This is a repository copy of *Estimating health state utility values for comorbid health conditions using SF-6D data*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/11049/

Article:

White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Estimating health state utility values for comorbid health conditions using SF-6D data

Ara Roberta, Brazier John

Health Economics and Decision Science, ScHARR, The University of Sheffield, 30 Regent Street, Sheffield S1 4DA. UK

Telephone: 44114 222 0788

Fax 44 114 272 4095

Email r.m.ara@sheffield.ac.uk

ABSTRACT

BACKGROUND: When health state utility values for comorbid health conditions are not available, analysts frequently use data from cohorts with single health conditions to estimate proxy scores. The methods used can produce very different results and there is currently no consensus on which is the most appropriate approach.

OBJECTIVE: The objective of the current study was to assess the accuracy of five different methods that have been used to estimated HSUVs for comorbid health conditions.

METHOD: Data collected during five Welsh Health Surveys (WHS) were subgrouped by health status. Mean SF-6D scores from cohorts with a particular health condition were used to estimate mean SF-6D scores for cohorts with two comorbid health conditions using:the additive, multiplicative, and minimum methods, and the adjusted decrement estimator. A linear model was obtained by regressing mean HSUV from subgroups with single health conditions onto mean HSUVs from subgroups with combined health conditions.

RESULTS: The pooled WHS data provided 64,437 cases with SF-6D scores. When subgrouped by self reported health condition(s), 32 groups (n>30) were identified with comorbid health conditions. The mean SF-6D for these subgroups ranged from 0.4648 to 0.6068. The linear model produced the most accurate HSUVs for the combined health conditions with 88% of values accurate to within the minimum important difference for the SF-6D. The additive method underestimated the actual SF-6D scores and produced some substantial errors in the estimated values. The minimum method overestimated all mean SF-6D scores but was more accurate when estimating higher values. The multiplicative and ADE methods both underestimated the majority of the actual SF-6D scores. However, both methods both performed better when estimating SF-6D scores smaller than 0.50 with 43% and 86% of estimated HSUVs accurate to within the MID for the multiplicative and ADE respectively.

This study makes an important contribution to the existing evidence as it is the first to compare five different methods on SF-6D data. Although the range in actual HSUVs was relatively small, the data covered the lower end of the index while the majority of previous research has involved actual HSUVs covering the upper end of possible ranges. While the linear model gave the most accurate results in our data, additional research is required to develop and validate the model.

BACKGROUND

Policy decision makers such as the National Institute for Health and Clinical Excellence (NICE) in the UK recommend that the results of economic evaluations in healthcare are presented in terms of quality adjusted life years (QALY).[1] The QALY quantifies both health related quality of life and life expectancy in a single metric and allows comparison across disparate diseases and interventions.[2] The health state utility values (HSUVs) used to weight the QALYs are obtained from preference-based measures of health such as the EQ-5D, and the SF-6D.[3,4]

While there is a large evidence base for HSUVs associated with single health conditions, due to the large number of possible combinations of health conditions, studies reporting HSUVs for comorbid health conditions are limited. When these data are not available HSUVs for comorbid health conditions are estimated using the HSUVs obtained from people with single conditions. For example, the mean HSUV for a comorbid health condition defined as both condition A and condition B would be estimated using the mean HSUVs obtained from cohorts with condition A (but not condition B) and the mean HSUV from cohorts with condition B (but not condition A). The three most frequently used techniques are the additive, multiplicative and minimum methods. The additive and multiplicative methods assume a constant absolute or constant proportional decrement respectively while the minimum method attributes no additional health decrement, taking the smallest HSUV from the single health conditions involved.

The evidence base describing empirical research in this area is limited and there is currently no consensus on the most appropriate approach. The multiplicative method gave a good fit on HUI3 data from the Canadian Community Health Survey;[5] and was more accurate than the additive method on EQ-5D data from the Medical Expenditure Panel Survey (MEPS).[6] The minimum method performed better than both the additive and the multiplicative methods on EQ-5D data from the MEPS.[7] More recently, a variation of the minimum method, the adjusted decrement estimator (ADE) has been proposed and was shown to outperform the three other methods on EQ-5D data from MEPS.[8]

The methods can produce very different HSUVs and it has been shown that these differences are great enough to potentially influence a policy decision based on a cost per QALY threshold.[9] This undermines the rational for consistent reimbursement recommendations and optimal resource allocation. The objective of the current study was to assess the accuracy of all four methods using SF-6D data collected in the Welsh Health Surveys. We compare these results with values predicted using a parametric model which maps from mean HSUVs

obtained from cohorts with single health conditions onto HSUVs for cohorts with comorbid health conditions.

METHODS

The Welsh Health Survey (WHS) is an annual survey which draws from a random sample of the population living in private households in Wales.[10-14] Responses collected during the surveys conducted in the years 2003, 2004, 2005, 2007 and 2008 were pooled for use in the current study. HSUVs were obtained using the SF-6D (v2) preference-based measure which is derived from responses to the SF-36 generic health questionnaire.[15] The SF-6D is a six-dimensional health state classification system assessing physical functioning, role-limitations, social functioning, pain, mental health and vitality. The classification system generates a total of 18,000 possible health states. Weights for the SF-6D preference-measure used in the current study were obtained from a random sample of the UK general population using anchors of zero and one to represent death and perfect health respectively.[3] The SF-6D is scored on a continuous index whereby 0.296 represents the maximum impaired level on all six dimensions and 1 represents the least impaired level.

In addition to questions on health related quality of life, respondents were asked to identify if limiting long term health conditions and the coded data details information on 39 individually categorised and 14 grouped limiting long term health conditions (see online Appendix A). All analyses are weighted using the individual level self-administered questionnaire weights which adjusts for non response.

The "actual" mean SF-6D scores were calculated for subgroups ($n \geq 30$) of respondents with comorbid pairs of health conditions (condition A and condition B), and for subgroups with condition A (and not condition B) or condition B (and not condition A). The latter were then used to estimate mean SF-6D scores for the cohorts with comorbid health conditions using the methods described below. The relationship between the SF-6D scores from cohorts with single health conditions and the actual SF-6D scores was also explored using ordinary least square regression (OLS). The OLS model incorporates the additive and minimum method with a multiplicative interaction term as described in the next section.

Methods used to estimate HSUVs

For the two health conditions, condition A and condition B, the following combinations are possible: condition A and condition B; neither condition A or condition B; condition A but

4

not condition B; condition B but not condition A. The HSUVs associated with these alternatives are defined to be $U_{A,B}$, $U_{nA,nB}$, U_A, and U_B respectively.

The additive method assumes a constant absolute detriment relative to the baseline. When assuming a baseline of full health the additive method is written as:

$$U_{A,B}^{add} = 1 - \left[(1 - U_A) - (1 - U_B) \right]$$
(1)

Using an adjusted baseline (see next section) the additive method is written as:

$$U_{A,B}^{add} = U_{nA,nB} - \left((U_{nA} - U_A) + (U_{nb} - U_B) \right)$$
(2)

The multiplicative method assumes a constant proportional detriment relative to the baseline. When assuming a baseline of full health, the multiplicative method is written as:

$$U_{A,B}^{mult} = U_A \cdot U_B$$
(3)

When using an adjusted baseline, the multiplicative method is written as:

$$U_{A,B}^{Mult} = U_{nA,nB} \cdot \left(\frac{U_A}{U_{nA}} \right) \cdot \left(\frac{U_B}{U_{nB}} \right)$$
(4)

The minimum method assumes the impact on HRQoL for a comorbid health condition is equivalent to the most severe of the single health conditions. I.e. there is no additional decrement associated with a second health condition. When assuming a baseline of full health, the minimum method is written as:

$$\hat{U}_{A,B}^{min} = \min(U_A, U_B) + \varepsilon$$
(5)

When using an adjusted baseline, the minimum method is written as:

$$U_{A,B}^{min} = \min(U_{nA,nB}, U_A, U_B)$$
(6)

The adjusted decrement estimator (ADE), proposed by Hu, assumes the HSUV for the comorbid health condition is bound by the minimum HSUV of the two HSUVs for the single health conditions and is written as:[8]

$$U_{A,B}^{ADE} = \min(U_A, U_B) - \min(U_A, U_B) \cdot (1 - U_A) \cdot (1 - U_B)$$
(7)

In addition to the methods described above, a simple linear model has been proposed.[16] Based on decision theory, multi-attribute utility functions,[17,18] and a prospect theory[19] the model incorporates terms that represent the additive, multiplicative and minimum methods.[16] The model is defined by:

$$U_{A,B}^{comb} = \beta_0 + \beta_1 \cdot \min\left((1 - U_A), (1 - U_B)\right) + \beta_2 \cdot \max\left((1 - U_A), (1 - U_B)\right) \\ + \beta_3 \cdot \left(U_A \cdot U_B\right) + \varepsilon \qquad (8)$$

whereby the beta coefficients are obtained using ordinary least square regressions and ε represents the residual. We use the following adaptation which incorporates an adjusted baseline:

$$U_{A,B}^{comb} = \beta_0 + \beta_1 \cdot \min\left((U_{nA} - U_A), (U_{nb} - U_B)\right) + \beta_2 \cdot \max\left((U_{nA} - U_A), (U_{nb} - U_B)\right) \\ + \beta_3 \cdot \left(U_{nA,nB} \cdot \frac{U_A}{U_{nA}} \cdot \frac{U_B}{U_{nb}}\right) + \varepsilon \qquad (9)$$

When using a baseline of full health it is assumed that if a particular health condition is alleviated, the HSUV for the health condition will revert to 1 on a preference based utility index. However, this assumption ignores the natural decline in health due to age and additional comorbidities and overestimates the decrement on health related quality of life associated with health conditions.[20] Consequently this may not be the most appropriate technique when estimating HSUVs for comorbid health conditions. Several alternatives have been suggested and these include: "purifying" data by dividing all HSUVs by the mean HSUV obtained from individuals with none of the health conditions,[5] or using HSUVs associated with not having specific health conditions.[6] We used age adjusted baseline HSUVs obtained from respondents who do not have any of the health conditions identified in the WHS.

The methods used to estimate HSUVs for the combined health conditions were assessed in terms of errors (actual minus estimated) in the estimated HSUVs. In addition to the statistics generally reported (mean absolute errors (MAE), mean squared errors (MSE), root mean squared error (RMSE)), the proportion of errors within the minimum important difference (MID) for the SF-6D (MID = 0.041)[21] were calculated and the magnitude of errors across the actual SF-6D scores were examined using scatter plots

RESULTS

The pooled data included 64,437 cases with SF-6D scores. The mean SF-6D for the full sample irrespective of health status was 0.7613 (range 0.301 to 1). The mean SF-6D for respondents (16414/64437) who reported having at least one limiting long term health condition was 0.6055 (se 0.0011) compared with 0.8104 (se 0.0006) for respondents who reported no limiting long term health condition. There were just 2,021 respondents who

reported two or more limiting long term illnesses and 32 subgroups (n≥30) with two concurrent conditions (see Appendix for details). The mean SF-6D scores (Figure 1) for these subgroups ranged from 0.4648 (se 0.0086) for respondents (n=140) who reported both a mental disorder and a musculoskeletal condition to 0.6068 (se 0.0269) for respondents (n=33) who reported both arthritis/ rheumatism/fibrositis and an unclassifiable complaint. As can be seen in Figure 1, the SF-6D scores are clustered around the mean (0.5301) with just 4/32 groups scoring less than 0.50 or greater than 0.60. When comparing mean SF-6D scores for subgroups, all scores from the groups with comorbid health conditions were smaller than those from the subgroups with the corresponding single health conditions.

INSERT FIGURE 1: Distribution of mean SF-6D scores for subgroups (n=32) with two comorbid health conditions

The linear model estimated using ordinary least squares is provided in Table 1. The coefficients for all three independent variables are negative as would be expected as they are the decrements associated with the health conditions. When comparing the magnitude of the coefficients, the coefficient for the condition with the maximum decrement is larger than the coefficient for the condition with the minimum decrement as might be expected. The weight associated with the interaction term (p = 0.661) is similar to that for the health condition with the minimum disutility.

INSERT TABLE 1: Results from the OLS combination model

A summary of the results obtained using the five alternative techniques is provided in Table 2. Overall, the HSUVs obtained using the linear model are the most accurate producing the smallest MAE (0.0191) and the smallest RMSE (0.0254) in the predicted mean SF-6D values. Although the average of the predicted mean SF-6D scores equals the actual value of 0.5301, the range is somewhat truncated (predicted range: 0.4935 to 0.5549, actual range: 0.4368 to 0.6068). However, all predicted HSUVs are within the MID for the SF-6D and 75% have errors smaller than |0.025|.

Of the four nonparametric methods, the ADE outperforms the other three having the smallest MAE (0.0419) and smallest RMSE (0.0471). When examining accuracy in predicting the individual mean SF-6D scores, the ADE does not compare favourably with the linear model and only 47% (25%) of estimated HSUVs are accurate to within the |MID| (|0.025|). The

additive, multiplicative and minimum methods perform less well with just 3%, 6%, and 13% of estimated HSUVs within the |MID| respectively.

INSERT TABLE 2: Comparing the predictive abilities of the four methods

Figure 2 shows the actual and estimated mean SF-6D scores. It is clear that the minimum method overestimates the actual SF-6D scores and the errors increase as the actual SF-6D score decreases. The additive, multiplicative and ADE methods underestimate the majority of the actual SF-6D scores. While the linear model produces the most accurate scores there is a tendency for the errors to be larger at the extremes of the range of actual scores.

INSERT FIGURE 2: Plot of actual and estimated SF-6D scores

When sub-grouped by actual SF-6D score (Table 3) it can be seen that the value of the SF-6D score being estimated can influence the accuracy of the methods. For example, while the minimum method was the least accurate in terms of mean errors overall, it performs better than all the other nonparametric methods when estimating actual SF-6D scores greater that 0.55 and 71% of these estimated values are accurate to within the MID. Similarly, when estimated SF-6D smaller than 0.50, the ADE produces 86% of HSUVs accurate to within the MID compared with 57% of values predicted using the linear model. The additive method does not perform well across the full range of actual SF-6D scores while the multiplicative method performs better when estimating SF-6D scores smaller than 0.50.

DISCUSSION

The objective of the current study was to add to the existing evidence base by comparing the accuracy of methods frequently used to estimate HSUVs for comorbid health conditions. Using SF-6D data obtained from respondents taking part in Welsh Health Surveys, we found that the linear model obtained using OLS regression out-performed the non parametric methods. Overall 88% of HSUVs predicted using the linear model were within the MID of the SF-6D. The additive method underestimated the actual SF-6D scores and produced some substantial errors with none of the estimated HSUVs within the MID for the SF-6D. Although the minimum method overestimated the actual HSUVs it performed better when estimating SF-6D scores greater than 0.55 with errors in estimated values increasing as actual SF-6D scores decreased. The multiplicative and ADE methods both underestimated the majority of the actual SF-6D scores. However, when looking at subgroups of actual SF-6D

8

scores, both methods both performed better when estimating SF-6D scores smaller than 0.50 with 43% and 86% of estimated HSUVs accurate to within the MID for the multiplicative and ADE respectively.

Our findings are similar to those reported in a recent publication using EQ-5D data obtained from the MEPS.[8] The range in actual HSUVs for the comorbid health conditions ranged from 0.62 to 0.88 and the authors reported the ADE model outperformed the additive, minimum and multiplicative methods when assessed by MEs and RMSE in estimated values. Charts of the estimated and actual EQ-5D scores showed the additive and multiplicative methods underestimated the actual EQ-5D scores and the magnitude of errors increased as the actual EQ-5D score increased for both methods. As in our data, the minimum method performed better for higher HSUVs with the magnitude of errors increasing as the actual HSUV decreased. While the ADE performed better than the other methods overall, the magnitude of errors in estimated values grew substantially as the actual EQ-5D score decreased.

There are three limitations relating to the data used in the current study. First, the range in actual mean SF-6D scores (0.4648 to 0.6068) for the comorbid health conditions covered only 24% of the possible range (0.29 to 1) and all values were in the bottom half of the SF-6D index (i.e. below 0.65). Actual mean HSUVs for comorbid health conditions reported in other studies tend to be in the upper range of the preference based indices. For example, Hu and Fu used data from MEPS and their actual EQ-5D scores ranged from 0.62 to 0.88 which equates to 24% of the possible range (-0.1 to 1) for the US EQ-5D index.[7,8,22] Janssen used a similar dataset and reported actual mean EQ-5D scores for comorbid health conditions ranging from (0.734 to 0.819).[6] The widest range (-0.01 to 1) of actual mean HSUVs for comorbid conditions was reported in a dataset of HUI3 scores obtained from the Canadian Community Health Survey (2001, 2003). However, the majority (184/278) of scores were greater than 0.80. One possible explanation for the differences in the ranges for the actual HSUVs is that the respondents in our dataset were asked to identify limiting long standing illnesses, while the respondents in the surveys for the other studies were asked to identify chronic health conditions. The consequence of this is that the respondents in the WHS may not have reported health conditions they did not perceive to affect their HRQoL. As the accuracy in the estimating methods has been show to vary depending on the range of the scores estimated in both this study and Hu's it is possible that different conclusions would be drawn if the methods were tested in datasets that covered the full ranges of the indices.

Second, we were only able to identify 32 subgroups with comorbid health conditions and the number of cases in some of the subgroups was relatively small (n= 30 to 346). As a consequence we did not estimate HSUVs for comorbid health conditions consisting of greater than two health conditions. While Flanagan et al. assessed the accuracy of the multiplicative method in estimating HSUVs for comorbid health conditions consisting of more than two health conditions, as far as we are aware, this is the only research in this area and no-one has compared results for multiple comorbid health conditions using alternative methods to date.[5]

Third, although we obtained a linear model to predict SF-6D scores for the comorbid health conditions, the number of cases used in the regression was small (n=32) and none of the coefficients in the model were statistically significant. As the model tends to over predict the lower SF-6D scores and under predict the higher SF-6D scores it is possible that a different model specification would produce more accurate results and additional research exploring alternatives is warranted.

CONCLUSION

Despite the limitations in the data, this study makes an important contribution to the evidence base. It is the first study to compare the five different techniques on SF-6D data and although the range of estimated scores was relatively small, they covered the lower end of the preference based index whilst the majority of other research in this area has involved actual HSUVs at the top end of the preference measures. While the linear model gave the most accurate results in our sample, additional research is required to develop and validate the model.

REFERENCES

1.   NICE Guide to the methods of technology appraisal. 2008. London, NICE

2.   Gold MR, Siegel JE, Cost-Effectiveness in Health and Medicine. Oxford University Press ISBN 0-19-510824-8

3.   Brazier JE, Roberts JF, Deverill MD. The estimation of a preference-based measure of health from the SF-36. J Health Econ 2002;21:271-92.

4.   Dolan P, Gudex C, Kind P, Williams A. The time trade-off method: results from a general population study. Health Econ 1996;5:141-54.

5.   Flanagan W, McIntosh C, Le Petit C, Berthelot J. Deriving utility scores for co-morbid conditions: a test of the multiplicative model for combining individual condition scores. Population Health Metrics 2006:4(13) doi:10.1186/1478-7954-4-13.

6.   Janssen M, Bonsel G Estimating preference weights for chronic multimorbidity: Don't add, multiply. Proceedings of the EuroQol Group 2008.

7.   Fu A, Katan M. Utilities should not be multiplied. Medical Care 2008;46:984-90.

8.   Hu B, Fu AZ. Predicting utility for joint health states: a general framework and a new nonparametric estimator. MDM (In press)

9.   Ara R, Brazier J, The issues involved in populating an economic model with utility values: moving towards better practice. ViH Online: Mar 10 2010 2:53AM DOI: 10.1111/j.1524-4733.2010.00700-x.

10.   National Centre for Social Research, Welsh Health Survey, 2008 [computer file]. Colchester, Essex: UK Data Archive [distributor], January 2010. SN: 6372.

11.   National Centre for Social Research, Welsh Health Survey, 2007 [computer file]. Colchester, Essex: UK Data Archive [distributor], October 2008. SN: 6052.

12.   National Centre for Social Research, Welsh Health Survey, 2005-2006 [computer file]. Colchester, Essex: UK Data Archive [distributor], November 2007. SN: 5750.

13.   National Centre for Social Research, Beaufort Research Limited and University College London. Department of Epidemiology and Public Health, Welsh Health Survey, 2004-2005 [computer file]. Colchester, Essex: UK Data Archive [distributor], September 2007. SN: 5693.

14.   National Centre for Social Research, Beaufort Research Limited and University College London. Department of Epidemiology and Public Health, Welsh Health Survey, 2003-2004 [computer file]. Colchester, Essex: UK Data Archive [distributor], September 2007. SN: 5692.

15.   Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). Conceptual framework and item selection. Medical Care 1992:30:473-483.

16.   Basu A, Dale W, Elstein A, Meltzer D. A linear index for predicting joint health state utilities from single health state utilities. Health Economics 2009;18:403-19.

17.   Keeney RL, Raifa H. 1993. The multiplicative utility function. In Decisions with Multiple Objectives: Preferences and Value Tradeoffs. Cambridge University Press: Cambridge; 288-292.

18.   Keeney RL, Raiffa H. 1976. Decision Making with Multiple Objectives. Wiley: New York.

19.   Tversky A, Kaheman D. 1992. Advances in prospect theory: cumulative representation of uncertainty. Journal of Risk and Uncertainty 5: 297-323.

20.   Fryback DG, Lawrence WF, Dollars may not buy as man QALYs as we think: A problem with defining quality of life adjustments, Med Decis Making 1997;17;276.

21.   Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D SF-6D. Qual Life Res 2005;14:1423-32.

22.   Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the DI valuation model. Med Care 2005;43:203-220.

Table 1 : Results from the OLS combination model

Independent variable	Coefficient	Robust Std. Err.	P>|t|
Maximum decrement	-1.049809	0.6162	0.099
Minimum decrement	-0.4797101	0.89592	0.597
Interaction of utilities	-0.4986031	1.12376	0.661
Constant	1.0606200	0.77913	0.184
R Sq	0.3472		

Maximum decrement	$\max\left((U_{nA}-U_A),(U_{nb}-U_B)\right)$
Minimum decrement	$\min\left((U_{nA}-U_A),(U_{nb}-U_B)\right)$
Interaction of utilities	$\left(U_{nA,nB}\cdot\dfrac{U_A}{U_{nA}}\cdot\dfrac{U_B}{U_{nb}}\right)$

Table 2: Comparing the predictive abilities of the four methods

	Actual	Additive	Multiplicative	Estimated Minimum	ADE (Hu)	Linear model
Mean SF-6D	0.5301	0.4092	0.4556	0.5848	0.4918	0.5301
Min SF-6D	0.4368	0.3453	0.4115	0.5620	0.4656	0.4935
Max SF-6D	0.6068	0.4794	0.5077	0.6053	0.5169	0.5549
Range:	0.1700	0.1341	0.0962	0.0433	0.0513	0.0614
Mean error		0.1209	0.0745	-0.0546	0.0383	0.0000
Maximum absolute error		0.1924	0.1496	0.1316	0.1196	0.0669
MAE		0.1209	0.0747	0.0563	0.0419	0.0191
MSE		0.0157	0.0064	0.0038	0.0022	0.0006
RMSE		0.1252	0.0799	0.0613	0.0471	0.0254
Proportion within |0.01|		0%	3%	0%	0%	31%
Proportion within |0.025|		0%	6%	6%	25%	75%
Proportion within MID |0.041|		3%	6%	13%	47%	88%

Table 3 : Errors in estimated HSUVs subgrouped by actual SF-6D score

Actual SF-6D score	n	Additive	Multiplicative	Minimum	ADE (Hu)	Linear model
Mean error						
SF-6D > 0.55	7	0.1399	0.0943	-0.0260	0.0636	0.0256
0.55 ≤ SF-6D < 0.50	18	0.1271	0.0762	-0.0528	0.0403	-0.0003
SF-6D ≥ 0.50	7	0.1105	0.0463	-0.0880	0.0080	-0.0249
Mean absolute error						
SF-6D > 0.55	7	0.1399	0.0943	0.0334	0.0636	0.0256
0.55 ≤ SF-6D < 0.50	18	0.1271	0.0762	0.0528	0.0403	0.0120
SF-6D ≥ 0.50	7	0.1105	0.0463	0.0880	0.0243	0.0308
Root mean squared error						
SF-6D > 0.55	7	0.1430	0.0978	0.0342	0.0683	0.0324
0.55 ≤ SF-6D < 0.50	18	0.1302	0.0788	0.0545	0.0430	0.0152
SF-6D ≥ 0.50	7	0.1172	0.0553	0.0916	0.0270	0.0360
Accurate to within the |MID|						
SF-6D > 0.55	7	0%	0%	71%	0%	86%
0.55 ≤ SF-6D < 0.50	18	0%	6%	17%	50%	100%
SF-6D ≥ 0.50	7	14%	43%	0%	86%	57%

Figure 1: Distribution of mean SF-6D scores for subgroups (n=32) with two comorbid health conditions
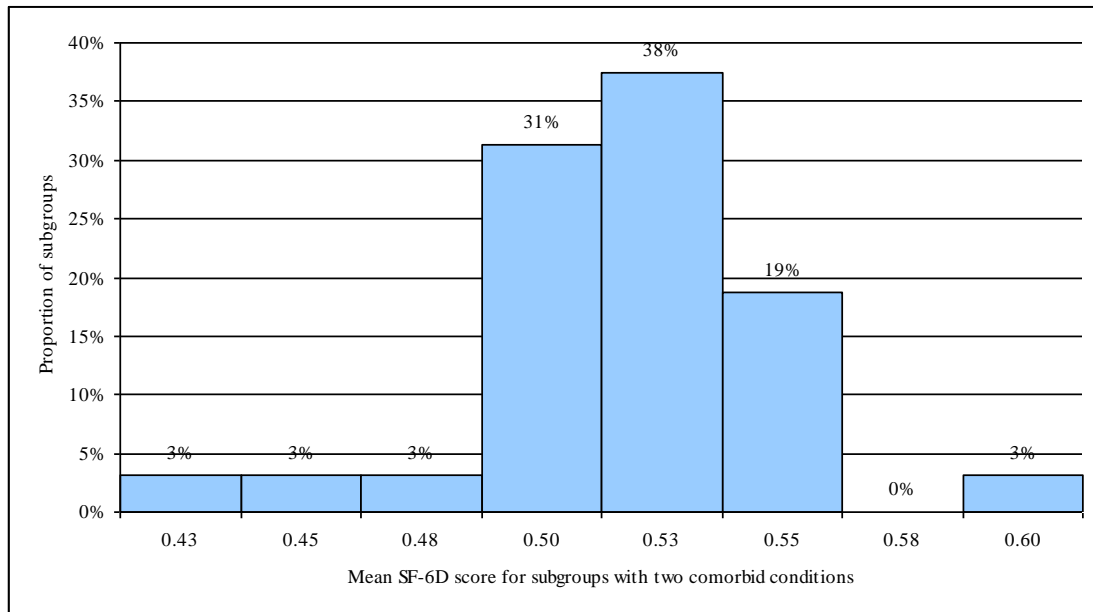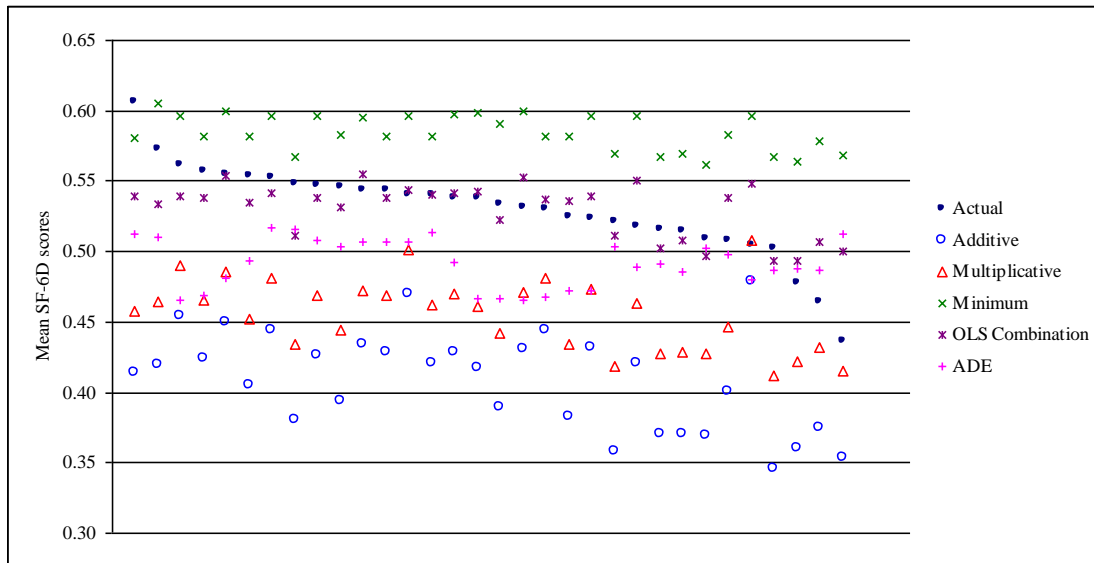
Figure 2: Actual and estimated mean SF-6D scores

APPENDIX A:

TableA1 description of comorbid conditions and mean SF-6D scores

		N		Mean SF-6D score		Age adjusted baseline (from respondents with none of health conditions	
				Condition			
Condition A	Condition B	N	Condition A&B (Actual)	Condition A (not Condition B)	Condition B (not Condition A)	Condition A	Condition B
endocrine and metabolic diseases	heart and circulatory	85	0.5553	0.6372	0.5993	0.7933	0.7808
endocrine and metabolic diseases	musculoskeletal	103	0.5523	0.6393	0.5967	0.7924	0.7922
mental disorders	nervous system	34	0.5020	0.5671	0.5862	0.8127	0.8033
mental disorders	heart and circulatory	45	0.5161	0.5672	0.5994	0.8132	0.7803
mental disorders	musculoskeletal	140	0.4648	0.5788	0.5985	0.8133	0.7919
nervous system	musculoskeletal	139	0.5340	0.5905	0.5972	0.8037	0.7920
eye complaints	musculoskeletal	33	0.5052	0.6691	0.5965	0.7801	0.7923
ear complaints	musculoskeletal	40	0.5405	0.6662	0.5964	0.7936	0.7922
heart and circulatory	respiratory system	85	0.5316	0.6002	0.6167	0.7808	0.7911
heart and circulatory	musculoskeletal	277	0.5377	0.6058	0.5981	0.7812	0.7927
respiratory system	musculoskeletal	175	0.5378	0.6222	0.5974	0.7903	0.7921
digestive system	musculoskeletal	82	0.5475	0.6250	0.5966	0.7978	0.7922
genito-urinary system	musculoskeletal	30	0.5232	0.6299	0.5964	0.7961	0.7921
musculoskeletal	skin complaints	38	0.5615	0.5963	0.6568	0.7921	0.8064
diabetes. incl. hyperglycemia	arthritis/rheumatism/fibrositis	34	0.5434	0.6351	0.5815	0.7895	0.7855
mental illness/anxiety/depression/nerves	arthritis/rheumatism/fibrositis	104	0.4368	0.5684	0.5837	0.8129	0.7851
mental illness/anxiety/depression/nerves	back problems/slipped disc/spine/neck	126	0.4778	0.5641	0.6049	0.8126	0.8043
mental illness/anxiety/depression/nerves	other problems of bones/joints/muscles	36	0.5093	0.5620	0.6099	0.8124	0.7920
other problems of nervous system	arthritis/rheumatism/fibrositis	49	0.5218	0.5700	0.5820	0.8012	0.7854
other problems of nervous system	back problems/slipped disc/spine/neck	30	0.5146	0.5693	0.6037	0.8005	0.8044
other problems of nervous system	other problems of bones/joints/muscles	158	0.5480	0.5676	0.6093	0.8008	0.7921
heart attack/angina	other heart problems	36	0.5437	0.6184	0.5951	0.7783	0.7793
heart attack/angina	arthritis/rheumatism/fibrositis	168	0.5405	0.6216	0.5816	0.7784	0.7857
hypertension/high blood pressure/blood p	arthritis/rheumatism/fibrositis	31	0.5576	0.6301	0.5814	0.7881	0.7857
other heart problems	arthritis/rheumatism/fibrositis	44	0.5077	0.6006	0.5824	0.7798	0.7859

other heart problems	other problems of bones/joints/muscles	51	0.5182	0.5966	0.6098	0.7795	0.7925
asthma	arthritis/rheumatism/fibrositis	113	0.5300	0.6567	0.5819	0.8034	0.7854
other respiratory complaints	arthritis/rheumatism/fibrositis	33	0.5247	0.5841	0.5815	0.7811	0.7854
arthritis/rheumatism/fibrositis	back problems/slipped disc/spine/neck	346	0.5465	0.5827	0.6068	0.7849	0.8052
arthritis/rheumatism/fibrositis	other problems of bones/joints/muscles	227	0.5538	0.5821	0.6122	0.7856	0.7928
arthritis/rheumatism/fibrositis	unclassifiable	102	0.6068	0.5809	0.6150	0.7857	0.7772
back problems/slipped disc/spine/neck	other problems of bones/joints/muscles	37	0.5730	0.6053	0.6122	0.8049	0.7913