The Convergent Validity of a Measure of the Quality of Life: Further Evidence

HESG Conference, Aberdeen University, July 14-16, 1999.

Douglas McCulloch
The National Centre of Pharmacoeconomics in Ireland,
St James's Hospital,
St. James's Street,
Dublin 8.
Telephone: 00 44 353 1 4531725;
Fax: 00 44 353 1 4730596

e-mail: dw.mcculloch@stjames.ie

SUMMARY

Quality of life data was collected from 105 patients who were waiting for, or who had had, invasive treatment for Ischaemic Heart Disease: the instruments were the Health Measurement Questionnaire and the Functional Limitations Profile. The resulting data set includes data collected before and after cardiac procedures, as well as a wide range of qol; this justifies the re-testing of the convergent validity hypotheses of Whynes and Neilsen[1], and makes possible the testing of further possibilities. The convergent validity of the Health Measurement Questionniare is again demonstrated, but the testing of medical expectations, and of a linear mdoel relating qol to age and sex, do not find convergent validity, and raise questions about the conventional methods for assessing the validity of QALY measures.

This is a draft paper; please do not quote from it without the author's permission

The Convergent Validity of a Measure of the Quality of Life: Further Evidence

The examination of evidence regarding the validity of a quality of life (qol) measure serves one of two purposes: either the validity of the measure is confirmed[1], or the basis of its validity is questioned, on grounds supported by evidence. The continuing importance of health outcome measures to health policy-makers[2] requires the testing of the validity of qol measures until either acceptable validity is achieved, for a particular measure, or the basis for the assessment of qol measures, itself, is re-examined.

Whynes and Neilson [W&N] define the project of qol measure assessment in terms of finding convergent validity between the Health Measurement Questionnaire [HMQ] and the Nottingham Health Profile [NHP][1]; in their work, convergent validity was found, and we develop the investigation of the HMQ's convergent validity with reference to three aspects:

1) Greater variation in qol scores
W&N report that the qol of the patients studied, measured by the mean scores on each of the NHP's dimensions, was not significantly different from the qol of the population (adjusted for the age and sex composition of the sample). A more severe test of the HMQ's convergent validity would be one which involved patients with greater variation in their qol scores. Given the well-recognised debilitating nature of heart disease[3], it seems clear that the mean qol of a group of cardiac patients, some of whom were waiting for, or had had, invasive treatment, is likely to be less than the mean qol of the patients in the W&N study, equivalent as this was to the age- and sex-adjusted NHP values for the population as a whole. Further, cardiac disease is severe in its effects on some patients, but its impact on others is less; this is indicated by the Canadian Cardiovascular Society's classification of angina[3]. It follows that some at least of a sample of cardiac patients will have a qol approaching "full health", and that the variation in qol values is likely to be larger, for a sample of cardiac patients, than for the patients in the W&N study. Accordingly, we may conclude that a sample of cardiac patients will offer a more rigorous test of convergent validity than the patients studied by W&N.

2) A Single Interval Scale, instead of a Set of Interval Scales
The purpose of the HMQ is to determine which Rosser health state applies to the respondent, so that the Rosser matrix health state value applying to the respondent can be determined. W&N tested the convergent validity of the HMQ with reference to the NHP, a set of interval scale measures on several separate dimensions, so that qol is measured using a set of scores, not a single overall score. The Functional Limitations Profile (FLP) is made up of 136 statements, similar in form to those of the NHP, which have been given weights on a single scale, so that the respondent accumulates a total score, made up of the number attached to each of the statements which (he or she affirms) is true of the present health state. This establishes patient qol on an interval scale, and in that sense the FLP is closer to the cardinal scale of the Rosser matrix than the NHP. Like the NHP, the FLP is a general health status measure, and is a suitable comparator for the Rosser HMQ questionnaire for similar

reasons. The FLP's original [United States] version, the Sickness Impact Profile[4,5] [SIP], has been successfully tested for reliability and validity[6-9], and has achieved widespread acceptability in use[10-15]. The FLP closely resembles the SIP, with minor modifications to take account of United Kingdom English expression; the testing of the FLP against the SIP[16] found them to be almost identical.

The use of the FLP simplifies the specification of the hypotheses required for convergent validity, because dimensions do not need to be considered separately; it also makes possible the measurement of changes in qol on a single scale.

3) Before and After Data

W&N's data came from the administration, on one occasion, of the HMQ and the NHP to discharged cancer patients; data from cardiac patients before and after their procedures permit the assessment of convergent validity using change correlations, linear models, and medical expectations, as described below.

METHOD

The original aims of the present investigation were to measure the qol impact, in QALYs, of the Routine Coronary Artery Bypass Graft [CABG], and to assess the validity of the Rosser/HMQ measure using the FLP. The unexpectedly small number of Routine CABGs during the project period required the extension of the enquiry to include Angioplasty and Emergency CABG patients, and the narrowing of its focus to the validation of the Rosser/HMQ measure. All of the patients were interviewed in their own homes, over a 12 month period; in cases where before and after data were collected, patients were interviewed up to eight weeks before the procedure, and six months afterwards. Waiting List patients, who did not get a procedure during the study, were interviewed between two and four times in all; in each of these cases, the last dataset collected was used in the analysis. The questionnaires cannot be used before the operation with Emergency CABG patients, because they are hospitalised; such patients were interviewed, once, in their homes, six months after the procedure.

The result was a database comprising FLP and Rosser HMQ data from 14 CABG Waiting List patients, who did not have the Routine CABG during the project (14 sets of data), 12 patients who did have the Routine CABG (24 sets of data, 12 before and 12 after the procedure), 31 patients who had an Angioplasty (62 sets of data, 31 before and 31 after), and 48 patients who had an Emergency CABG six months before interview (48 sets of data).

To avoid bias, we need to separate the before and after groups of Routine CABG and Angioplasty patients, so that they are not included twice in the analysis; it follows that we can test W&N's hypotheses with reference to two sets of data, Group A, comprising all those patients who had not (at the time of interview) had a procedure [14 (Waiting List) + 12 (pre-CABG) + 31 (pre-Angioplasty) = 57 datasets], and Group B, those patients who had had a procedure at the time of interview [12 (post-CABG) + 31 (post-Angioplasty) + 48 (post-Emergency CABG)

= 91 datasets]. Both of these groups included patients recording "zero limitations" on the Functional Limitations Profile, 5 in Group A (9%), and 14 in Group B (15%)

Adapting W&N's hypotheses to our context, we wish to apply the following criteria to our data:

a) there should exist a significant *negative* correlation between total FLP score and the Rosser matrix value assigned as a result of the HMQ responses,

b) there should exist a significant *positive* correlation between increasing weighted scores on the FLP and the level of each dimension (disability and distress) of the Rosser matrix;

c) the same correlations should be found (as in the previous study[1]) among the age and sex sub-groups of the data.

As additional tests of convergent validity, the following are proposed:

d) for those patients who had a procedure, there should exist a *negative* correlation between the *changes* in FLP score and the *changes* in Rosser value;

e) if there is a negative correlation between the total FLP score and the Rosser matrix value, it should also be found in each of the patient groups, there being no *prima facie* reason for supposing the groups to be different from the whole sample;
f) if medical expectations are correct, CABG waiting list and routine CABG patients should have significantly different qol values, and both the CABG and the Angioplasty should produce significant improvements in qol;

g) there should be a relationship between Rosser values, age and sex, in a linear model; this should also be present for transformed FLP scores, age and sex.

RESULTS
(In what follows, the 1% level of significance is used, unless otherwise stated).

Table 1 shows that, despite the (assumed) greater variation in qol among the cardiac patients, compared to the cancer patients of the W&N study[1], criteria a) and b) are again met. W&N's co-efficients, shown for comparison, relate the absolute number of NHP responses, as an overall measure of qol, to the relevant Rosser matrix variables.

Turning to criterion c), a breakdown of the data into cases above and below the median age, and between the sexes, offers a comparison with W&N's data[1], though the age and sex groups are not evenly distributed in the present study. The results are shown in Table 2.

Small sample sizes could be referred to as reasons for a lack of significance in all those cases where it occurs except one, that is, males under the median age show a non-significant co-efficient between Rosser Distress Level and Total FLP scores. It is argued here that this result is of subsidiary importance, because the Rosser and FLP measures are interval scales; the main feature is that there is indeed a significant correlation, for groups larger than 20, between the Rosser values and the Total FLP scores, as W&N also found.

Moving on to criterion d), we would expect that, in respect of patients whose qol was recorded before and after a procedure, there should exist a *negative* correlation between the *changes* in FLP score and the *changes* in Rosser value; as qol diminishes, these qol values move in opposite directions. There should also be a *positive* correlation between the changes in the Disability and Distress levels, and the change in the FLP score.

These data come from the Routine CABG and Angioplasty patients, numbering 43 in all. For the purpose of generating QALY values, one could argue that this is one of the more important aspects of the analysis, because ultimately, the intention is to compare the (marginal) cost per (marginal) QALY of different interventions. The correlations are shown in Table 3; all are significant, and add to the weight of the evidence for the convergent validity of the Rosser measure.

However, testing criterion e) produces mixed results. *Prima facie,* we have no reason to expect that the correlation co-efficients will not be significant for each procedure group; that is, we would expect that the group chosen would have no effect on significance, and that either all groups, or none, would be significant at the 1% level. Table 4 shows that all the co-efficients are significant except those for the Pre- and Post-Routine CABG groups; this is our first finding against convergent validity, though it is not a strong one, given the relatively small numbers of patients in the relevant groups (12) for which the co-efficient is not significant, and the fact that the same people are giving the responses associated with each non-significant result.

Medical Expectations

Given the condition and different treatments of the patients studied, we can frame medical expectations, namely,

1) *Waiting List and Pre-Routine CABG Patients*

Those who are selected for a CABG from a waiting list should be the patients who are the most ill; it follows that, on the average, the patients who [during the project] remained on the Routine CABG waiting list should have a higher average qol value than those who went forward for a CABG procedure.

2) *The impact of the CABG and PTCA on qol*

For routine CABG and PTCA patients, the mean qol value should be higher after the procedure than before.

These expectations are ad hoc, and are specific to the cardiac speciality, but they are widely held[3,18]; to our knowledge, similar tests have not previously been carried out for the Rosser measure, even in the context of other diseases.

The differences shown in Table 5, on their face, confirm the medical expectation specified; the question is, are these differences significant? We may apply a t-test only if the samples are taken from a distribution which is normal; this assumption is contradicted by Rosser and Kind (1974), though this judgement itself is based on a sample size of 70. In fact the three Rosser t-test results are not significant, though only one of the three FLP results is not significant.

Linear Modelling

While the project data collection was not designed for this purpose, the convergent validity of the Rosser QALY may also be tested by assuming there is a relationship between other variables and qol, and examining the Rosser data for evidence of that relationship; if appropriate, this evidence can be compared with results from FLP data, to determine whether the Rosser values show the relationship to the same extent.

As age increases, we would expect qol to fall, and people of different gender may be expected to have different health experiences; it follows that we would expect there to be some relationship between qol, age and sex, such that:

$$\text{Qol} = f(\text{Age, Sex})$$

which both the FLP scores and the Rosser values should reflect.

Given such a relationship, variations in the Rosser values and FLP scores in the data set should be explained, to some extent, by variations in Age and Sex; the criterion proposed is that the Rosser values will be validated if they achieve the same degree of explanation as the FLP scores.

Now, while the Rosser values are already based on a ratio scale, the FLP scores, as they stand, are from an interval scale; the linear model approach requires the transformation of the FLP data into data from a ratio scale, so that each datum comes from a scale with the same zero. For this to be true, we have to make the assumption that all the judges in the original FLP research who created the FLP by valuing its statements had in mind the same zero, from which the FLP total score is measured.

This assumption seems reasonable enough, though the zero has no empirical foundation; it amounts to assuming that the range of health state values is the same for all individuals. This assumption has been described as "unrealistic"[17]; however, the equal length of the QALY scale between persons can be defended as a merit of the QALY approach, assuming as it does that the same range of health states applies to all. In the literature, the same assumption is made by those using the FLP data in percentage terms[14,15].

The transformation necessary is to divide each FLP score by the maximum possible FLP score, 10154 (rounded to 10000 for ease of calculation), and subtracting the result from one, that is, calculating [1 - (x / 10000)], where "x" is the relevant FLP score, to create a new data set. This transformation is possible because of the interval scale of the FLP; obviously, it was not available to W&N, given their reliance on the NHP. Since there is no obvious reason why data from the PTCA and Routine CABG groups should show different relationships between Age, Sex, and qol, we combine these two groups in the analysis that follows, to give a combined sample size of (31+12 =) 43.

There are a number of possible ways we can relate qol to a linear relationship involving age and sex; since we found a statistically significant Spearman co-efficient in respect of changes in FLP total scores related to changes in Rosser values, we first relate changes in Rosser values and (transformed) changes in FLP scores, separately, to Age and Sex. We then relate qol *after* each procedure to Age, Sex, and qol beforehand, separately for the Rosser values and FLP scores.

The following models are therefore proposed:

Incremental Changes

$\Delta RQol = \alpha 1 + \alpha 2.Age + \alpha 3.Sex + \varepsilon$..(1)

$\Delta Fqol = \beta 1 + \beta 2.Age + \beta 3.Sex + \varepsilon$..(2)

Post-Procedure Health States

$RqolA = \alpha 1 + \alpha 2.Age + \alpha 3.Sex + \alpha 4RqolB + \varepsilon$....................(3)

$FqolA = \beta 1 + \beta 2.Age + \beta 3.Sex + \beta 4.FqolB + \varepsilon$......................(4)

where:

Δ: incremental change;

Rqol: qol measured in Rosser values;

Fqol: qol measured in FLP scores;

A: After a procedure;

B: Before a procedure.

What may we expect from the results of such estimations? It has to be remembered that the maximum value of both the qol variables is 1.0; if the constant term in the estimated equation is close to 1.0, the model will not explain much variation in the dependent variable. Also, the values of the equation co-efficients are adding to a total whose maximum value is 1.0; a co-efficient equal to (for example) 0.0296 would add nearly 3% for every unit change in the independent variable. Finally, we would expect the Age co-efficients to be negative; as age increases, we would expect qol to fall.

The next section presents the "Incremental Change" Models, for the CABG and PTCA respectively, and the following section reports the results, in the same format, from the "Post-Procedure Health State" Models.

The "Incremental Change" Models

The first question is, are Age and Sex related to the change in qol? Only if this is true for the FLP data can we test the Rosser matrix's convergence to the FLP; the results are shown in Table 6.

The R^2 co-efficients are low; neither Age nor Sex explains much of the variation in qol, however measured. If (as seems reasonable) the expected relationship between Age, Sex, and qol does exist, the data we have do not permit us to unmask it. If we had data regarding (for example) the length of time on the waiting list for a procedure, or the type of heart disease, we could include such variables in the equation. Some of this variation may be accounted for when we include "Health state before the procedure" in the equation, as a determinant of the health state after the procedures, because these other, so far unexamined, variables will have an influence on this additional variable. This is the basis of the model discussed in the next section.

The Post-Procedure Health State Models

The co-efficients for these models are shown in Table 7; the equations are:

$$RqolA = \alpha 1 + \alpha 2.Age + \alpha 3.Sex + \alpha 4RqolB + \varepsilon \quad (3)$$
$$FqolA = \beta 1 + \beta 2.Age + \beta 3.Sex + \beta 4.FqolB + \varepsilon \quad (4)$$

In respect of the FLP, we find the effect expected; QolB [quality of life Before the procedure] has a coefficient of 0.3046 in the model, and therefore is the variable most closely related to QolA [quality of life After the procedure]; the same applies for the Rosser data, though the overall R^2 is smaller. For both sets of data, we again do not find a strong relationship between qol, and age and sex, and we have been unable to assess the Rosser QALY's convergent validity by the means proposed.

These results suggest an interesting question: if qol variables are not linearly related to age or sex (especially to age), what does this mean for our definition of qol ? It seems unlikely that patients are deliberately adjusting to allow for the effects of age; FLP statements either apply to patients or they do not, and the factual kind of statements involved in the HMQ similarly do not seem to allow much room for bias of this kind.

However, the phenomenon may be a subtle one; as people age, their perceptions change, in ways of which they may be unaware, and their lifestyle (especially the type of accommodation) may be adapted to their more limited capabilities. If a person never needs to climb a flight of stairs, an inability to do so may not be apparent, or may be taken for granted, and the same applies to walking up or down hills, or routine social activity. Similarly, the statement emphasised at the start of the FLP questionnaire is "Because of my health"; do respondents separate the effects of their illness from the effects of old age? These complex matters merit further investigation, but they are outside the scope of the present study.

DISCUSSION

According to the same criteria as those adopted in the previous study[1], we found that the Rosser QALY has "convergent" validity with the Functional Limitations Profile. The linear modelling approach did not generate evidence for or against convergent validity, because the relationship expected between Age and Sex, and qol, was not found. Regarding medical expectations, the data showed the expected differences; the results were not statistically significant, though when the underlying distribution is not normal, the wider meaning of these findings is hard to determine. Overall, the picture is unclear, and the issue as to whether the Rosser measure has convergent validity remains unresolved.

One useful conclusion may be that the Spearman correlation co-efficient is not a rigorous enough test of QALY validity, even when we are testing for convergent rather than criterion validity. The rank-ordering of health cases means taking decisions about who shall live (longer), and who shall suffer (less); only a Spearman co-efficient of 1.00 would mean identical orderings of patients, and therefore identical sets of opportunities for life- and hrqol-enhancing treatments, for the patients involved.

It is hard to see that a Spearman co-efficient of 1.00 would be possible in a study of this kind, given that the FLP and Rosser measures are designed differently, from different premisses and contexts, and therefore are entirely likely to differ in their rankings, even if respondents are 100% accurate in their replies. The Spearman test for significance works by comparing the correlation found with what could have been expected to result from chance; it might be more appropriate to compare the correlation with complete agreement plus a margin of error - for instance, "our results are consistent with complete agreement +/- 5% average error in the responses". What distribution might we expect for such an error term?

When health priorities are being decided, it is not the distance from randomness that is important, but the distance from complete agreement; in this area, statistical analysis has the potential to make the difference between more and less pain, or between longer or shorter lives, and we should demand higher standards of statistical tests than are available from the Spearman statistics.

CONCLUSION

Our results indicate that the Rosser/HMQ QALY measure has convergent validity when tested against the FLP, in the same manner as the previous study. However, convergent validity was not found using medical expectations or linear models; in the former case the changes were in the expected direction, though not statistically significant. When a QALY is being tested, it seems appropriate to use an interval scale measure, and to apply tests which rely on this property to achieve "convergent" validity. The continuing search for a "valid" QALY may require either further testing against an interval scale measure, or a more rigorous test of ranking than the Spearman correlation coefficient provides.

REFERENCES

1. Whynes, D. and Neilson, A.R. Convergent Validity of Two Measures of the Quality of Life *Health Economics* 1993; **2**(3): 229-235.

2. Department of Health *The New NHS: Modern, Dependable* HMSO, Cmnd 3807, 1997.

3. Gunnell, D. and Smith, L. *The Invasive Management of Ischaemic Heart Disease*. Bristol: Health Care Evaluation Unit, University of Bristol 1994

4. Bergner M., Bobbitt, R.A., Kressel, S., *et al.* The Sickness Impact Profile: conceptual formulation and methodology for the development of a health status measure *International Journal of Health Services* 1976 **6:** 393-415

5. Gilson, B.S., Gilson, J.S., Bergner, M., *et al.* The Sickness Impact Profile: Development of an outcome measure of health care *Am. Jou. of Public Health* 1975; 1304-1310

6. Bergner M., Bobbitt, R.A., Pollard, W.E., et al The Sickness Impact Profile: validation of a health status measure *Medical Care* 1976; **14:** 57-67

7. Carter, W.B., Bobbitt, R., Bergner, M., Gilson, B.S. Validation of an Interval Scaling: the Sickness Impact Profile *Health Services Research* 1976; 516-528

8. Pollard, W.E., Bobbitt, R.A., Bergner, M., *et al.* The Sickness Impact Profile: Reliability of a Health Status Measure *Medical Care* 1976; **XIV**(2):146-155

9. Pollard, W.E., Bobbitt, R.A., Bergner, R.M. Examination of variable errors of measurement in a survey-based social indicator *Social Indicators Research* 1978 **5**: 279-301

10. Bergner, M., Bobbitt, R.A., Carter, W.B., and Gilson, B.S. The Sickness Impact Profile: development and final revision of a health status measure *Medical Care* 1981 **21**: 787-805

11. de Bruin, A.F., de Witte, L.P., Stevens, F., and Diedericks, J.P.M. Sickness Impact Profile: The State of the Art of a Generic Functional Status Measure *Social Science and Medicine* 1992 **35**(8): 1003-1014

12. Ott, C.R., Sivarajan E.S., Newton, K.M., et al A controlled randomised study of early cardiac rehabilitation: the Sickness Impact Profile as an assessment tool *Heart and Lung* 1983; **12**(2)162-70

13. Patrick, D.L., and Erickson, P. *Health Status and Health Policy* Oxford University Press: New York, 1993

14. Fletcher, A., Hunt, B., Bulpitt, C.J. Evaluation of Quality of Life in Clinical Trials of Cardiovascular Disease *Journal of Chronic Disease* 1987 **40**(6)

15. Fletcher, A., McLoone, P., Bulpitt, C. Quality of Life on Angina Therapy: a randomised controlled trial of transdermal blyceryl trinitrate against placebo *The Lancet* 1988: 216-230

16. Patrick, D.L. Standardisation of comparative health status measures: using scales developed in America in an English-speaking country In *Health Survey Research Methods 3rd Biennial conference*, US Dept of Health and Human Services: Hyattsville, Maryland, Publication no. PHS 81-236 1981

17. Broome, J. QALYs *Journal of Public Economics* 1993 **50.2**

18. Rushworth, R.L., Rob, M.I., Rubin, G.L. Assessing utilisation of health technologies: coronary artery bypass graft surgery in New South Wales *Australian Journal of Public Health* 1994; **18**(1): 101-105

Table 1. Spearman Rank Correlation Coefficients between total FLP score, and Disability level, Distress level, and Rosser matrix value, for Groups A and B, and for W&N data.

	Disability Level	Distress Level	Rosser Value
Group A (n=57)	0.62	0.43	-0.56
Group B (n=91)	0.54	0.60	-0.65
W&N (n=221)	0.68	0.54	-0.71

Table 2 Spearman correlation coefficients with respect to Total FLP scores, for four sub-samples

Group A (n=57)	Disability Level	Distress Level	Rosser Value	n =
Male < 58 years	0.57	0.40*	-0.57	22
Male = or > 58 years	0.69	0.53	-0.69	23
Female <58 years	0.80*	0.90*	-0.80*	4
Female = or > 58 years	0.49*	0.39*	-0.65*	8
Group B (n=91)				
Male <59 years	0.43	0.47	-0.50	36
Male = or >59 years	0.56	0.72	-0.74	39
Female <59 years	0.61*	0.42*	0.69*	7
Female = or >59 years	0.71*	0.60*	-0.66*	9
W&N				
Male <71 years	0.52	0.62	-0.61	50
Male >71 years	0.75	0.62	-0.76	55
Female <71 years	0.68	0.50	-0.67	59
Female >71 years	0.69	0.43	-0.69	56

"*" : Correlation co-efficient which is not significant at the 1% level.

Table 3: Spearman co-efficients for changes in FLP scores related to changes in Rosser Distress and Disability Levels, and to Rosser values, from patients undergoing a Routine CABG or PTCA.

	Change in Disability Level	Change in Distress Level	Rosser Value
Change in FLP Score (n=43)	0.55	0.38	-0.54

Table 4: Spearman co-efficients between FLP total score and Rosser values, for procedure sub-groups.

Patient Group	**Co-efficient**	**Significant at 1% level?**	**Number of cases**
CABG Waiting List Patients	-0.699	Yes	14
Pre-Routine CABG	-0.018	No	12
Post-Routine CABG	-0.579	No	12
Pre-PTCA	-0.656	Yes	31
Post-PTCA	-0.481	Yes	31
Emergency Patients	-0.717	Yes	48

Table 5: Qol of - Waiting List and Routine CABG patients, Before and After a Routine CABG, and Before and After an Angioplasty

	Mean Rosser Value	Mean FLP Score
Waiting List	0.965	1637.3
Pre-Routine CABG	0.861	2593.2
Difference	0.104	965.9
Pre-Routine CABG	0.861	2593.2
Post-Routine CABG	0.986	789.3
Difference	0.125	1803.9
Pre-Angioplasty	0.957	1227.0
Post-Angioplasty	0.986	588.9
Difference	0.029	638.1

Table 6: Incremental Change Model Co-efficients, for (transformed) FLP data, and Rosser data, from the combined PTCA and Routine CABG patient groups.

	FLP	Rosser
n	43	43
R^2	0.041	0.018
Constant	0.154	-0.127
Constant t-value	1.382	-0.788
Age co-efficient	-0.0018	0.0019
Age t-value	-0.997	0.714 (S)
Sex co-efficient	0.0389	0.963
Sex t-value	-0.0313	-0.538

[S = Significant; unless otherwise stated, t-values are not statistically significant.]

Table 7: Post-Procedure Health State Model Co-efficients

	FLP	Rosser
n	43	43
R^2	0.412	0.0990
Constant	-0.0279	0.9768
Constant t-value	-0.508	56.185 (S)
Age	0.0013	-2.360E-04
Age t-value	1.412 (S)	-1.072
Sex	-0.0234	0.0076
Sex t-value	-1.200	1.569 (S)
QolB	0.3046	0.0155
QolB t-value	5.052 (S)	1.200 (S)