-8D mean (SD) utility valuation for cancer was 0.655 (0.22), and for heart disease, it was 0.667 (0.23) [48].

Therefore, our current study's findings particularly revealed that the preoperative AQoL-8D utility valuation for our cohort of severely obese long-term waitlisted patients was over 0.15 utility points less than that for a study population with cancer or heart disease. In other words, people who languish for long periods on the public waiting list can endure the same substantially diminished HRQoL status as someone with metastatic cancer or prolonged heart disease.

As an independent measure of HRQoL, there is emerging literature that suggests that utility valuations could be independent predictors of health outcomes. A study that investigated the predictive qualities of utility valuations derived from the EQ-5D in patients with diabetes found that they were useful in predicting for health events, including cardiovascular events (e.g. stroke, hospitalisation for angina), other major diabetes-related complications (e.g. heart failure, amputation, renal dialysis and lower extremity ulcer) and death from any other cause [21]. Bariatric surgery patients carry complex physical and psychosocial comorbidity loads, and the assessment of utility valuations in routine clinical care could provide a better understanding of this complexity at an individual patient level, informing preoperative and ongoing postoperative care. Prediction is more likely to be accurate when the instrument used for prediction takes account of the full range of the complex physical and psychosocial problems associated with the problem. Our study's findings suggest that the AQoL-8D is more likely to provide correct prediction than the EQ-5D-5L.

4.3 AQoL-8D's Individual and Super-Dimension Scores

Another key finding of our current study was the substantially lower AQoL-8D dimensional scores before surgery and improvements in these dimensional levels after surgery. Happiness and Coping improved the most after surgery and indeed approached population norms. Additionally, Self-worth also revealed a substantial change. All other individual dimensions improved, but did not substantially approach population norms. Recent evidence has found that body weight is only one contributing factor to the complex physical and psychosocial HRQoL needs of bariatric surgery patients [8].

4.4 Integrating Patient-Reported Outcomes in Clinical Practice

The International Society for Quality of Life Research has developed a clinical users guide to encourage the routine collection of PROs which "are rarely collected in routine clinical practice" [49]. Recent evidence has also found that

integrating PROs in clinical practice has the potential to enhance patient-centred care. Within this broader and evolving context of patient-centredness in clinical care, our exploratory study highlighted the clinical relevance of MAUI analyses for long-term waitlisted patients who subsequently undergo bariatric surgery.

This study found that psychosocial health drove a relatively lower utility valuation for the AQoL-8D, despite clinical improvements. We suggest that bariatric clinicians could also further investigate and subsequently integrate and implement utility valuation's predictive qualities, and individual and super-dimension scores to further enhance patient-centred clinical care. Further studies could assess the feasibility of adopting a MAUI that preferentially captures and assesses physical and psychosocial HRQoL into the routine clinical assessment of these patients. We previously identified in our earlier published work that the AQoL-8D preferentially captured physical and psychosocial health for patients who had undergone bariatric surgery (in the longer term) [9], a position reinforced by our current analysis. Through MAUI analyses, our current study established clinically significant changes in psychosocial health (albeit from a relatively low baseline to post-surgical dimensional scores that were still relatively low) that warrant additional attention after surgery to improve overall postoperative health. Additionally, our current study's dimensional comparisons highlighted the EQ-5D-5L's relative insensitivity in distinguishing between health states close to full (or perfect) health for long-term waitlisted patients who had very recently undergone bariatric surgery.

4.5 Limitations

There are limitations to our study. The first limitation is small sample size. Nevertheless, our study was exploratory and we were provided with a novel opportunity to recruit participants from the long-term waitlisted patients subsequently fast-tracked for bariatric surgery through a government policy decision to reduce waiting lists. Our exploratory study of long-term waitlisted patients should inform larger confirmatory studies to test the validity of the EQ-5D-5L and AQoL-8D, and the short-term health impacts for long-term waitlisted patients. Nevertheless, we also acknowledge that a substantial commitment would need to be made at the public policy level to recruit a similar cohort of long-waiting patients. Other MAUIs such as the SF-6D could also be considered for larger confirmatory studies. The second limitation is that all participants were operated on by the same surgeon in the same hospital. This could affect the generalisability of our results if scaled up to all bariatric surgery patients. On the other

hand, this circumstance could also be a strength given the homogenous nature of the sample.

The third limitation is that there is no control arm in the study. The observational nature of our study did not enable the recruitment of a control arm to elicit utility valuations; however, the key objective of this study was to compare the two MAUI. The final limitation is that the sample is at risk of participant selection bias, which could also affect the generalisability of our results.

A relative strength of our study is the high overall response rate of 74% to the questionnaires across the two time points.

The limitations of our study concur with our complementary study of the same cohort [11].

5 Conclusions

Within the small sample limitations of our exploratory study and to address the key objective of our study, which was a head-to-head comparison of the instruments, compared to the EQ-5D-5L, the AQoL-8D preferentially captured the complex physical and psychosocial short-term health changes for long-term publicly waitlisted patients who very recently underwent bariatric surgery. Importantly, researchers should understand a MAUI's descriptive/classification system and the innate sensitivities of the MAUI in regard to the particular study population, in this case long-term waitlisted patients who then undergo bariatric surgery. We recommend the AQoL-8D as a preferred MAUI over the EQ-5D-5L for bariatric surgery patients, given their complex physical and psychosocial needs.

In regard to our secondary objectives, utility valuations and dimensional scores (AQoL-8D only) revealed substantially lower health status for long-term waitlisted patients both before and after surgery, but with clinical short-term HRQoL improvements even 3 months after surgery. AQoL-8D preoperative utility valuation particularly revealed our study population's HRQoL was substantially lower than that of people with cancer or heart disease.

Dimensional comparisons, utility valuations, and individual and super-dimension scores could provide the clinician with both individual patient and cohort valuations that could lead to improved patient-centred care by identifying health domains requiring additional attention.

Routine integration of comprehensive MAUI analyses could provide clinicians with additional and independent assessments and predictors of HRQoL and in turn, enhance patient-centred care.

Data Availability Statement The dataset used for this study contains the following: the participant-reported responses to the EQ-5D-

5L and the AQoL-8D MAUIs' questionnaires; the individual utility valuations (both instruments) and utility scores (AQoL-8D only) that were generated with the instruments' specific algorithm; the participant-reported EQ-VAS scores; and participants' socio-demographic and clinical data. The corresponding author will provide a de-identified dataset upon reasonable request for all or part of the data.

Author Contributions Julie Campbell contributed to study design, data collection, data verification and analysis, manuscript preparation, and final approval. Martin Hensher contributed to study design, manuscript review and final approval. Amanda Neil contributed to study design, manuscript review and final approval. Alison Venn contributed to study design, manuscript review and final approval. Petr Otahal contributed to the data analysis, manuscript review and final approval. Stephen Wilkinson contributed to study design, manuscript review and final approval. Andrew Palmer contributed to study design, data verification and analysis, manuscript review and final approval.

Compliance with Ethical Standards

The University of Tasmania's Health and Medical Human Research Ethics Committee (HMHREC) approved this study, and informed consent was obtained from all participants in accordance with the HMHREC's guidelines. This study was performed in accordance with the ethical standards of the Declaration of Helsinki.

Funding This work was supported by a National Health and Medical Research Council (NHMRC) Partnership Project Grant (APP1076899). Alison Venn is supported by an NHMRC fellowship grant. Amanda Neil is supported by a Select Foundation Research Fellowship.

Conflict of interest The authors, Julie Campbell, Martin Hensher, Amanda Neil, Alison Venn, Petr Otahal, Stephen Wilkinson and Andrew Palmer, declare they have no conflict of interest.

Informed Consent Informed consent was obtained from all individual participants included in the study.

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Owen-Smith A, Kipping R, Donovan J, Hine C, Maslen C, Coast J. A NICE example? Variation in provision of bariatric surgery in England. *BMJ* (Online). 2013;346:f2453. doi:10.1136/bmj.f2453.
- Padwal RS, Majumdar SR, Klarenbach S, Birch DW, Karmali S, McCargar L, et al. Health status, quality of life, and satisfaction of patients awaiting multidisciplinary bariatric care. *BMC Health Serv Res*. 2012;12(1):139.
- Curtis AJ, Russell CO, Stoelwinder JU, McNeil JJ. Waiting lists and elective surgery: ordering the queue. *MJA*. 2010;192(4):217–20.
- Sharman MJ, Hensher M, Wilkinson S, Campbell JA, Venn AJ. Review of publicly-funded bariatric surgery policy in Australia—lessons for more comprehensive policy making. *Obes Surg*. 2016;26(4):817–24.
- Christou NV, Efthimiou E. Bariatric surgery waiting times in Canada. *Can J Surg*. 2009;52(3):229–34.
- Padwal RS, Sharma AM. Treating severe obesity: morbid weights and morbid waits. *CMAJ*. 2009;181(11):777–8.
- Padwal RS, Rueda-Clausen CF, Sharma AM, Agborsangaya CB, Klarenbach S, Birch DW, et al. Weight loss and outcomes in wait-listed, medically managed, and surgically treated patients enrolled in a population-based bariatric program: prospective cohort study. *Med Care*. 2014;52(3):208–15.
- Herpertz S, Muller A, Burgmer R, Crosby RD, de Zwaan M, Legenbauer T. Health-related quality of life and psychological functioning 9 years after restrictive surgical treatment for obesity. *Surg Obes Relat Dis*. 2015;11(6):1361–70.
- Campbell JA, Palmer AJ, Venn A, Sharman M, Otahal P, Neil A. A head-to-head comparison of the EQ-5D-5L and AQoL-8D multi-attribute utility instruments in patients who have previously undergone bariatric surgery. *Patient*. 2016;9(4):311–22.
- Borisenko O, Adam D, Funch-Jensen P, Ahmed A, Zhang R, Colpan Z, et al. Bariatric surgery can lead to net cost savings to health care systems: results from a comprehensive European decision analytic model. *Obes Surg*. 2015;25(9):1559–68.
- Campbell JA, Hensher M, Neil A, Venn A, Wilkinson S, Palmer AJ. An exploratory study of long-term publicly waitlisted bariatric surgery patients' quality of life before and 1 year after bariatric surgery, and considerations for healthcare planners. *PharmacoEcon Open*. 2017;. doi:10.1007/s41669-017-0038-z.
- Richardson J, Iezzi A, Khan MA. Why do multi-attribute utility instruments produce different utilities: the relative importance of the descriptive systems, scale and 'micro-utility' effects. *Qual Life Res*. 2015;24:1–9.
- Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20(10):1727–36.
- Richardson J, Iezzi A, Khan MA, Maxwell A. Validity and reliability of the Assessment of Quality of Life (AQoL)-8D multi-attribute utility instrument. *Patient*. 2014;7(1):85–96.
- Richardson J, Sinha K, Iezzi A, Khan M. Modelling utility weights for the Assessment of Quality of Life (AQoL)-8D. *Qual Life Res*. 2014;23(8):2395–404.
- Brazier J, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Econ*. 2004;13(9):873–84.
- Cunillera O, Tresserras R, Rajmil L, Vilagut G, Brugulat P, Herdman M, et al. Discriminative capacity of the EQ-5D, SF-6D, and SF-12 as measures of health status in population health survey. *Qual Life Res*. 2010;19(6):853–64.
- Chen G, Khan MA, Iezzi A, Ratcliffe J, Richardson J. Mapping between 6 Multiattribute Utility Instruments. *Med Decis Making*. 2016;36(2):147–59.
- Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes*. 3rd ed. Oxford: Oxford University Press; 2015.
- Baxter S, Sanderson K, Venn A, Otahal P, Palmer AJ. Construct validity of SF-6D health state utility values in an employed population. *Qual Life Res*. 2014;24(4):851–70.
- Clarke PM, Hayes AJ, Glasziou PG, Scott R, Simes J, Keech AC. Using the EQ-5D index score as a predictor of outcomes in patients with type 2 diabetes. *Med Care*. 2009;47(1):61–8.
- Skinner EH, Denehy L, Warrillow S, Hawthorne G. Comparison of the measurement properties of the AQoL and SF-6D in critical illness. *Crit Care Resusc*. 2013;15(3):205.
- Warkentin LM, Majumdar SR, Johnson JA, Agborsangaya CB, Rueda-Clausen CF, Sharma AM, et al. Weight loss required by

- the severely obese to achieve clinically important differences in health-related quality of life: two-year prospective cohort study. *BMC Med.* 2014;12:175.
24. Hawthorne G, Osborne R. Population norms and meaningful differences for the Assessment of Quality of Life (AQoL) measure. *Aust N Z J Public Health.* 2005;29(2):136–42.
 25. Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Qual Life Res.* 2005;14(6):1523–32.
 26. Horsman J, Furlong W, Feeny D, Torrance G. The Health Utilities Index (HUI): concepts, measurement properties and applications. *Health Qual Life Outcomes.* 2003;1(1):54. doi:10.1186/1477-7525-1-54.
 27. Alanne S, Roine RP, Räsänen P, Vainiola T, Sintonen H. Estimating the minimum important change in the 15D scores. *Qual Life Res.* 2015;24(3):599–606.
 28. Vainiola T, Pettilä V, Roine RP, Räsänen P, Rissanen AM, Sintonen H. Comparison of two utility instruments, the EQ-5D and the 15D, in the critical care setting. *Intensive Care Med.* 2010;36(12):2090–3.
 29. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol.* 2003;56(5):395–407.
 30. Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Control Clin Trials.* 1989;10(4):407–15.
 31. Hachem A, Brennan L. Quality of life outcomes of bariatric surgery: a systematic review. *Obes Surg.* 2016;26(2):395–409.
 32. Lindekilde N, Gladstone BP, Lübeck M, Nielsen J, Clausen L, Vach W, et al. The impact of bariatric surgery on quality of life: a systematic review and meta-analysis. *Obes Rev.* 2015;16(8):639–51.
 33. Torquati A, Lutfi RE, Richards WO. Predictors of early quality-of-life improvement after laparoscopic gastric bypass surgery. *Am J Surg.* 2007;193(4):471–5.
 34. Nguyen NT, Goldman C, Rosenquist CJ, Arango A, Cole CJ, Lee SJ, et al. Laparoscopic versus open gastric bypass: a randomized study of outcomes, quality of life, and costs. *Ann Surg.* 2001;234(3):279–89.
 35. Basch E, Spertus J, Dudley RA, Wu A, Chuahan C, Cohen P, et al. Methods for developing patient-reported outcome-based performance measures (PRO-PMs). *Value Health.* 2015;18(4):493–504.
 36. Coulman KD, Abdelrahman T, Owen-Smith A, Andrews RC, Welbourn R, Blazeby JM. Patient-reported outcomes in bariatric surgery: a systematic review of standards of reporting. *Obes Rev.* 2013;14(9):707–20.
 37. Richardson J, Khan MA, Iezzi A, Maxwell A. Comparing and explaining differences in the magnitude, content, and sensitivity of utilities predicted by the EQ-5D, SF-6D, HUI 3, 15D, QWB, and AQoL-8D multiattribute utility instruments. *Med Decis Making.* 2015;35(3):276–91.
 38. National Partnership Agreement on Improving Health Services in Tasmania. Australian Council on Federal Financial Relations. 2014.
 39. Thornton CM, Rozen WM, So D, Kaplan ED, Wilkinson S. Reducing band slippage in laparoscopic adjustable gastric banding: the mesh plication pars flaccida technique. *Obes Surg.* 2009;19(12):1702–6.
 40. Campbell J, Venn A, Neil A, Hensher M, Sharman M, Palmer A. Diverse approaches to the health economic evaluation of bariatric surgery: a comprehensive systematic review. *Obes Rev.* 2016;17(9):850–94.
 41. Maxwell A, Ozmen M, Iezzi A, Richardson J. Deriving population norms for the AQoL-6D and AQoL-8D multi-attribute utility instruments from web-based data. *Qual Life Res.* 2016;25(12):3209–19.
 42. Kakarla VR, Nandipati K, Lalla M, Castro A, Merola S. Are laparoscopic bariatric procedures safe in superobese (BMI \geq 50 kg/m²) patients? An NSQIP data analysis. *Surg Obes Relat Dis.* 2011;7(4):452–8.
 43. Devlin N, Shah K, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: an EQ-5D-5L value set for England. *OHE Research Paper 16/01.*
 44. Feng Y, Devlin N, Shah K, Mulhern B, Van Hout B. New methods for modelling EQ-5D-5L value sets: an application to English data. *OHE Research Paper 16/02.*
 45. Janssen B, Szende A. Population norms for the EQ-5D. Self-reported population health: an international perspective based on EQ-5D. Netherlands: Springer; 2014. p. 19–30.
 46. Tshiplova K, Pullenayegum E, Cooke T, Xie F. EQ-5D-derived health utilities and minimally important differences for chronic health conditions: 2011 Commonwealth Fund Survey of Sicker Adults in Canada. *Qual Life Res.* 2016;25(12):3009–16.
 47. Luo N, Johnson JA, Coons SJ. Using instrument-defined health state transitions to estimate minimally important differences for four preference-based health-related quality of life instruments. *Med Care.* 2010;48(4):365–71.
 48. Mitchell PM, Venkatapuram S, Richardson J, Iezzi A, Coast J. Are quality-adjusted life years a good proxy measure of individual capabilities? *Pharmacoeconomics.* 2017;35(6):637–46. doi:10.1007/s40273-017-0495-3.
 49. Snyder CF, Aaronson NK, Choucair AK, Elliott TE, Greenhalgh J, Halyard MY, et al. Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. *Qual Life Res.* 2012;21(8):1305–14.