Measuring outcome of total knee replacement using quality of life indices

Robert F Drewett DPhil

Lecturer

Department of Psychology, University of Durham

R Julian Minns PhD

Bioengineer

T Frank Sibly FRCS

Research Fellow in Orthopaedics

Dryburn Hospital, Durham

Key words: Knee prosthesis; Quality of life

A series of 26 patients was studied before and after total knee replacement to determine the change in their quality of life, expressed as a gain in Quality Adjusted Life Years (QALYs). Global health status was assessed using the Nottingham Health Profile, disability by the Harris scale, pain by the McGill Pain Questionnaire, and anxiety and depression by the Hospital Anxiety and Depression Scale. Substantial reductions were found in pain, anxiety and depression, and a significant improvement was found in mobility.

These data were used to generate a 'QALY' (Quality Adjusted Life Year) comparable to the measure used by Williams (1) in his comparative evaluation of medical and surgical treatments. A gain of only 0.42 QALY was found for knee replacement, which is about one-tenth the figure given by Williams for hip replacement. The difference appears to be attributable not to differences in the success of the operations, but rather to the scope that QALY calculations give for allocating comparable patients to different quality of life states. Tighter criteria are needed if QALYs are to be used to guide resource allocation.

Since economic resources are limited, it is important to determine which medical and surgical procedures give the greatest benefit for the least cost. To do this it is necessary to have a universal method of patient assessment, which can be applied to any form of ill health. The best known measure in Britain is the Quality Adjusted Life Year, or QALY (1). It adjusts life expectancy for the quality of life, which depends, for example, on mobility, ability to care for oneself, and absence of pain. The

Correspondence to: Mr T F Sibly, Department Orthopaedics, Derbyshire Royal Infirmary, Derby DE1 2QY

QALY gain of patients undergoing a variety of procedures has been calculated by Williams (1), basing his QALY calculations on the published clinical results of others (Table I).

Hip replacement is widely regarded as a successful and cost-effective operation. This view is supported in the QALY calculations of Williams (1), who states that 'of all treatments examined so far, hip replacement comes near the top of the league . . .'. The QALY approach is well suited to the evaluation of joint replacement since the main benefits of the operation are on mobility and pain, and these correspond simply to the two dimensions of disability and distress in the classification of Rosser and Watts (2) on which QALYs are based. In this paper we have measured prospectively the success of total knee replacement in the same terms. This operation came into use in the early 1970s and is now regarded as similarly successful and cost-effective.

Table I. Benefit of various procedures when measured in QALYs

	QALYs gained	Cost in £000s	Cost (£000s) per QALY gained
Coronary artery bypass graft	2.75	3	1.04
Heart transplant	4.5	23	5
Kidney transplant	5	15	3
Haemodialysis	5	66	13
Hip replacement	4	3	0.8

From Williams (1)

Measures for use in evaluating Quality of Life can be derived from questionnaires specifically designed to generate QALYs, such as the self-completion questionnaire given in Gudex and Kind *(3)*, or condition specific scales, such as the Arthritis Impact Measurement Scales *(4)*, or scales designed for more general use, such as the McGill Pain Questionnaire *(5)*. We have used the third type of measure. There is a large amount of data available on the performance of scales for pain, anxiety and depression, and they are already in common use for clinical and research purposes. It seems to us desirable as far as possible to capitalise on this, rather than to develop new scales *ad hoc* to assess the same states.

Patients and methods

Patients undergoing total knee replacement were interviewed before surgery and at least 2 months after operation. In nearly all cases interviews were conducted in the patient's home. Data were collected on 26 subjects, 5 male and 21 female. Age at operation ranged from 49 years to 84 years with a mean of 72 years (69 years for the men and 73 years for the women).

Assessment was carried out with one global health scale (the Nottingham Health Profile *(6)* and three symptom-specific scales. The NHP comprises six subscales (physical mobility, pain, social isolation, emotional reactions, energy and sleep) each weighted to give a score out of 100, with greater ill-health represented by a higher score. The McGill Pain Questionnaire *(5)* was used to assess pain, with a 1-week recall period as in Burckhardt *(7)*. Rating of the worst pain was requested. The Present Pain Intensity Scale of the McGill Pain Questionnaire returns scores from 1 (mild pain) to 5 (excruciating pain). Complete absence of pain scored as 0. Disability was assessed using the self-care items from the interview schedule of Harris *(8)* (appendix G, question 26, 1-29). A score from 1 to 8 (9 for men) gives the number of different areas of self-care that the respondent has difficulty with. Data on anxiety and depression were collected using the Hospital Anxiety and Depression (HAD) scale *(9)*, which returns scores of 0 to 21 for each. Symptoms of anxiety and depression overlap with symptoms of physical illness, so that people who are physically ill tend to have high scores on, for example, the General Health Questionnaire *(10)*. The HAD scales were developed for use with people who are physically ill, and concentrate on the psychic rather than the somatic symptoms of anxiety and depression.

Results

Scores before and after operation were compared (Table II) and the significance of differences was tested using the Wilcoxon test *(11)*. Significant improvements were found on the Present Pain Intensity Scale of the McGill (median 3 'distressing' before operation reduced to 0 'no pain' after the operation; $Z = -4.1$; $P<0.001$), and on the

Table II. Median values of outcome measures before and after total knee replacement, together with the statistical significance of the difference (P value)

	Maximum score of scale	Preop.	Postop.	P
Pain (McGill)	5	3	0	<0.001
Immobility (NHP)	100	54.6	27.6	<0.001
Self-care problems (Harris)	8	7	5.5	<0.001
Anxiety (HADS)	21	8.5	3.5	<0.01
Depression (HADS)	21	5.5	3.0	<0.01

100-point pain scale of the Nottingham Health Profile (medians 66.4 before surgery, 11.2 afterwards; $Z = -4.3$; $P<0.001$). Physical mobility improved on the Harris scale (median number of areas giving difficulty 7 before and 5.5 after operation; $Z = -3.4$; $P<0.001$); and the Nottingham Health Profile (54.6 before surgery, 27.6 afterwards; $Z = -3.8$; $P<0.001$). There were significant reductions in both anxiety and depression scores on the HAD scale (8.5 to 3.5 and 5.5 to 3.0 for anxiety and depression respectively; $Z = -2.75$. $P<0.01$ for each).

'Emotional reactions' were also significantly reduced on the Nottingham Health Profile (22.3 to 3.5; $Z = -3.66$; $P<0.001$). Results in terms of Quality of Life measures were very similar for patients with osteoarthritis and those with rheumatoid arthritis. Although some of these patients had multiple joint arthritis, none were on the waiting list for further joint surgery, so no further improvement in their Quality of Life can be anticipated.

To estimate QALYs, the McGill and HAD scales were rescaled to give scores from 0 to 100, and then used to allocate the patient to a Rosser distress category following the procedure in Gudex and Kind *(3)*. Data from the Nottingham Health Profile and the Harris scale were used to allocate each subject to one of the eight Rosser disability categories. The category definitions are shown in Table III; the allocation of patients to the various cells in the distress/disability matrix is shown in Table IV.

The QoL (quality of life) scores before and after operation were then calculated as the sum of the products of the numbers in each disability/distress cell and the evaluation coefficients for that cell *(12)*. Quality of Life was found to be 0.910 before and 0.974 after operation.

The mean age of the patients in the series was 69 years (men) and 73 years (women). The expectation of life of men aged 70 years is 10.3 years; of women aged 75 years, 10.4 years *(13)*. As it is probably slightly reduced in patients with rheumatoid arthritis, a figure of 10 years is resonable for both sexes. Undiscounted, and with no complications taken into account, the QALYs gained are derived by multiplying the life expectancy by the gain in quality of life. This gives a figure of 0.64 [ie $10 \times (0.974 - 0.910)$]. Discounting at 5% per year is suggested by Williams to give greater impact to treatment for the first

Table III. Rosser's classification of illness states

Disability

    I No disability

    II Slight social disability

    III Severe social disability and/or slight impairment of performance at work

    IV Choice of work or performance at work severely limited
Housewives and old people able to do light housework only, but able to go shopping

    V Unable to undertake any paid employment
Unable to continue any education
Old people confined to home except for escorted outings and short walks and unable to do shopping

    VI Confined to chair or wheelchair or able to move around in the house only with support from an assistant

    VII Confined to bed

    VIII Unconscious

Distress

    A No distress

    B Mild

    C Moderate

    D Severe

Table IV. Location of subjects before and after operation on the Rosser disability/distress matrix. Number postoperatively is shown in brackets

	Distress				
	A	B	C	D	Totals
Disability					
I	(3)				(3)
II	(6)	5		1	6 (6)
III	(8)	6 (4)		2	8 (12)
IV				1	1
V	1 (1)	5 (3)	1 (1)	4	12 (5)
VI					
VII					
VIII					
Totals	1 (18)	16 (7)	1 (1)	8	

few years *(1)*. This reduces the QALY gain to 0.49 ($7.722 \times (0.974 - 0.910)$).

A further reduction is needed for an anticipated peroperative mortality of 1%. One patient in a hundred will die, and hence lose all of the 7.03 QALYs the patient would have expected to enjoy (7.722×0.910). We must therefore reduce the gain in QALYs for the whole group by 0.07 (ie $1\% \times 7.03$). This leaves an overall gain in QALYs of 0.42. The service cost of the operation in this district is £3280, giving a figure of £7810 per QALY gained.

Discussion

The improvement in patients' mobility and their reduction in pain, anxiety and depression were all highly significant, despite the relatively small numbers. However, when this improvement is expressed in terms of QALYs gained, the result is disappointing (0.42 QALYs gained).

This is about one-tenth of the figure quoted for hip replacement by Williams *(1)*. Even though his calculations concerned hip replacement and ours concerned knee replacement, we found the size of this difference surprising.

Differences in QALYs gained for the two operations come from various sources:

1 *Life expectancy of the patients.* The patients in our series were slightly older than those in Wilcock's paper *(14)*, on which Williams' calculations were based (72 years rather than 68 years). Assuming that the same life expectancy were used, this difference in age at operation would give a weight of 9.899 rather than 7.722 for the multiplier.

2 Adjustment is necessary in principle for failure of the prosthesis. The failure rate depends on time since implantation and ranges from 0.65% per annum *(15)* to 1.8% per annum *(16)*. Because of this wide range we have not adjusted for failure, and the adjustment for this in Williams *(1)* is made in the cost of the operation rather than in the Quality of Life estimation.

3 *Differences in the calculation of the QoL.* The data given in Wilcock *(14)* upon which the calculations in Williams were based did not include measures of depression and anxiety, but these changed the allocations we made only slightly from those that would have been given on the basis of the pain scale alone. Postoperatively, three patients were classified in distress level B who would have been classified in distress level A on their pain scores alone.

Table V. Valuations for the 29 health states

| | *Distress* | | | |
	A	B	C	D
Disability				
I	1.000	0.995	0.990	0.967
II	0.990	0.986	0.973	0.932
III	0.980	0.972	0.956	0.912
IV	0.964	0.956	0.942	0.870
V	0.946	0.935	0.900	0.700
VI	0.875	0.845	0.680	0.000
VII	0.677	0.564	0.000	−1.486
VIII	−1.028	—	—	—

1 = Healthy; 0 = Dead

The main difference between the QALY gain seen in total knee replacement and the score given by Williams for hip replacement appears to result from none of these. The Rosser category to which patients were allocated preoperatively is quite different between the two series. None of our cases were allocated to a disability level greater than V, and discounted at 5% over a period of 10 years the maximum QALY gain per patient would therefore be 2.317 (or discounted over a period of 14 years, 2.970). In order to arrive at an average gain of 4 overall, substantial numbers of patients in the hip study must have been allocated preoperatively to Rosser disability VI (confined to chair or wheelchair or able to move about the house only with support from an assistant) or VII (confined to bed). We suggest that the series used may have been unrepresentative of the disability suffered by patients before undergoing hip or knee joint replacement.

The criteria for allocation of cases to categories are ill-defined, and must be made more consistent, especially where it matters most. Within those health states valued over 0.9 (I–III A to D, and IV–V A to C in Table V) misallocation to an adjacent category alters the quality of life valuation by less than 0.042. Once the more severe states are involved, misallocation to an adjacent category can change the quality of life score by as much as 1.705. It would be beneficial to reduce the number of disability/distress categories, and to define them more closely.

We conclude that a figure of 0.42 quality adjusted life years is a realistic estimate of the gain accrued from knee replacement. Although the QALY gain is not great, it still compares favourably with other procedures on a cost per QALY basis.

References

1 Williams A. Economics of coronary artery bypassing. *Br Med J* 1985;291:326–9.
2 Rosser RM, Watts VC. The measurement of hospital output. *Int J Epidemiol* 1972;1:361–8.
3 Gudex C, Kind P. The QALY toolkit. University of York Centre for Health Economics. Discussion Paper 38, 1986:1–36.
4 Meenan RF, Gertman PM, Mason JH. Measuring health status in arthritis. *Arthritis Rheum* 1980;23:146–52.
5 Melzack R. The McGill pain questionnaire: major properties and scoring methods. *Pain* 1975;1:227–9.
6 Hunt SM, McKenna SP, Williams J. Reliability of a population survey tool for measuring perceived health problems: a study of patients with osteoarthritis. *J Epidemiol Community Health* 1981;35:297–340.
7 Burckhardt CS. The use of the McGill pain questionnaire in assessing arthritis pain. *Pain* 1984;19:305–14.
8 Harris AI. *Handicapped and Impaired in Great Britain.* Part 1. London: HMSO, 1971;1:1–330.
9 Zigmond AS, Snaith RP. A Hospital Anxiety and Depression scale. *Acta Psychiatr Scand* 1983;67:361–70.
10 Goldberg D, Williams P. *A User's Guide to the General Health Questionnaire.* Windsor: NFER-Nelson, 1988:1–129.
11 Siegal S. *Non-parametric Statistics for the Behavioral Sciences.* Tokyo: McGraw-Hill–Kogakusha Ltd, 1956:1–312.
12 Kind P, Rosser R, Williams A. Valuation of quality of life: some psychometric evidence. In: Jones-Lee MW ed. *The Value of Life and Safety.* Amsterdam: Elsevier North-Holland, 1982:159–70.
13 Griffin T, Church J eds. *Social Trends 18.* London: HMSO, 1988;114.
14 Wilcock GK. Benefits of total hip replacement to older patients and the community. *Br Med J* 1978;277:37–9.
15 Scuderi GR, Insall JN, Windsor RE, Moran MC. Survivorship of cemented total knee replacements. *J Bone Joint Surg* 1989;71B:798–803.
16 Tew M, Waugh W, Forster IW. Comparing the results of different types of knee replacement: a method proposed and applied. *J Bone Joint Surg* 1985;67B:775–9.

Invited comment

It is pleasing to see that an attempt is being made to prospectively evaluate the benefit to the patient of what has now become a commonly performed but relatively expensive procedure. It is a pity that the method of assessment needs to be so cumbersome, though the need to use a system which allows comparison with other