EVALUATION OF NEW TECHNOLOGIES FOR TOTAL HIP REPLACEMENT

ECONOMIC MODELLING AND CLINICAL TRIALS

W. J. GILLESPIE, B. PEKARSKY, D. L. O'CONNELL

From the University of Newcastle, New South Wales, Australia

Cost is a factor in the choice of prosthetic components in joint replacement. For a given performance, the least expensive components are the most cost-effective. When evaluating a new prosthesis with an unknown outcome, the use of an economic model allows estimation of potential cost-effectiveness. We used published data for the survival of cemented total hip replacements from Sweden, and cost and demographic information from New South Wales, Australia, in such a model.

In young active total hip recipients a new prosthetic design which offered a 90% improvement in survivorship over 15 years and a 15% reduction in the cost of revision surgery, could be sold at a price of 2 to 2.5 times that of conventional cemented components such as the Charnley Low Friction Arthroplasty and still be cost-effective. Using more likely estimates of the improved performance of new technology, however, the upper limit of cost-effectiveness is an increase of 1.5 to 1. Only a very small increase in the cost of a prosthesis could ever be justified for older patients of either sex. Most of the potential benefits of a better level of survivorship appear towards the end of the 15-year period. The results of modelling may be incorporated in clinical trial design.

Given the known performance of some well-established and relatively inexpensive designs of prostheses, very large randomised studies would be required to prove an improvement in performance.

J Bone Joint Surg [Br] 1995;77-B:528-33.

Received 6 April 1994; Accepted after revision 18 August 1994

W. J. Gillespie, ChM, FRCS Ed, FRACS, Professor of Orthopaedic Surgery
Clinical Research Unit, Department of Orthopaedic Surgery, Princess Margaret Rose Orthopaedic Hospital, Fairmilehead, Edinburgh EH10 7ED, UK.

B. Pekarsky, BEc, DipHEc, Associate Lecturer in Health Economics
D. L. O'Connell, BMaths, PhD, Senior Lecturer in Biostatistics
Centre for Clinical Epidemiology and Biostatistics, Faculty of Medicine, University of Newcastle, Callaghan, New South Wales 2308, Australia.

Correspondence should be sent to Professor W. J. Gillespie.

©1995 British Editorial Society of Bone and Joint Surgery
0301-620X/95/4959 \$2.00

Information on long-term survival indicates that cemented total hip replacement (THR) is a very effective procedure, particularly in the elderly patients who constitute most of its recipients (Ahnfelt et al 1990). Many new implant designs have been introduced. Most have avoided comparative evaluation of outcome by randomised controlled trials by application of the principle of 'substantial equivalence' (Faro and Huiskes 1992) and have been put into use in the expectation of improved results. Although based on extensive biological and biomechanical research their expectations may not be fulfilled, as illustrated by the history of surface replacement of the hip in the early 1980s (Dorey and Amstutz 1989). The price of many newer designs is over three times that of proven cemented components, yet the proposed benefits remain unsubstantiated.

Surgeons should therefore consider carefully what the individual patient and society at large are receiving for their investment, and whether the additional costs of the prosthesis are justified (Bulstrode et al 1993). Perhaps the new technology should be withheld from general use until outcome data, including analysis of cost-effectiveness, are available from comparative studies, as is now the case for new drugs in many countries (Freemantle et al 1995). Another approach is to model cost-effectiveness, comparing the hypothetical performance with direct observation of existing prosthetic designs.

We have used a model described by Daellenbach et al (1990) to examine long-term outcome data, and the choices which orthopaedic surgeons and others must make.

METHOD

The model assumes that the clinical result of a THR remains acceptable until revision is required and that increased expenditure on the prosthesis is only justifiable if accompanied by a proportional improvement in survival. The possible benefits of a new THR include reduced costs of implantation (including prosthetic and all other hospital costs), longer clinical survival and easier, less expensive revision surgery. In this simple form of the model, equivalence of hospital costs other than that of the prosthesis is assumed. The Present Value (PV) of the implantation of a prosthesis of known cost and survivorship is compared with that of a new prosthesis, for which the pricing indicates extra cost in the current period. Estimates of future savings

due to increased survival and easier revision are modelled. The probability of requiring reoperation depends upon the age, gender, and life expectancy of the patient. This allows estimation of the additional prosthetic costs which would be justified were any given improvement in performance actually achieved. The model also indicates the effects on health costs of different choices of prosthesis made by surgeons treating a population of defined characteristics. 'Cross-over' ages may also be calculated; these are the ages below which a given increase in prosthetic expenditure would be justified (Daellenbach et al 1990).

For this particular application of the model, the PV of conventional cemented total hip arthroplasty (PV_C) was calculated for men and women in the age categories < 55, 55 to 64, 65 to 74 and over 75 years, using the survivorship data presented in the Swedish Hip Review study (Ahnfelt et al 1990) from a cross-section of university and community hospitals, which are assumed to be equivalent to those achieved using the same prostheses in Australia. The theoretical PV (PV_E) was calculated for prostheses of defined cost, but undetermined outcome, using a range of assumptions of longer trouble-free prosthetic life and easier, cheaper revision. Estimates of life expectancy were drawn from life tables published by the Australian Bureau of Statistics. The pattern of total hip replacement by age and gender and the total hospital costs of primary prosthetic arthroplasty and of revision arthroplasty were taken from the statistics of the New South Wales Health Department for 1989. PV was calculated according to the following formula:

$$PV_C = C_c + H + \sum_{i=3}^n \frac{l_{im} p_{cim} (C_c + R_c + H)}{(1+r)^{(i-1)}}$$

where PV_C = expected present value of cost of treating a patient with prosthesis C over n years; C_c = cost of pro-

thesis C; H = hospital costs of first operation; l_{im} = probability of individual in age group m being alive at year i ; p_{cim} = probability that reoperation will be required between years $i-2$ and i if individual of age group m has prosthesis type C; R_c = additional costs of reoperation over and above H if original prosthesis type is C; and $(1+r)^{(i-1)}$ = discount rate for cash flows at year i .

For a new prosthetic system to be cost-effective, the PVs must be equal or the new system must have a lower value ($PV_C > PV_E$). PV represents the present-day value of the intervention, but depends on events which have yet to happen. It is calculated from both costs expected to occur over a given period into the future, and those which are certain in the current period. The time periods which we have chosen are 5, 9 and 15 years. The question to be answered is that if we decide today to do A or B, and measure cost-effectiveness in X years time, which of the two options can we expect to have cost less for the same outcome, or have delivered a better outcome for the same cost.

RESULTS

Break-even prosthetic price ratios. Tables I and II show by gender, the prosthetic price ratios which achieve the break-even point for options C and E at 5, 9 and 15 years, and are derived by using each of four hypothetical rates of prosthetic failure and two assumptions about the cost of revision surgery. We modelled equivalent failure rates ($F_E = F_C$) and three possible outcomes in which the failure rate of the new technology was 80%, 50% or 10% of that of the conventional at each point of measurement over 15 years (i.e., $F_E:F_C = 0.8, 0.5$ and 0.1). We assumed that the costs of revision were equal ($R_E = R_C$) or that the cost of revision surgery was reduced by \$1500, approximately

Table I. Break-even analysis for prosthetic price ratio C_E/C_C^* of total hip replacement in females

Age group (yr)	Proportional reduction of failure rate (%)	Reoperation costs equal ($R_E=R_C$) at			Costs \$1500 less after prosthesis ($R_E=R_C - \$1500$) at		
		5 years	9 years	15 years	5 years	9 years	15 years
< 55	0	-	-	-	1.0	1.1	1.2
	20	1.0	1.1	1.2	1.1	1.2	1.4
	50	1.1	1.4	1.6	1.1	1.4	1.7
	90	1.2	1.7	2.1	1.2	1.7	2.2
55 to 64	0	-	-	-	1.0	1.1	1.2
	20	1.0	1.1	1.2	1.1	1.2	1.3
	50	1.1	1.2	1.4	1.1	1.3	1.5
	90	1.1	1.5	1.8	1.1	1.5	1.8
65 to 74	0	-	-	-	1.0	1.1	1.2
	20	1.0	1.1	1.2	1.1	1.2	1.3
	50	1.1	1.3	1.4	1.1	1.3	1.5
	90	1.2	1.5	1.8	1.2	1.5	1.8
> 75	0	-	-	-	1.0	1.0	1.0
	20	1.0	1.0	1.0	1.0	1.0	1.1
	50	1.0	1.1	1.1	1.0	1.1	1.1
	90	1.1	1.1	1.2	1.1	1.1	1.2

* break-even cost is the price ratio (experimental prosthesis/conventional prosthesis; $C_E = C_C$) which would prove cost-effective at the end of a period if prosthesis E were to perform more effectively as measured by rate of revision during that period

Table II. Break-even analysis for prosthetic price ratio C_E/C_C * of total hip replacement in males

Age group (yr)	Proportional reduction of failure rate (%)	Reoperation costs equal ($RE=RC$) at			Costs \$1500 less after prosthesis ($RE=RC - \$1500$) at		
		5 years	9 years	15 years	5 years	9 years	15 years
< 55	0	-	-	-	1.1	1.1	1.3
	20	1.1	1.1	1.3	1.1	1.3	1.5
	50	1.2	1.4	1.7	1.2	1.5	1.9
	90	1.3	1.7	2.3	1.3	1.8	2.4
55 to 64	0	-	-	-	1.1	1.2	1.3
	20	1.1	1.2	1.3	1.1	1.3	1.5
	50	1.2	1.4	1.8	1.2	1.5	2.0
	90	1.4	1.8	2.5	1.4	1.8	2.5
65 to 74	0	-	-	-	1.0	1.1	1.1
	20	1.0	1.1	1.1	1.0	1.2	1.3
	50	1.1	1.2	1.4	1.1	1.3	1.5
	90	1.1	1.5	1.7	1.1	1.5	1.7
> 75	0	-	-	-	1.0	1.0	1.0
	20	1.0	1.0	1.0	1.0	1.1	1.1
	50	1.1	1.1	1.1	1.1	1.1	1.1
	90	1.1	1.2	1.2	1.1	1.2	1.2

* see Table I

15% of the total, if the new technology had been used for the primary replacement ($R_E = R_C - \$1500$). The ratio expresses the price C_E which could be justified. Over periods of 5 and 9 years only modest increases in price ratio would prove cost-effective and over 15 years price ratios of over 2 to 1 would be acceptable if the revision rate was reduced by 90%. Significant reduction in the cost of revision, however, is unlikely, and a fall in the present failure rate of more than 50% is improbable using any components for which early results are available. Thus, only price rises of the order of 70% over those of good, established 'first-generation' components are likely to prove cost-effective in the younger age groups. In older patients, only very small increases in expenditure on prosthetic components could ever be justified, which is what would be expected in the context of shorter life expectancy and high survivorship of the replacement.

Effect of varying prosthetic cost on the output of an orthopaedic unit. In Table III it is assumed that a new

prosthesis E is available at the same price as prosthesis C but has an improved survival profile. For each of i years 100 first operations of present value PV_{Cfi} have been performed after which a number of reoperations, dependent on the published survival profile for the age group for prosthesis C, of present value PV_{Cfi} , will be necessary. Had prosthesis E been used, its improved performance would have reduced the number of revisions necessary in the period, allowing more first operations for the same overall expenditure. The benefits from additional first operations would be significant.

The effect of introducing a new prosthesis E available at twice the price of prosthesis C is very different. Table IV shows that, despite improved survival profiles, a policy of using E rather than C is reflected in a reduction in the numbers of first operations for each reoperation prevented. The additional costs incurred by the choice of prosthesis cannot be recouped by the reduction in the number of reoperations. For those who prefer not to talk in terms of

Table III. Additional first operations possible under a policy of using prosthesis E which costs the same as prosthesis ($C_E = C_C$) and has improved survival profiles

Age group (yr)	Proportional reduction of failure rate (%)	Females calculated at			Males calculated at		
		5 years	9 years	15 years	5 years	9 years	15 years
< 55	20	1	5	15	2	6	17
	50	2	12	38	4	14	44
	90	4	22	72	7	27	84
55 to 64	20	1	3	10	2	6	19
	50	2	8	27	4	16	49
	90	3	15	50	7	30	94
65 to 74	20	1	4	11	0	3	10
	50	2	10	28	1	7	26
	90	4	18	53	1	14	48
> 75	20	0	1	3	1	2	4
	50	1	3	7	1	5	10
	90	1	5	12	3	9	19

Table IV. First total hip replacements foregone per reoperation prevented under a policy of using prosthesis E which costs twice as much as prosthesis ($C_E = C_C$) and has improved survival profiles

Age group (yr)	Proportional reduction of failure rate (%)	Females calculated at			Males calculated at		
		5 years	9 years	15 years	5 years	9 years	15 years
< 55	20	81	37	25	60	34	20
	50	31	13	8	23	12	6
	90	17	6	2	12	5	1
55 to 64	20	126	58	38	50	30	18
	50	49	22	13	19	10	5
	90	26	11	5	10	4	1
65 to 74	20	85	52	41	143	54	43
	50	33	19	14	57	20	14
	90	18	9	6	31	10	6
> 75	20	295	224	247	164	159	179
	50	117	88	95	64	61	67
	90	49	77	107	35	32	34

cost saving, the ratio of first operations foregone per reoperation prevented has increased.

The role of the randomised controlled trial. The model can assist in the design of clinical trials of new components. It can predict the cost or survival limits outside which a component system will not prove cost-effective in a given recipient group. These limits, once established, define the smallest relevant improvement in outcome that is worth detecting, and thus determine the size of a trial. The opportunities for cost-effective improved performance indicated by the economic model are limited. Thus, although the argument for insisting on the use of randomised controlled trials (RCT) in the introduction of any new intervention or technology is strong, there are practical difficulties in their use. They would need to be performed over a minimum of 15 years since at that time survival of a conventional prosthesis is still about 80%, but manufacturers would be unwilling to delay the general introduction of new components for such a period. An alternative would be to conduct a 15-year trial with interim analyses starting at five years. The study would need to be large enough to have a reasonable chance of detecting a difference at five years and intermediate reviews would be needed. Pocock (1984) suggested that to achieve an overall nominal level of

significance of $\alpha = 0.05$ over 15 years, the p value at any intermediate assessment should be no larger than 0.01 to declare a significant difference. The break-even price at 15 years should be used to decide if the trial is justified. The information in Table I shows that a randomised trial recruiting men and women under the age of 75 years would be appropriate in the case of a new component system priced at 1.4 times C_C for which a plausible claim is made of a 50% reduction in prosthetic failure rate at 15 years. A claim for a 20% reduction in prosthetic failure rate for a prosthesis priced at $2C_C$ would not justify either an RCT or licensing of the prosthesis. The same projected reduction in failure rate with a C_E set at $1.2C_C$, however, would lead to a trial being considered in men under 65 and women under 55 years of age.

The method of Lakatos (1988) can be used to calculate sample size taking account of staggered entry, loss to follow-up through death and different probabilities of failure in the new and standard prosthesis groups. Using the two examples given, and assuming loss to follow-up due to death, and recruitment over a two-year period, the number of patients required in the studies, using three different trial designs, are given in Table V, assuming $\alpha = 0.05$ and 90% power.

Table V. Sample sizes required for randomised controlled trials of new prostheses

Group	Reduction in failure rate (%)*	Length of study (yr)	Sample size†
Women < 55 years	20	5	50060
		15	5760
		15‡	8175
Men < 65 years	20	5	30696
		15	4891
		15‡	6942
Men and women < 75 years	50	5	5516
		15	929
		15‡	1319

* reduction in failure rate for experimental compared with conventional prosthesis

† $\alpha = 0.05$ (two-sided), 90% power, two-year recruitment period

‡ 15-year study with multiple looks at the data

Table VI. Sensitivity of break-even price ratios (C_E/C_C) to the ratio of prosthetic cost/total intervention cost, using four published estimates, assuming reoperation costs are equal ($R_E = R_C$)

Break-even at (yr)	Proportional reduction of failure rate (%)	Ratio			
		1:4	1:8	1:12	1:15
Males aged < 55					
5	20	1.0	1.1	1.1	1.1
	90	1.2	1.3	1.4	1.5
9	20	1.1	1.1	1.2	1.3
	90	1.4	1.7	2.0	2.2
15	20	1.2	1.3	1.4	1.5
	90	1.8	2.3	2.9	3.3
Females aged 65 to 70					
5	20	1.0	1.0	1.1	1.1
	90	1.1	1.2	1.3	1.4
9	20	1.1	1.1	1.1	1.2
	90	1.3	1.5	1.7	1.8
15	20	1.1	1.2	1.2	1.3
	90	1.5	1.8	2.1	2.3

DISCUSSION

A value-based approach to the assessment of quality in total joint replacement has recently been described as subjective, lacking well-defined limits, and of little practical guidance (Faro and Huiskes 1992). We do not agree. Since the performance of a number of prosthetic designs is well established, the opportunities for improvement are clearly defined. Modelling allows rational evaluation of the probability of cost-effectiveness of new designs.

The model is sensitive to the actual information which is entered, and, in particular, to the ratio between the total cost of the intervention and the cost of the established prosthetic components. Sensitivity analysis (Table VI) shows, however, that the upper limit of cost-effectiveness for C_E is relatively stable for likely improvements in prosthetic performance over a range of estimates of costs. For the largest potential improvement in performance, and the lowest ratio of prosthetic to total costs, substantially higher prices for C_E could be justified. We emphasise that current local figures should be entered if use is made of the model for planning purposes.

Prosthesis selection. The results indicate that in THR possible future savings due to improved survival and lower costs of reoperation do not usually justify the use of components which cost substantially more than a conventional cemented prosthesis of average market price. Current trends in hip replacement indicate widespread use of designs priced well above the limits suggested by the model.

The greatest part of the cost of THR is independent of that of the prosthetic components unless these allow significant reduction in hospital stay and morbidity, which is

unlikely in primary replacement. The use of uncemented components, however, would obviate the need for removal of cement in revision arthroplasty leading to significant savings in operating time and associated morbidity. At present this advantage is theoretical and has not been established. Sensitivity analysis indicates that it would be a relatively unimportant variable unless new technology was to offer the prospect of effective revision surgery cheaper than the initial procedure.

Evaluation of new components. RCTs which show a significant impact on survivorship to revision are large and expensive to mount, since the need for the earliest possible evaluation of outcome demands a multicentre design with rapid recruitment. A minimum follow-up of 15 years seems necessary for thorough assessment since it is only as revision rates increase that the potential of more effective implants becomes apparent. If new technology was subjected to a mandatory randomised trial, however, the use of the approach which we have described could lead to permission for unrestricted marketing as soon as a statistically significant improved performance was established. Alternative outcome measures with early identification of femoral loosening or acetabular wear would allow a more rapid decision about general release of the new technology but the true survivorship profile would require confirmation by continuing surveillance to a minimum of 15 years.

Implications for design, manufacture, marketing and legislation. New designs offer their recipients the unproven possibility of better outcome, but, equally, no certainty that it will not be worse. If the risk of a worse outcome appears small, most surgeons and patients would opt for the possibility of improved efficacy, providing that neither party has to consider costs. In these circumstances, there is little professional incentive to conduct an RCT.

For the manufacturer the possibility of improved performance has a market value. Unless there is a regulatory requirement, the decision on whether or not to institute an RCT is an economic one. If the costs of RCT are expected to 'pay off' by increasing market share and perhaps market price of the prosthesis, then conducting an RCT will be profitable. If these commercial benefits can be achieved on the basis of possible, but unproven improvement, the additional costs of RCT may not be justified. Therefore few randomised trials have been performed. Yet the use of new implant designs outside the context of a rigorous comparative study can only be justified if cost as well as design and material composition are indeed 'substantially equivalent'. Were governments to insist on RCTs before licensing any new components, this would, in the first instance, place the costs of RCTs upon the supplier, who might recover them if the RCT were to establish improved performance.

An alternative approach might be the negotiation between purchaser and supplier of a contractual guarantee of improved performance which, if not met, would result in price adjustment. Such contracts occur in other sectors of commerce.

Both these approaches present difficulties. The prosthesis itself is only one of many variables affecting the outcome; surgical judgement and skill, for example, are outside the manufacturer's control. The costs of confirming effectiveness by RCT could be greater than the possible benefits, given the large sample size and the time and cost involved. The need to stabilise health expenditure and the emerging appreciation of what constitutes cost-effectiveness in joint replacement may slow the multiplication of new designs, making RCT a more cost-effective option. For designers and manufacturers this model indicates an upper limit for the market price of cost-effective new technology, given achievable effectiveness, which should focus future research, development and marketing strategies, and stimulate alternative approaches to component design and manufacture.

Many countries enforce detailed standards relating to the manufacture, packaging, and sterilisation of prostheses (Faro and Huiskes 1992), but few demand other than loosely interpreted 'substantial equivalence' for prosthetic components. Proof of cost-effectiveness is now mandatory for new drugs and new applications of existing drugs in Australia and elsewhere (Freemantle et al 1995). The long-term outcome of interventions for chronic disease such as joint replacement presents a more difficult problem, since the parties involved cannot wait for long periods before realising the potential of new technology. Use of the clinical trial design which we have described could substantially shorten the period of evaluation of new designs with a projected survival profile which looks promising enough to test.

Implications for purchasers and providers. The selection of favoured components by surgeons and purchasing policies pursued by management should be mutually negotiated with the principles of cost-effectiveness in mind. The model indicates that the potential of more expensive but cost-effective new components would be realised 10 to 15 years after implantation, and before that time, assessment of cost-effectiveness would appear less encouraging. Therefore a commitment to long-term outcome measurement, and to budgeting strategies which take account of distant benefit will be required. In a climate of cost-containment, this may be politically difficult.

The approach which we have described raises professional, ethical and political issues. Decisions upon limits to health spending and resource allocation within health lie in the hands of the politicians. Health workers should be free

to participate in that debate, but professional responsibility includes a commitment to make the best use of the resources that society chooses to allocate. The freedom granted to surgeons to exercise choice depends on the style of health system in which they work. Many see control of that choice as interference in their ability to deliver contemporary standards of care. We believe that choice should not be the subject of government or bureaucratic regulation but a professional decision, based on known outcomes and plausible alternatives, rather than dictated by fashion or by unproven expectation.

It is the surgeon's role to act on the patient's behalf should there appear to be conflict between managerial goals and the interest of patients. That role reasonably includes consideration of the interests of those candidates for joint replacement who have yet to present, and particularly for those who are waiting for surgery (Smith et al 1990). Where limited funds are available for health care, the use of information about cost-effectiveness must influence the choice of prosthesis.

We believe that this model should be used in research and development to allow direction towards increased affordability for a given level of performance. Increase in prosthetic costs can only be justified by proportional improvement in outcome.

No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

REFERENCES

- Ahnfelt L, Herberts P, Malchau H, Andersson GBJ. Prognosis of total hip replacement: a Swedish multicenter study of 4,664 revisions. *Acta Orthop Scand* 1990;61 Suppl 238:1-26.
- Bulstrode CJK, Murray DW, Carr AJ, Pynsent PB, Carter SR. Designer hips: don't let your patient become a fashion victim. *BMJ* 1993;306:732-3.
- Daellenbach HG, Gillespie WJ, Crosbie P, Daellenbach US. Economic appraisal of new technology in the absence of survival data: the case of total hip replacement. *Soc Sci Med* 1990;31:1287-93.
- Dorey F, Amstutz HC. The validity of survivorship analysis in total joint arthroplasty. *J Bone Joint Surg [Am]* 1989;71-A:544-8.
- Faro LMC, Huiskes R. Quality assurance of joint replacement: legal regulation and medical judgement. *Acta Orthop Scand* 1992;60 Suppl 250.
- Freemantle N, Henry D, Maynard A, Torrance G. Promoting cost-effective prescribing. *BMJ* 1995;310:955-6.
- Lakatos E. Sample sizes based on the log-rank statistic in complex clinical trials. *Biometrics* 1988;44:229-41.
- Pocock SJ. *Clinical Trials: a practical approach*. Chichester: John Wiley & Sons, 1984.
- Smith A, Maynard A, Grimley Evans J, Harris J. The ethics of resource allocation. Proc 33rd Annual Scientific Meeting of the Society for Social Medicine, Manchester, 1989. *J Epidemiol Community Health* 1990;44:187-90.