MONASH University
Business and Economics

Centre for Health Economics

Working Paper 04-14

Subjective Wellbeing versus Utility

Incommensurable or mismeasured constructs

Professor Jeff Richardson

Foundation Director, Centre for Health Economics
Monash University

Dr Gang Chen

Research Fellow, Flinders Health Economics Group
Flinders University

Dr Munir A Khan

Adjunct Research Fellow, Centre for Health Economics
Monash University

Angelo Iezzi

Research Fellow, Centre for Health Economics
Monash University

October 2014

Centre for Health Economics
ISSN 2204-0218
ISBN 1 921187 90 5

Subjective wellbeing versus utility: incommensurable or mismeasured constructs

Jeff Richardson*[1], Gang Chen*, Munir A Khan*, Angelo Iezzi*

*Centre for Health Economics, Monash University

30 September 2014

Synopsis

It has been suggested that in the economic evaluation of services which improve the quality of life (QoL) the role of utility should be replaced, or at least supplemented by the use of the subjective wellbeing (SWB) of health states: that priority should be assigned according to the improvement in SWB rather than the increase in utility. While the change may be advocated on purely normative grounds (SWB *should* be the goal of health policy), the case is strengthened if it can be shown that measured utility cannot take account of SWB. As it may be assumed that people prefer higher to lower SWB it would be anomalous to base health policy upon procedures which could not take account of health related changes in SWB.

Criticism of utility has arisen from what is described here as the 'incommensurability hypothesis': that decision utility, as measured before the event, cannot accurately take account of SWB experienced after the event.

The theme of the present paper is that the alleged failure of utility may not be a consequence of a fundamental incommensurability of measurement before and after the event but be attributable to the description of the health states which are evaluated before the event. In particular, SWB is sensitive to psycho-social components of a health state, but these components are poorly described in the main multi attribute utility (MAU) instruments used by economists. Descriptive systems for these instruments reflect the historical, linguistic convention of equating 'health' with physical problems. But these components are comparatively unimportant determinants of SWB.

[1] Corresponding author: Jeff Richardson; jeffrey.richardson@monash.edu; +61 399050754; Centre for Health Economics, Building 75, Monash University, Clayton 3800, Victoria, Australia.

The present paper employs data from a large multi instrument comparison survey which included three measures of SWB and six MAU to investigate three questions. These are:

1. The correlation between the two sets of instruments – MAU and SWB;
2. The determinants of the variation in utility and SWB using dimension scores from the SF-36; and
3. The extent to which variation in SWB associated with seven groups of chronically ill patients may be explained by the MAUI.

The prior expectation was that (i) the correlation between SWB and utility would rise with the psycho-social content of the MAUI; (ii) that the psycho-social dimensions of the SF-36 would be primary determinants of SWB and the physical dimensions of the SF-36 the primary determinants of variation in utility predicted by the major MAUI; and (iii) that the capacity of MAUI to explain illness related loss of SWB would rise with the instrument's psycho-social content.

Results are consistent with these expectations. The QWB and EQ-5D are least able to explain SWB. The AQoL-8D fully explains SWB supporting the view that the 'incommensurability hypothesis' is wrong and that utility before the event may take into account SWB measured after the event.

Subjective wellbeing versus utility: incommensurable or mismeasured construct

Introduction

A number of economists have suggested that the improvement in the quality of life (QoL) attributable to a health service should be assessed in economic evaluation studies using subjective wellbeing (SWB) rather than utility (Cummins, 1998; Dolan, 2008a, b, 2011; Dolan & Kahneman, 2008; Dolan & Metcalfe, 2012; Kahneman & Krueger, 2006; Kahneman et al., 1997; Layard et al., 2012). The argument is based upon the assumption that people prefer higher to lower SWB and that utility, as commonly measured, does not satisfactorily reflect SWB.

Criticism of utility has focused upon the ability of individuals to imagine hypothetical states and to assess them as if they were truly experiencing them. Kahneman et al. (1997), for example, argue that 'systematic errors in the evaluation of past events and decisions that do not maximise future experience utility can be observed in decision makers whose cognitive functions are normal' (p 376) and Smith et al. note that 'an entire subfield within psychology deals with the errors and biases that occur when people attempt to forecast what their quality of life would be like in different circumstances' (Smith et al., 2008 p 86). A particular problem arises because, as Dolan and Kahneman (2008) note, 'people adapt to changes in their circumstances but they often fail to appreciate the degree to which they will adapt to these changes' (p230). In support of the argument the authors cite a review by De Wit et al. (2000) which found that 23 studies reported health states to be rated more highly by patients than by the public; two studies where the reverse was true and eleven where no difference occurred.

In short, people may have a preference for higher SWB but the construct measured before a health service occurs (decision utility) and the construct measured when a person experiences the new health state (SWB) are incommensurable owing to people's cognitive limitations.

Despite these arguments neither the theoretical differences nor the empirical evidence necessarily indicates that utility is unable to take account of SWB. In the empirical studies employed by CUA the assessment of utility is based upon a health state description and the extent to which the assessment takes account of SWB necessarily reflects the description presented to raters. The failure of utility to reflect subsequent SWB may therefore be the result of the health state description provided by an MAUI rather than the incommensurability of the two constructs. In particular, in every case where it has been studied mental health has been found to be the most important determinant of individual happiness (Helliwell et al., 2013). Consequently, the strength of the association between an MAUI and SWB is likely to be associated with the extent of the description of psycho-social health in the instrument's descriptive system.

The present article employs three measures of SWB and six MAUI to examine these hypotheses. It presents two tests. The first is the correlation between measures of SWB and utility. The prior expectation is that the correlation will rise with the psycho-social content of the instrument's descriptive system which is independently measured. The second test is the extent to which variation in the measures of SWB can be explained by the MAUI.

While the analysis is limited to utility measured by these MAU instruments the hypothesis is more general. Utility may be assessed directly by the evaluation of a specific health state description in conjunction with a utility scaling instrument such as the standard gamble or time trade-off. While it is not tested here, the study hypothesis is that in these cases, also, a failure of utility to account for SWB may be a result of the description used.

The section below describes the methods, instruments and data used to investigate the two study questions, viz, (i) the correlation between three measures of SWB and six MAUI, and (ii) the extent to which variation in the measures of SWB can be explained by the MAUI. Results are presented and discussed in the following two sections.

Data and Methods

Methods: Three tests were conducted using data from a large multi instrument survey of the healthy public and patients in seven chronic disease areas. The survey included six MAU and three SWB instruments. The first test was to compare the extent to which utilities measured using the six MAU instruments respond to changes in three measures of SWB. Unadjusted beta coefficients were calculated which indicate the change in utility, measured in standard deviations, with a change in SWB of one standard deviation. Secondly, results were used to test the hypothesis that the correlation between utility and SWB reflects the psycho-social content of the utility instrument where the content was derived from the regression of MAUI results upon the physical and psycho-social dimensions of the SF-36. The third test replicated the methods used by Böckerman et al. (2011). The three measures of SWB were firstly regressed upon socio demographic variables and seven dummy variables each of which indicated that a respondent had one of the seven chronic diseases. Coefficients therefore indicated the average differences between the SWB of respondents in a disease group and respondents from the healthy public. Secondly, utility as measured by one of the six MAU instruments was included in the regression. If an MAU instrument took full account of variation in SWB then dummy variables in the second regression would be insignificant. The diminution in the magnitude of the coefficient on each dummy variable indicates the responsiveness of the MAU instrument to SWB in the corresponding patient group.

Questionnaires: The three SWB instruments are described in Table 1. The SWLS is a widely used instrument whose reliability has been repeatedly established (Diener et al., 1985; Dolan et al., 2006). Similarly, the PWI is an established instrument which has been subject to extensive testing (Cummins, 1998, 2014). In contrast, the instrument labelled here the 'ONS' is new and has no official name. It was introduced into the UK 'Integrated Household Survey' of the Office for National Statistics (ONS) in 2011 and the first results published in December 2011 and February 2012 (Hicks, 2012). Despite seeking to measure the same concept, the form and content of the instruments differ. The SWLS is focused exclusively upon satisfaction with present and past life. Statements primarily alter the way in which the two questions are asked. The PWI is similarly focused upon satisfaction. Unlike the SWLS it nominates the major life domains and does not refer to past life. The ONS has only one satisfaction question. It seeks to incorporate all dimensions of SWB by including questions to measure eudemonia and both positive and negative affect.

The six MAU instruments are described in Table 2 and reviewed in Richardson et al. (2014a). They differ significantly in size and content. The smallest instrument – the EQ-5D – has four of its five items in the physical domain. In contrast, five of the eight dimensions and 25 of the 35 items of the largest instrument, the AQoL-8D, relate to psycho-social health.

Data: The Multi Instrument Comparison (MIC) survey was carried out in six countries: Australia, Canada, Germany, Norway, UK and USA. It was administered online by a global panel company, CINT Pty Ltd. The survey was approved by the Monash University Human Research Ethics Committee, Monash University Melbourne Australia (Ref No: CF11/3192-2011001748). Respondents were administered the three SWB instruments before subsequent questioning to preclude the contamination of answers by contemplation of other questions. Subsequently, respondents were asked to indicate if they had been diagnosed as currently having one of the chronic illnesses in the study and also to rate their overall health on a VAS numerical scale where 0.00 represented death and 100 'best possible health' (physical, mental and social). Quotas were then used to obtain a demographically representative sample of the healthy public, defined by the absence of a chronic disease and a VAS score above 70. Quotas were also applied to obtain a target number of respondents in each of the seven disease areas in the study, viz, arthritis, asthmas, cancer, depression, diabetes, hearing loss and heart disease. Respondents with a chronic disease were asked to complete a disease specific questionnaire to determine the severity of the condition. For five of the MAU instruments utilities were calculated using algorithms provided by the instruments' authors: SF-6D (Brazier et al., 2002), HUI 3 (Feeny et al., 2002), 15D (Sintonen & Pekurinen, 1993), QWB (Kaplan et al., 1976) and AQoL-8D (Richardson et al., 2014b) . The five level EQ-5D-5L utilities were obtained from the crosswalks published by the EuroQoL Group (Rabin et al., 2011).

Results

Data were obtained from 9,665 individuals. Edit procedures, based primarily upon responses to repeated questions, resulted in the removal of 17 percent of the total. Table 3 presents the age-gender and educational status of the remaining 8,022 respondents. Because quotas were imposed the proportion of respondents from each country is similar (Australia 17.8 percent, USA 18.2 percent, UK 16.9 percent, Canada 16.6 percent, Norway 14.7 percent, and Germany 15.8 percent). For the same reason, the age, gender and educational profile of respondents within each country is similar. The numbers recruited from the disease areas varied from 772 for cancer to 943 for heart disease. The 1,760 'public' respondents were obtained from country samples which closely matched the age-gender profile in each country. Except in Norway and Germany where the QWB was not administered (reducing the response for the QWB to 5,576) each of the 8,022 respondents completed the six MAU instruments along with socio-demographic questions. There were few missing data as the online program did not permit respondents to proceed until questions were completed. (Only 14 individuals did not complete the final question.) Details of the sample administration and editing in each country are provided in country specific reports (Richardson et al., 2012a-f).

Table 4 reports summary statistics for the six MAU and three SWB instruments. With the exception of the QWB and 15D, mean values for the MAU are similar, varying from 0.83 to 0.88 in the public sample and from 0.68 to 0.74 in the full sample. Other characteristics of the sample differ more significantly. In the full sample the standard deviation of the observations varies by 100 percent from 0.27 for HUI 3 to 0.13 for 15D and 0.14 for SF-6D. Ceiling effects (U = 1.00) vary from 19.1 percent (EQ-5D) to 0.3 percent (AQoL-8D) and the percentage with a utility below 0.4 varies from 0.3 for the 15D and 1.3 percent for the SF-6D to 13.9 percent for HUI 3 and 14.7 percent for AQoL-8D.

Frequency distributions for the instruments are shown in Supplementary Figure 1(a-h).

Test 1: Results from the first test are reported in Table 5. Consistent with the differences in descriptive systems there is significant difference in the magnitude of the beta coefficients. For three of the MAU instruments – SF-6D, HUI 3 and 15D – the beta are similar and vary from 0.49-0.56. Coefficients for the EQ-5D and QWB are, on average 0.07 and 0.12 points lower and coefficients for the AQoL-8D are 0.16 higher than the average for the three. Table 5 also reports the beta coefficients for the linear relationship between the three measures of SWB. Consistent with the differences in their content, the relationship between them is imperfect with the strength of the association similar to the strength of the association between the MAU instruments reported in the literature (Richardson et al., 2014a). In Supplementary Table S.1 it is shown that the relationship between the three instruments is non-linear implying that the instruments have different interval properties.

Test 2: Beta coefficients from the regression of both utilities and SWB upon the dimensions of the SF-36 are reported in Table 6 and the physical and psycho-social content of each instrument in the final two rows. Content was defined as the sum of the relevant (physical or psycho-social) beta coefficients divided by the sum of all beta coefficients. The latter represents the increase in the dependent variable, measured in standard deviations, when each independent variable is increased by one standard deviation. The physical and psycho-social content are therefore defined by the percentage of this increase attributable to the physical and psycho-social variables respectively.

Results from Tables 5 and 6 are combined in Figure 1 which plots the average correlation between MAUI and SWB against the psycho-social content of the MAUI.

Test 3: Results from the third test are given in Table 7. Each column reports results from three regression equations with SWLS, ONS and PWI as dependent variables. Full regression results are reported for the PWI. For the SWLS and ONS only the coefficients on the utility and disease dummy variables are shown. Equations in column 1 omit utility from the list of independent variables. Subsequent columns introduce utility as measured by the MAU instrument shown at the top of the column.

The three measures of SWB give consistent results with respect to the effect of the seven diseases. From the first column, depression has a significantly greater effect upon SWB than any other disease. The reduction in the index of SWB is between 0.2 and 0.25 which is more than double the next highest effect which is associated with diabetes and cancer. The smallest effect is consistently for patients with hearing problems followed by patients with asthma. The average loss of SWB associated with these diseases is about 20 and 25 percent of the loss associated with depression respectively. The effects of the seven diseases upon the PWI are shown in Figure 2.

Inclusion of the different estimates of utility in the three equations has a broadly similar effect. The magnitude of the coefficient upon the dummy variables is reduced but with several exceptions remains negative. Exceptions occur with respect to hearing loss and arthritis where the coefficients generally become insignificant. The other exceptions are the equations which include AQoL-8D. These eliminate the negative coefficient on variables in every disease category and with every measure with the single exception of depression, where there remains a marginally significant but small negative coefficient of 0.014 in the ONS equation.

A simple index of the extent to which the different MAU instruments take account of SWB may be obtained by summing the number of results where the negative sign on a disease dummy variable is eliminated or becomes statistically insignificant. The result of this exercise, in order of success, is AQoL-8D (20/21); 15D (14/21); SF-6D (12/21); QWB (10/21); HUI 3 (7/21); EQ-5D (3/21). While the inclusion of utility did not fully account for SWB in the majority of cases it did result in a diminution of the b coefficient indicating a partial response to SWB. The magnitude of this effect for the PWI is shown in Figure 3.

Using the same criteria as above – the elimination of a significant negative coefficient – the diseases where the MAU instruments were least able to account for the loss of SWB were depression and diabetes where only AQoL-8D fully accounted for SWB. These diseases were followed by cancer and heart diseases (8 coefficients eliminated in at least one regression by the SF-6D, 15D, QWB and AQoL-8D); asthma (14; all instruments except EQ-5D) and hearing loss (15; all instruments except EQ-5D).

With the inclusion of utility there are a number of cases in Table 6 where the coefficient on the dummy variable switches from negative to significantly positive. This implies that utility over-predicts the loss of SWB in the disease group and that the positive coefficient on the dummy variable is needed to offset this. Switching primarily occurs with the inclusion of AQoL-8D in SWLS and ONS regressions. However switching also occurs for arthritis in ONS regressions with every MAU instrument; and it occurs for hearing loss in SWLS regressions when utility is measured by the HUI 3 or QWB. Over prediction is discussed further in the following section.

Discussion

The case for replacing or supplementing utility with SWB in the assessment of health services has drawn upon the argument that decision utility and experienced utility – SWB – are largely incommensurable: that the perspectives and experiences of individuals prior to entering a health state do not permit them to properly take account of the SWB that they would experience in that state. In particular, it has been argued that they would fail to fully anticipate hedonic adaptation: their capacity to adjust to the new circumstances. The present paper has sought to test a contrary view, namely, that these problems are largely attributable to the failure of the health state description to provide information on attributes directly associated with SWB and in particular information relating to psycho-social health.

The paper presented three tests which related to these alternative hypotheses. The first test was the correlation between the two sets of measures. Results confirm that the relationship between MAUI and SWB varies with the composition of the instrument's descriptive system and, in particular, with the psycho-social content of the descriptive system. The SF-6D and HUI 3, which have 3 of 6 and 2 of 8 of their items in the psycho-social domain result in larger beta coefficients than the EQ-5D with only one of its five items in the psycho-social domain. The 15D has similar beta coefficients to the SF-6D and HUI 3, and three of the 15D items are dedicated to depression. Beta coefficients for AQoL-8D are significantly larger than for other instruments and seven of its items are dedicated to mental health and four to happiness, the items most closely related to SWB.

The third test investigated the relationship more systematically. Using the dimension scores from the SF-36 as explanatory variables it was found that the most widely used MAUI, the EQ-5D, has the least psycho-social content. Its beta coefficients, reported in Table 6, indicate that two dimensions – pain and physical function – account for 64 percent of the instrument's content. In striking contrast, neither of these dimensions affect any of the SWB instrument scores after mental health and vitality have been included in the equation. The correlation between SWB and the six MAUI rises with psycho-social content.

These results are consistent with the psychological literature. As noted earlier, in every case in which it has been studied mental health has been found to be the most important determinant of SWB (Helliwell et al., 2013).

The third test was the extent to which the reduction in SWB associated with seven chronic conditions could be explained by utility as measured by each of the MAUI. Results vary with the measure of SWB, reflecting the non-linear relationship between the three instruments. Nevertheless they are broadly consistent and suggest that utility may fully explain the reduction in SWB. There are, however, several important caveats. First, the result is a necessary but not sufficient condition for the conclusion. It is possible that utility only partially but systematically takes account of SWB and that the coefficient upon utility inflates the partial effect to fully account for SWB in the regression equations. Secondly, and as discussed further below, the ability of utility to account for SWB does not negate the normative argument for the use of SWB as a measure of outcome.

Third, data for the study were obtained from an online survey. Despite the editing described in Section 2 respondents are not representative of the general community. However there are no *a priori* reasons for believing that their self-selection would systematically change the relationship between utility and SWB. The inclusion of patients from seven disease areas resulted in a very wide cross section of health states and the diversity of respondent experience was more important for this study than population representativeness.

The most important caveat relates to the ability of instruments to account for adaptation. *Prima facie*, it may appear implausible that an instrument with a fixed utility formula could account for variation in a person's preferences because of adaptation. However, in principle, this may be achieved if the descriptive system includes dimensions which reflect the effect of adaptation. As noted above SWB is primarily associated with mental health and the inclusion of appropriate psycho-social dimensions has the potential to accommodate the effects of hedonic adaptation; that is, pre and post-adaptive health states may be measured with the same instrument as adaptation will alter the responses to the psycho-social items. Whether or not this is achieved is therefore an empirical question.

Contrary to the concern in the literature, the empirical evidence indicates that for the most widely used utility instruments, the problem is not the inability of the instrument to take account of improved SWB when hedonic adaptation occurs. Rather, it is the inability of these instruments to fully explain the reduction in SWB associated with chronic illness. That is, irrespective of hedonic adaptation, the instruments generally take too little, rather than too much, account of SWB. As noted, exceptions to this generalisation occur in the case of hearing loss and arthritis and when the AQoL-8D is included in equations. A likely explanation for the results for hearing loss and arthritis is that the elements affecting SWB are adequately represented in the descriptive systems of the MAUI. Arthritis is dominated by pain and physical function and each of the instruments includes items describing these. The only MAUI which does not fully account for hearing loss is the EQ-5D-5L which contains no item directly associated with hearing loss.

Interpretation of the results is complicated by the over-prediction of the loss of SWB associated with arthritis in all of the equations using the ONS and also when utility is measured by AQoL-8D. The present evidence does not permit a fully satisfactory explanation of these anomalies. It is possible that during the interview used to obtain utility weights the importance of pain and hearing loss were exaggerated by respondents. Consequently, utilities would be depressed. If so, the over-prediction of the loss of SWB would depend upon the strength of the correlation between utility and SWB. In the case of arthritis this is consistent with the results in Table 6 when SWB is measured by the ONS. The rank order of the over-prediction closely follows the rank order of the correlation between the ONS and utilities in Table 5. A similarly exaggerated response might explain the over-prediction of the effect of hearing loss by HUI 3 for the SWLS and AQoL-8D for each of the three measures of SWB.

This tentative explanation supports the argument that – at least in these cases – utilities obtained from the public do not fully reflect the experience of patients. This may be attributable to a degree of adaptation in these disease areas or to the failure of imagination by respondents to a utility scaling survey. While responses to psycho-social variables may accurately reflect the post-adaptive health state and are weighted appropriately the pain (arthritis) and hearing variables may retain their pre-adaptive utility weights which exaggerates their importance. If subsequent research confirmed this explanation one solution would be to adjust the weights to reflect the importance of the variables after adaptation. This is illustrated in Supplementary Note 1. Importantly, however, over-prediction occurs for a minority of results and, with the exception of the AQoL-8D result for arthritis, the effects are relatively small.

In a rare analysis of how well utility captures the negative effects of chronic disease upon SWB Böckerman et al. (2011) conclude that the utility instruments they employ – the EQ-5D and 15D – fail to capture the effects of some chronic conditions upon SWB and, in particular, that using utility as the basis for resource allocation is likely to result in the underfunding of psychiatric problems where utility appeared least able to account for variation in SWB. The present results suggest that this conclusion was a reflection of the choice of utility instruments and not a necessary result arising from the concept of utility. The ability of utility instruments to account for SWB varies with the quality of the description of mental health. With the larger description contained in the AQoL-8D, utility fully accounted for variation in SWB.

The importance of the composition of an instrument's descriptive system also varied with the disease. The primary effect of arthritis is upon pain and physical activity which are well described by the EQ-5D. In the corresponding equations, arthritis is the single disease where the EQ-5D accounts for variation in SWB. Similarly, two of the eight HUI 3 items are dedicated to senses and HUI 3 fully accounts for variation in SWB associated with hearing.

As noted earlier, the ability to account for SWB does not imply that utility is the appropriate criterion for evaluating health programs. Beta coefficients in Table 5 are less than 1.00 implying that a change in utility is associated with a smaller shift in SWB (relative to the standard deviation of each measure). This suggests that while people take account of SWB it is not the only influence upon utility. A normative argument might be made for disregarding these other influences, especially if it could be demonstrated that these influences reflected poor judgement. However there is no agreement in the literature that SWB is the appropriate social objective. There is also a normative argument for the use of utility to preserve autonomy and, as elsewhere, to require people to bear the consequences of their own decisions.

Notwithstanding the normative arguments, results here indicate that the transition to SWB as a unit of outcome might be premature. The three measures of SWB employed in this study differ in content and their association is imperfect. The non-linear relationship reported between them (in supplementary material) implies a varying interval property for the units of the three measures. This is indicative of the unsettled question in the literature about what constitutes 'SWB' which parallels the question in the economics literature concerning the appropriate content and properties of MAUI.

Conclusion

There have been relatively few empirical studies of the relationship between utility and SWB. The multi instrument survey data employed here provide a unique opportunity for their comparison. They confirm the low correlation between the two concepts as measured by any of the instruments in the survey. However the results do not suggest that utility and SWB are incommensurable concepts. Variation in SWB between health states may be largely explained by utility. While the results here apply specifically to utilities measured by MAU instruments, a single example of an instrument's capacity to account for SWB suggests that the result is more general. Consistent with the literature, the strength of the association and the capacity of MAU instruments to account for variation in SWB varies significantly with the detail relating to mental health that is in the instrument's descriptive system. With its uniquely limited coverage of mental health, the EQ-5D is least able to account for variation in SWB and its widespread use implies the systematic disadvantaging of therapies which primarily increase SWB and, in particular, psychiatric services.

While there has been concern in the literature with respect to the inability of utility to take account of adaptation, results here suggest that this concern has largely arisen from the use of MAUI which do not contain items which can reflect adaptation. While this remains an area where further research is needed, supplementary material to the paper suggests one approach which has the potential to accommodate adaptation when this is known to occur.

Finally, while similar conclusions were reached with each of the three measures of SWB used in the study, results varied with the choice of SWB instrument. The non-linear relationship between them indicates that – as with the measurement of utility – there is no simple and unambiguous measure of SWB and further research in this area is needed to establish its role in economic evaluation.

References

Böckerman, P., Johansson, E., & Saarni, S.I. (2011). Do established health-related quality-of-life measures adequately capture the impact of chronic conditions on subjective well-being? *Health Policy*, 100, 91-95.

Brazier, J., Roberts, J., & Deverill, M. (2002). The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics*, 21, 271-292.

Cummins, R.A. (1998). The second approximation to an international standard of life satisfaction. *Social Indicators Research*, 43, 307-334.

Cummins, R.A. (2014). The Personal Wellbeing Index in Australia. In A. Michalos (Ed.), *Encyclopedia of Quality of Life and Well-Being Research*. New York: Springer.

De Wit, A.A., Busschbach, J.J., & De Charro, F.T. (2000). Sensitivity and perspective in the valuation of health status: Whose values count? *Health Economics*, 9, 109-126.

Diener, E., Emmons, R.A., Larsen, R.J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment*, 49, 71-75.

Dolan, P. (2008a). Developing methods that really do value the "Q" in the QALY. *Health Economics, Policy and Law*, 3, 69-77.

Dolan, P. (2008b). In defence of subjective well-being. *Health Economics, Policy and Law*, 3, 93-95.

Dolan, P. (2011). *Using Happiness to Value Health*. London: Office of Health Economics, http://www.ohe.org/object/display.cfm?serv=2&id=98 [Accessed 24 January 2014].

Dolan, P., & Kahneman, D. (2008). Interpretations of utility and their implications for the valuation of health. *Economic Journal*, 118, 215-234.

Dolan, P., & Metcalfe, R. (2012). Valuing health: A brief report on subjective well-being versus preferences. *Medical Decision Making*, Jul/Aug, 578-582.

Dolan, P., Peasgood, T., & White, M. (2006). *Review of research on the influences on personal well-being and application to policy making*. London: http://tna.europarchive.org/20081112122150/http://defra.gov.uk/sustainable/government/publications/pdf/WellbeingProject2.pdf [Accessed 24 January 2014].

Feeny, D., Furlong, W., Torrance, G., Goldsmith, C., Zhu, Z., DePauw, S., et al. (2002). Multi attribute and single attribute utility functions for the Health Utilities Index Mark 3 System. *Medical Care*, 40, 113-128.

Helliwell, J., Layard, R., & Sachs, J. (2013). *World Happiness Report 2013*. New York: http://issuu.com/earthinstitute/docs/worldhappinessreport2013_online [Accessed 23 January 2013].

Hicks, S. (2012). Case Study: ONS Measuring Subjective Wellbeing: The UK Office for National Statistics Experience. In J. Helliwell, R. Layard, & J. Sachs (Eds.), *World*

Happiness Report 2012. New York: The Earth Institute, Columbia University http://www.earthinstitute.columbia.edu/articles/view/2960 [Accessed 30 April 2012].

Kahneman, D., & Krueger, A.B. (2006). Developments in the measurement of subjective well-being. *Journal of Economic Perspectives,* 20, 3-24.

Kahneman, D., Wakker, P., & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *Quarterly Journal of Economics,* 112, 375-405.

Kaplan, R., Bush, J., & Berry, C. (1976). Health status: Types of validity and the index of wellbeing. *Health Services Research,* 11, 478-507.

Layard, R., Clark, A., & Senik, C. (2012). The causes of happiness and misery. In J. Helliwell, R. Layard, & J. Sachs (Eds.), World Happiness Report 2012. New York: The Earth Institute, Columbia University http://www.earthinstitute.columbia.edu/articles/view/2960 [Accessed 30 April 2012].

Rabin, R., Oemar, M., Oppe, M., Janssen, B., & Herdman, M. (2011). *EQ-5D-5L User Guide: Basic information on how to use the EQ-5D-5L instrument*. Rotterdam: EuroQoL Group, http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/UserGuide_EQ-5D-5L.pdf. [Accessed 30 April 2012].

Richardson, J., Khan, M.A., Iezzi, A., & Maxwell, A. (2012a-f). *Cross-national comparison of twelve quality of life instruments, Research Papers 78, 80-83, 85. MIC Report: 2: Australia; 3: UK; 4: USA; 5: Canada; 6: Norway; 7: Germany*. Melbourne: Centre for Health Economics, Monash University.

Richardson, J., McKie, J., & Bariola, E. (2014a). Multi attribute utility instruments and their use. In A.J. Culyer (Ed.), *Encyclopedia of Health Economics* pp. 341-357). San Diego: Elsevier Science.

Richardson, J., Sinha, K., Iezzi, A., & Khan, M.A. (2014b). Modelling utility weights for the Assessment of Quality of Life (AQoL) 8D. *Quality of Life Research,* DOI: 10.1007/s11136-014-0686-8.

Sintonen, H., & Pekurinen, M. (1993). A fifteen-dimensional measure of health related quality of life (15D) and its applications. In S. Walker, & R. Rosser (Eds.), *Quality of Life Assessment*. Dordrecht: Kluwer Academic Publishers.

Smith, D.M., Brown, S.L., & Ubel, P.A. (2008). Are subjective well-being measures any better than decision utility measures? *Health Economics, Policy and Law,* 3, 85-91.

Table 1 Three SWB instruments[1]

Instrument	Description
SWLS: Satisfaction with Life Scale	5 satisfaction questions, 7 response categories: strongly dis agree to strongly agree ♦ life … close to ideal ♦ have the important things ♦ life … conditions excellent ♦ would not change past life ♦ satisfied with life
PWI: Personal Well-being Index	8 satisfaction questions, 11 response categories: 0 … 10 Common stem: how satisfied are you with… Domains ♦ standard of living ♦ safety ♦ health ♦ participation in the community ♦ achievement ♦ future security ♦ personal relations ♦ spirituality
ONS: Office of National Statistics	4 questions, 11 response categories: not at all … completely Question/type ♦ satisfaction with life … satisfaction ♦ do worthwhile things … eudemonia ♦ happy yesterday … positive effect ♦ anxious yesterday … negative effect

(1) In each case unweighted scores, S, are obtained by summing item responses, r, and rescaling to a (0.00-1.00) scale using the formula $S=(r-r_{min})/(r_{max}-r_{min})$.

Table 2 Comparison of the dimensions and content of six MAU instruments

	Dimension	EQ-5D-5L	SF-6D	HUI 3	15D	AQoL-8D	QWB[(1)]
Physical	Physical Ability/Mobility/ Vitality/Coping/Control	●	●	●●	●●	●●	●●
	Bodily Function/Self Care	●			●●●	●	
	Pain/Discomfort	●	●	●	●	●●	
	Senses			●●	●●	●●	
	Usual Activities/Work	●	●		●	●●	
	Communication			●	●	●	
Psycho-social	Sleeping				●	●●	
	Depression/Anxiety/Anger	●	●	●	●●●	●●●●●●●	
	General Satisfaction					●●●●	
	Self-esteem					●●●●●	
	Cognition/Memory Ability			●			
	Social Function/Relationships		●			●●●●●	●
	(Family) Role		●			●	
	Intimacy/Sexual Relationships				●	●	
	Total items/symptoms	5	6	8	15	35	
	Health states described	3125	18,000	972000	$3.1x10^{10}$	$2.4x10^{23}$	

(1) QWB has 3 items relating to mobility, physical and social health plus 27 symptom groups which include, inter alia, 'spells of feeling upset, being depressed or crying', 'trouble sleeping' and 'excessive worry or anxiety'.

Table 3 Respondent characteristics

Country	Excluded (%)		Composition of Final Sample																	
			Public (%)							Patient (%)							Education			Total (n)
	Pub	Pat	18-24	25-34	35-44	45-54	55-64	65+	Male	18-24	25-34	35-44	45-54	55-64	65+	Male	High school	Diploma or certificate or trade	University	
Australia	36.5	15.3	11.3	18.1	18.9	18.5	14.7	18.5	46.4	2.1	8.0	10.3	19.5	32.6	27.5	50.4	35.8	35.1	29.1	1430
USA	17.1	11.2	10.3	17.8	18.1	20.2	16.2	17.4	45.2	4.8	8.8	13.1	25.0	25.5	22.8	36.4	36.1	29.3	34.6	1460
UK	18.8	13.2	11.4	15.4	20.1	18.1	14.4	20.5	47.7	7.1	12.7	9.7	16.4	29.0	25.1	51.4	38.1	30.2	31.7	1356
Canada	9.4	19.2	12.8	18.3	16.2	20.1	16.8	15.9	47.3	5.8	15.1	18.0	19.1	27.3	14.8	34.8	29.2	47.6	23.2	1330
Norway	19.1	19.1	12.8	16.0	16.7	18.4	15.6	20.5	50.3	6.2	8.2	10.2	16.8	26.0	32.6	63.6	28.0	48.5	23.5	1177
Germany	24.4	17	6.5	20.0	18.5	23.1	17.7	14.2	50.4	5.2	8.3	17.5	31.4	24.4	13.2	54.2	19.6	55.0	25.4	1269
Total	21.2	15.7	11.0	17.6	18.0	19.7	15.9	17.8	47.8	5.1	10.1	13.1	21.4	27.6	22.6	48.0	31.4	40.4	28.2	8022

Table 4 Summary statistics for the 9 instruments(1)

	Public			Total						
	Mean	**SD**	**n**	**Mean**	**SD**	**Min**	**Range**	**percent U=1.00**	**percent U<0.4**	**n**
	A. MAUI									
EQ-5D-5L	0.88	0.13	1760	0.74	0.23	-0.51	1.51	19.10	8.90	8022
SF-6D	0.80	0.11	1760	0.71	0.14	0.30	0.70	1.30	1.30	8021
HUI 3	0.88	0.14	1760	0.71	0.27	-0.34	1.34	7.10	13.90	8021
15D	0.94	0.06	1760	0.85	0.13	0.25	0.75	6.90	0.30	8021
QWB*	0.74	0.14	1212*	0.63	0.15	0.15	0.85	2.40	6.50	5576*
AQoL-8D	0.83	0.14	1760	0.68	0.22	0.10	0.90	0.30	14.70	8022
	B. SWLS									
SWLS	0.65	0.21	1760	0.56	0.25	0.00	1.00	1.87	25.39	8008
PWI	0.72	0.16	1760	0.65	0.19	0.00	1.00	1.02	10.68	8008
ONS	0.72	0.18	1760	0.64	0.22	0.00	1.00	1.91	13.19	8008

(1) QWB was not administered in Norway or Germany

Table 5 Beta coefficients

	EQ-5D	SF-6D	HUI 3	15D	QWB	AQoL-8D	PWI	ONS
SWLS	0.43	0.49	0.50	0.49	0.39	0.66	0.79	0.81
PWI	0.45	0.52	0.52	0.52	0.40	0.67		0.78
ONS	0.46	0.56	0.52	0.52	0.40	0.70		
Average	0.45	0.52	0.51	0.51	0.40	0.68		

Table 6 Content of MAUI: Beta coefficients[1]

Independent	SF36 Regressions*: Dependent = MAU_i						Dependent = SWB		
SF-36 Dimension	EQ-5D	SF-6D[2]	HUI3	15D	QWB	AQoL-8D	SWLS	PWI	ONS
Physical									
Gen H	0.07	0.01*	0.09	0.19	0.13	0.14	0.14	0.23	0.8
Phys	0.29	0.15	0.31	0.28	0.19	0.09	ns	ns	ns
Role P	0	0.13	0	0.01*	0.03*	0	ns	ns	ns
B Pain	0.41	0.22	0.25	0.19	0.20	0.16	ns	ns	ns
Psycho-social									
Vital	0.01*	0.13	0.04	0.18	0.20	0.25	0.10	0.08	0.08
Social	0.09	0.21	0.11	0.09	0.05	0.06	ns	0.02	0.02
Role E	0.02*	0.18	0	0.04	0.03	0.05	ns	ns	ns
MH	0.22	0.19	0.27	0.14	0.08	0.38	0.42	0.43	0.60
Constant	0.09	0.33	-0.07	0.49	0.30	0.02	0.20	0.25	0.70
R^2	0.70	0.88	0.67	0.77	0.53	0.81	0.41	0.44	0.54
F	1652	5342	1414	2283	778	2971	687	783	1909
Content[3]									
Physical	0.71	0.42	0.61	0.60	0.68	0.36			
Psy-Soc	0.24	0.58	0.39	0.40	0.32	0.64			

(1) All results are significant at the 0.01 level except those with '*'; n = 8022 (QWB: n = 5576)

(2) Comparisons with SF-36 biased as it is a subset of the dimensions

(3) Content $= \sum_{j=1}^{4} Beta / \sum_{i=1}^{8} Beta_i$ when j = physical dimensions or psycho-social dimensions

SF-36 dimensions: Gen H=general health; Phys=physical function; Role P=role limit physical; B Pain=bodily pain; Vital=vitality; Social=social functioning; Role E=role limit emotional; MH=mental health

AQoL-8D dimensions: Ind Liv=Independent living; Pain=pain; Sense=senses; Happy=happiness; Mental=mental health; Cope=coping; Relation=relationship; Worth=self-worth

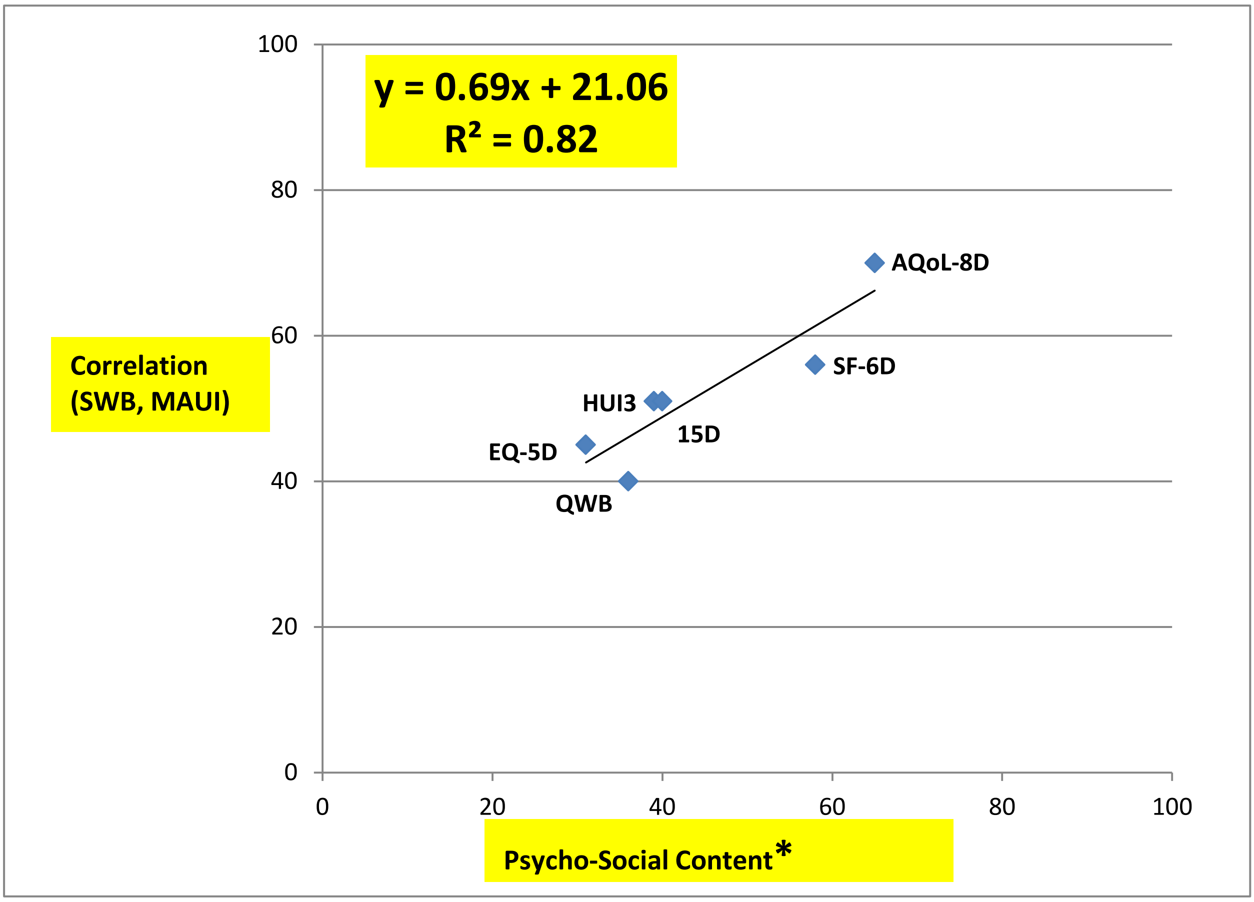
Table 7 Regression of SWB on disease dummy variables and MA utilities
'b' coefficient from regression: SWB = a $\sum_{i=1} b_i$ independent $_i$

SWB \ MAU		EQ-5D	SF-6D	HUI 3	15D	QWB	AQoL-8D
SWLS							
Coefficient on MAU	No MAU	0.394**	0.761**	0.412**	0.856**	0.557**	0.716**
Coefficient on dummy							
Depression	-0.258**	-0.145**	-0.113**	-0.115**	-0.103**	-0.127**	-0.003
Diabetes	-0.117**	-0.052**	-0.034**	-0.035**	-0.027**	-0.046**	0.009
Cancer	-0.111**	-0.044**	-0.020**	-0.030**	-0.007	-0.021	0.017*
Arthritis	-0.094**	-0.004	0.005	0.010	0.009	-0.006	0.044**
Heart disease	-0.093**	-0.031**	-0.010	-0.019*	0.003	-0.020	0.025**
Asthma	-0.072**	-0.025**	-0.005	-0.018*	0.008	-0.012	0.018*
Hearing loss	-0.043**	-0.010	0.001	0.032**	0.012	0.023*	0.044**
R^2	0.148	0.251	0.288	0.305	0.295	0.213	0.453
ONS							
Coefficient on MAU	No MAU	0.384**	0.797**	0.386**	0.804**	0.494**	0.666**
Coefficient on dummy							
Depression	-0.251**	-0.141**	-0.099**	-0.117**	-0.106**	-0.138**	-0.014*
Diabetes	-0.103**	-0.039**	-0.016*	-0.026**	-0.018*	-0.035**	0.015*
Cancer	-0.109**	-0.043**	-0.013	-0.032**	-0.010	-0.027**	0.011
Arthritis	-0.070**	0.017*	0.034**	0.027**	0.026**	0.016	0.058**
Heart disease	-0.094**	-0.033**	-0.006	-0.024**	-0.004	-0.024*	0.016*
Asthma	-0.065**	-0.019*	0.006	-0.014	0.010	-0.005	0.020**
Hearing loss	-0.058**	-0.026**	-0.012	0.012	-0.006	0.003	0.023**
R^2	0.173	0.298	0.370	0.350	0.340	0.243	0.513
PWI							
Coefficient on MAU	No MAU	0.325**	0.629**	0.330**	0.695**	0.418**	0.552**
Coefficient on dummy							
Depression	-0.201**	-0.108**	-0.081**	-0.087**	-0.076**	-0.105**	-0.005
Diabetes	-0.102**	-0.048**	-0.033**	-0.036**	-0.028**	-0.052**	-0.004
Cancer	-0.099**	-0.044**	-0.024**	-0.034**	-0.014	-0.028**	0.000
Arthritis	-0.077**	-0.002	0.005	0.007	0.007	-0.005	0.030**
Heart disease	-0.092**	-0.040**	-0.023**	-0.032**	-0.014*	-0.034**	-0.001
Asthma	-0.063**	-0.024**	-0.007	-0.019**	0.002	-0.018*	0.007
Hearing loss	-0.049**	-0.022**	-0.013	0.011	-0.004	0.001	0.018**
Male	-0.012**	-0.016**	-0.022**	-0.011**	-0.017**	-0.022**	-0.025**
Age 45-54	-0.008	0.008	-0.005	0.007	0.005	-0.006	-0.002
Age 55-64	0.026**	0.038**	0.021**	0.036**	0.035**	0.026**	0.013**
Age ≥ 65	0.086**	0.085**	0.065**	0.085**	0.081**	0.079**	0.039**
Diploma or certificate	0.023**	0.018**	0.019**	0.013**	0.014**	0.024**	0.008*
University Degree	0.054**	0.039**	0.038**	0.029**	0.031**	0.040**	0.021**
Canada	0.017*	0.016*	0.010	0.014*	0.016**	0.014*	0.009
Germany	-0.038**	-0.036**	-0.043**	-0.039**	-0.040**		-0.040**
Norway	0.056**	0.041**	0.036**	0.030**	0.034**		0.028**
UK	-0.025**	-0.015*	-0.018**	-0.014*	-0.013*	-0.020**	-0.006
US	0.000	0.005	0.001	0.003	0.005	0.001	-0.002
Constant	0.684**	0.401**	0.199**	0.400**	0.039*	0.380**	0.262**
R^2	0.170	0.287	0.329	0.337	0.332	0.231	0.473
Observations	7,919	7,919	7,918	7,918	7,918	5,473	7,919

Notes:

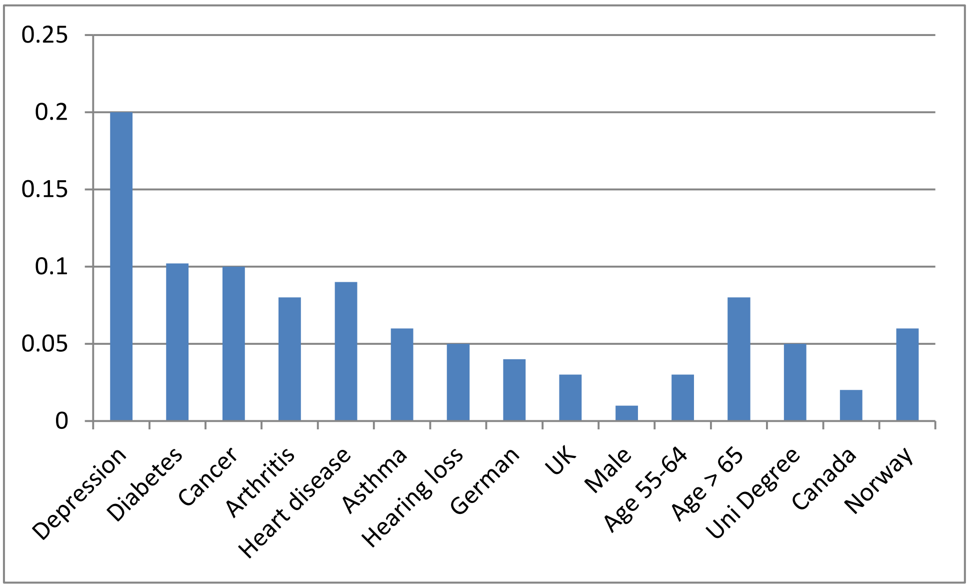
Dependent variable: SWB. ** p<0.01, * p<0.05. For the first two panels where SWLS and ONS were dependent variables, socio-demographic variables and country dummies shown in the third panel were all included.

Figure 1 Correlation (SWB, MAUI) <u>vs</u> Psycho-Social Content of MAUI (%)



$$*\left[\sum_{i=1}^{4} B\ e(Pas-yS\ o)\right] / \left[\sum_{i=1}^{8} B\ e\ S\ t\ dF\ 6D\ i\ m\ e\ n\right]$$

Figure 2 Reduction in PWI associated with independent variables

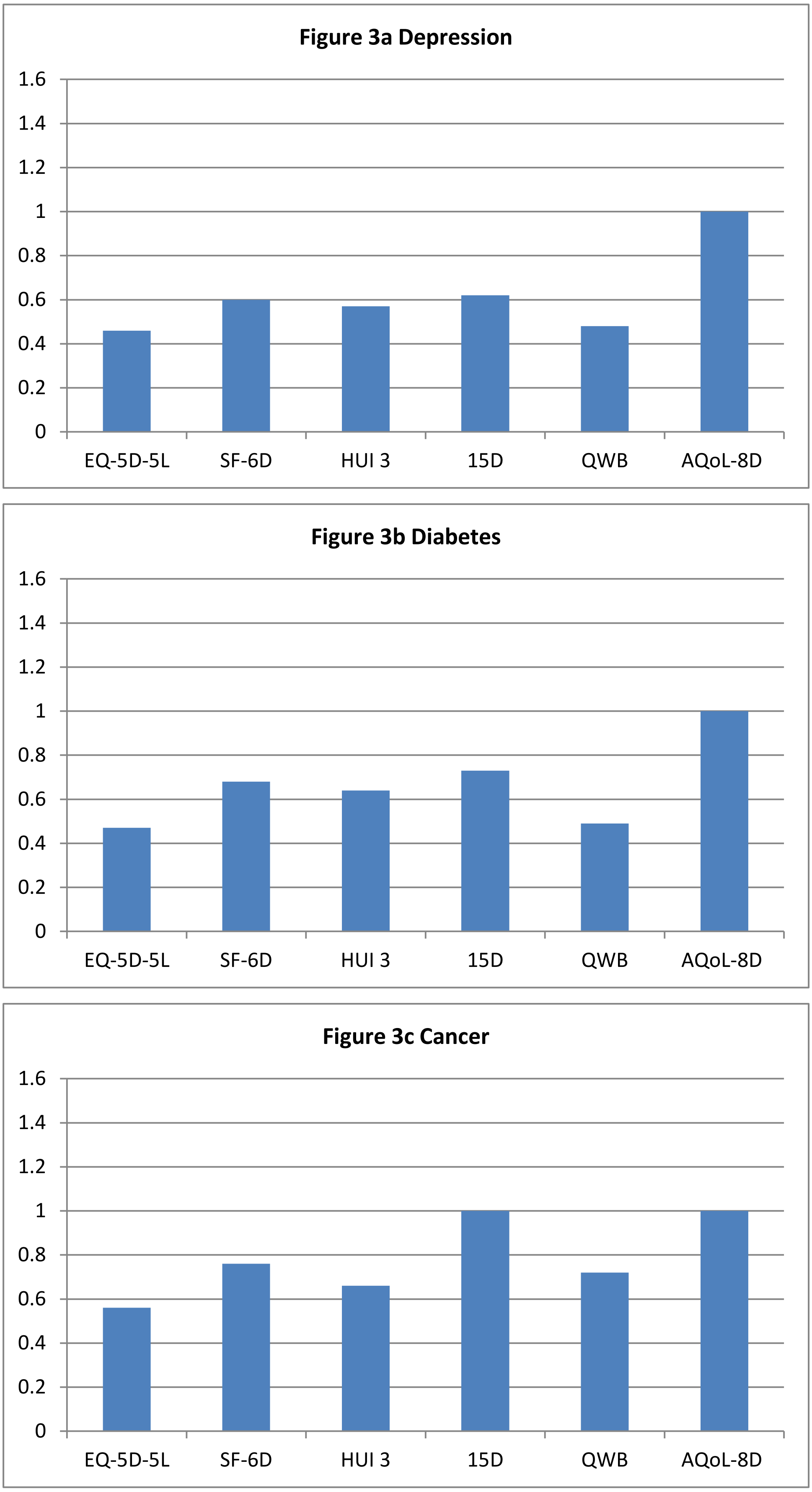


Source: Table 6 column 1

Notes:

Negative reductions represent higher PWI scores

Figure 3 Proportion of reduced PWI explained by MAU instruments



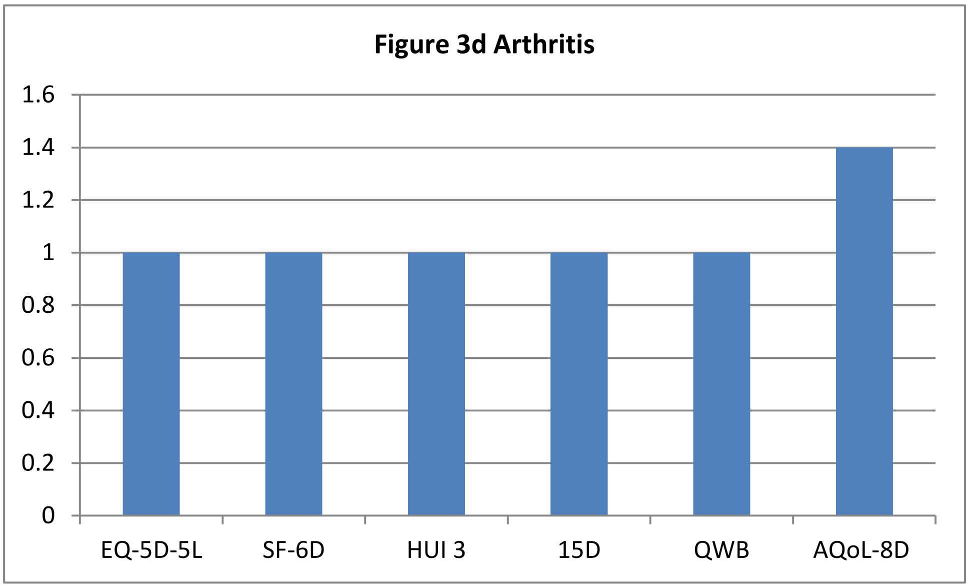


Figure 3d Arthritis

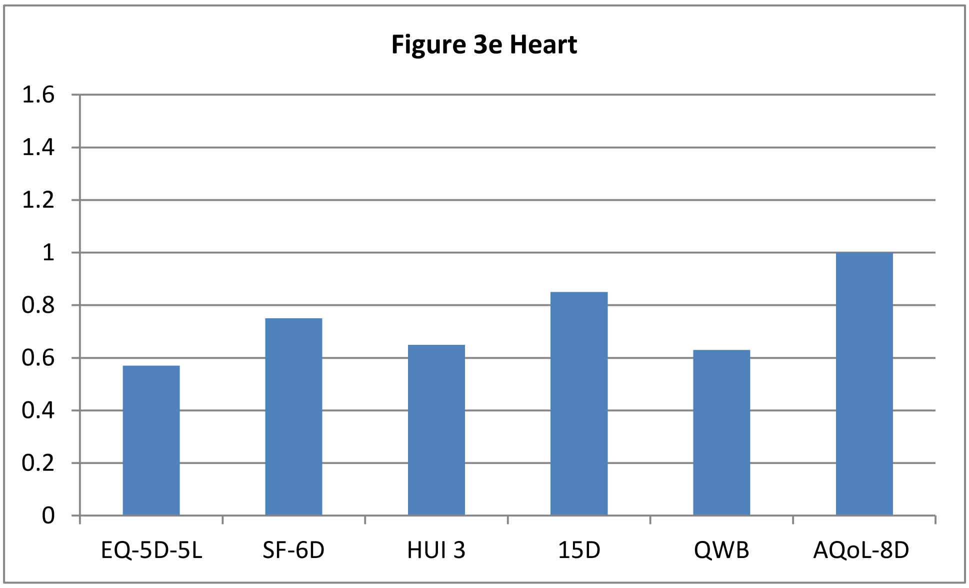


Figure 3e Heart

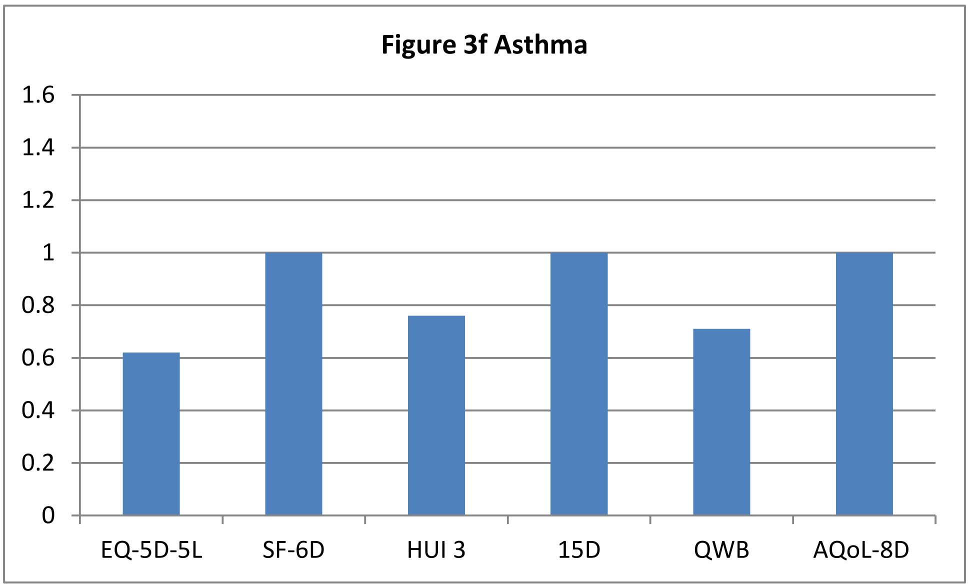


Figure 3f Asthma

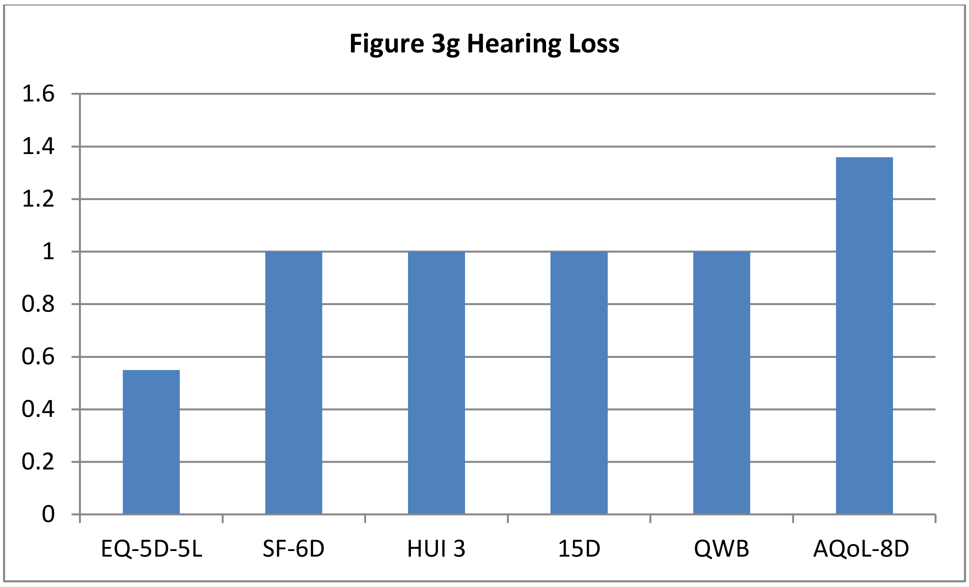


Figure 3g Hearing Loss
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
EQ-5D-5L
SF-6D
HUI 3
15D
QWB
AQoL-8D

Supplementary material

Supplementary Table S.1 Regression of PWI upon SWLS and ONS (n=8008)

SWLS	= 0.02 +	0.56 PWI (0.04)	+ 0.39 PWI^2 (0.04)	$R^2 = 0.63$
ONS	= 0.11 +	0.73 PWI (0.04)	+ 0.14 PWI^2 (0.03)	$R^2 = 0.61$

Supplementary Table S.2 SWB equation with AQoL-8D separately modified for patients with arthritis and hearing loss[(1)]

	No utility	Dep Var = SWLS		Dep Var = ONS		Dep Var = PWI	
	1	2	3	4	5	6	7
AQoL-8D		0.716**		0.666**		0.552**	
AQoL-8D_adjusted			0.679**		0.661**		0.548**
Depression	-0.221**	-0.003	-0.026**	-0.014*	-0.016*	-0.005	-0.006
Diabetes	-0.101**	0.009	0.003	0.015*	0.014*	-0.004	-0.005
Cancer	-0.096**	0.017*	0.010	0.011	0.010	0.000	-0.001
Arthritis	-0.081**	0.044**	0.0045**	0.058**	0.022**	0.030**	-0.000
Heart disease	-0.080**	0.025**	0.019*	0.016*	0.015*	-0.001	-0.001
Asthma	-0.062**	0.018*	0.013	0.020**	0.019**	0.007	0.006
Hearing loss	-0.037**	0.044**	0.033**	0.023**	0.001	0.018**	-0.000
R-squared	0.148	0.453	0.450	0.513	0.509	0.473	0.470

**p<0.01, *p<0.05

(1) Age and country coefficients are not reported.

Supplementary Figure S.1 (a-f) Frequency distributions

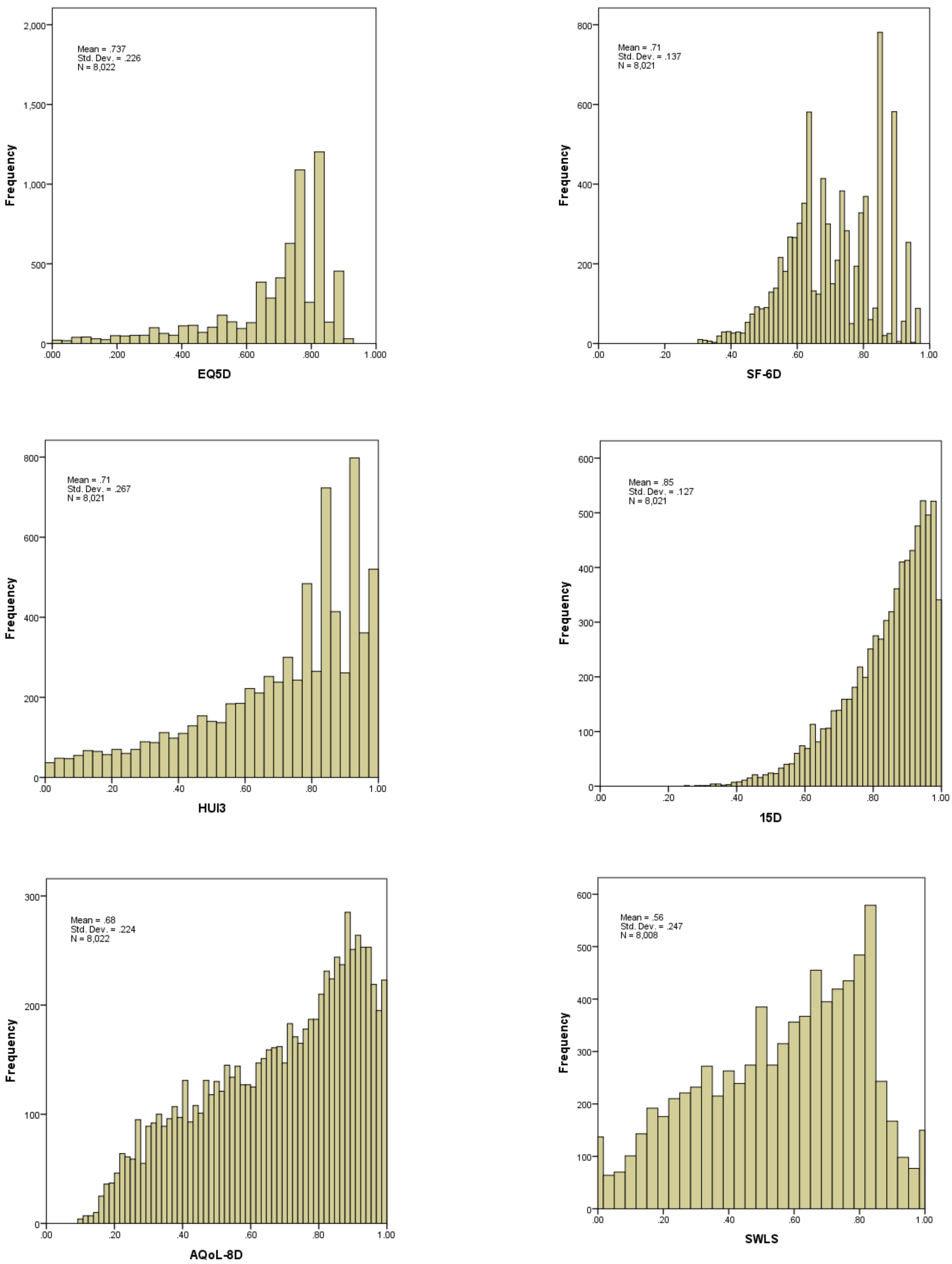
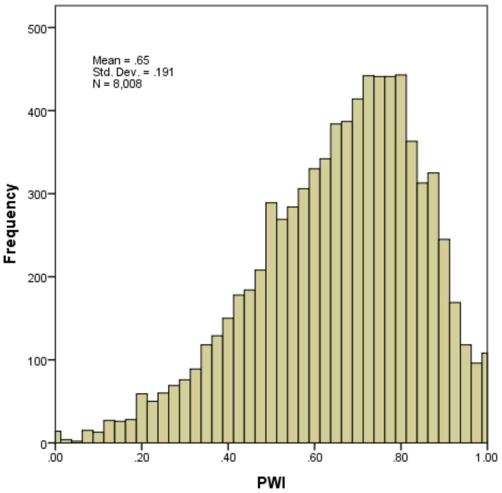
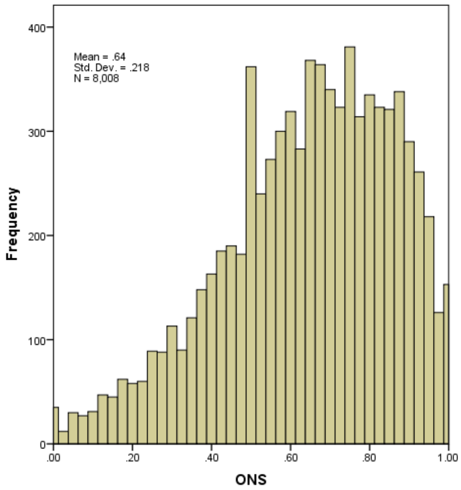


Figure S.1 (g-h) Frequency distributions





Supplementary Note 1

Adjusting for Hedonic Adaptation

The linear relationship between SWB and the utility predicted by an MAU instrument in Test 2 is given by equation S.1

$$SWB = a_i\ MAU_i^1 + \sum_{i-1}^{7} b_j\ DUM_j \qquad \text{… eq S.1}$$

where MAU_i^1 = original estimate of utility from MAU instrument i, and DUM_j = dummy variable for disease j.

Hedonic adaptation is identified by a significant and positive b_j which indicates that $a_i\ MAU_i^1$ underestimates SWB by an amount b_j. If $a_i\ MAU_i^2$ provides a correct estimate of SWB, then

$$a_i\ MAU_i^2 = a_i\ MAU_i^1 + b_j \text{ where } b_j>0$$

$$MAU_i^2 = MAU_i^1 + b_j / a_i$$

An adjustment to the utility algorithm which increases MAU_i^1 by b_j/a_i therefore will correct the error.

Adjusting AQoL-8D using PWI

The utility algorithm for AQoL-8D takes the form given in equation S.2 (Richardson et al., 2014b).

$$\text{AQoL-8D} = .(\text{Mult})^{\alpha} \qquad \text{… eq S.2}$$

$$\alpha = a_i\ = + \sum_{i=1}^{8} c_i\ Dim_i$$

where Mult = the multiplicative combination of AQoL-8D dimension scores and Dim_i = the 8 dimension scores of the AQoL-8D.

From the last column of Table 6, a_i = 0.552 and b_j for arthritis and hearing loss are 0.03 and 0.018 respectively. The AQoL-8D formula, equation S.2 must therefore be adjusted to an increase predicted utilities for arthritis and hearing loss patients by b_j/a_i = 0.03/0.552 = 0.054

and 0.018/0.552 = 0.033 respectively. This is achieved when the coefficient, c_i , in equation S.2 for pain is adjusted from -0.0406 to -0.1486 and for senses from -0.1495 to -0.2315. AQoL-8D scores using the two revised formula were used to re-estimate the utilities for patients in the two disease groups. In each case, the revised utilities were combined with unadjusted AQoL-8D utilities for other patients and the public. Re-estimated regressions are reported in Table S.2. Dummy variables in the equation for PWI are now insignificant. Over-prediction occurs in the equations for SWLS and ONS reflecting the differences between the three measures of SWB. However similar adjustment could be made for either of these measures if it emerged as the preferred measure of SWB in the literature.