Work in Progress: Please do not quote without permission

Assessing the stability of cost-effectiveness estimates across geographical locations.

The case of laparoscopic versus standard hysterectomy

A. Manca[1], MJ Sculpher[1], N.Rice[1], A.Briggs[2]

1.Centre for Health Economics, University of York, UK
2. Health Economics Research Centre, University of Oxford, UK

Abstract

Purpose: To explore the use of regression modelling to assess the stability of cost-effectiveness estimates across different centres.

Methods: These methods were developed as part of an analysis of a large randomised controlled trial undertaken in 32 UK centres, to compare laparoscopic hysterectomy with standard (abdominal or vaginal) hysterectomy in women with non-malignant gynaecological disease. Data were collected, over a period of up to one-year post-operation, on patient-specific resource use (e.g. time in theatre, days in hospital, management of complications) and health status (using EQ-5D). The latter was used to calculate patient-specific quality-adjusted life-years (QALYs) over the period of follow-up. Centre-specific unit costs were collected to value resource use. Using alternative assumptions about society's willingness to pay for an additional QALY, net monetary benefits (NMB) were calculated, for each individual patient. A regression model was developed with NMB as the dependent variable regressed against the patient-level variable age. Using multi-level modelling, we also allowed for a *centre effect* (i.e. hospital) within the model. Cost-effectiveness was expressed in terms of cost-effectiveness acceptability curves specific to patient sub-groups and centre types.

Results obtained using a "standard" statistical analysis across all centres were compared with those obtained with the use of regression analytic techniques. This approach to the analysis of patient-level cost and effectiveness data collected as part of multi-centre randomised controlled trials (RCTs) is a potentially valuable way to assess variability in cost-effectiveness between locations.

Paper presented to the 60th UK Health Economists' Study Group Conference, University of East Anglia, Norwich, UK, 7-9 January 2002

Introduction

Randomised controlled clinical trials are traditionally analysed under the implicit assumption that the only elements that could differ between patients allocated to alternative forms of management (other than the treatments they undergo) are patient-level characteristics such as age, gender and co-morbidities. This *modus operandi* stems from the belief that standard randomisation should ensure equal distribution of these individual-level characteristics between patients in alternative treatment groups. Although the above is (partly) true in the case of single centre clinical trials, it has been argued that in the context of multi centre and multi national studies the analytical problem is compounded by the consideration that factors other than those related to the patient may systematically affect cost and/or effectiveness results. The centre the patient is treated in and the patient's clinician (Localio *et al.* 2001) are examples of such factors. In this case, understanding the role (and quantifying the impact) of these *centre-specific* variables on the relative cost-effectiveness of treatment options becomes central to the analysis when assessing the generalisability of trial-based economic analysis *by location*. Although several approaches have been suggested to handle this type of variability in economic evaluation, particularly in international trials, very few studies have attempted to formally quantify it.

The use of regression analytic techniques using patient-level net benefit as the dependent variable (Hoch J, *et al.* in press) potentially opens up a range of opportunities for modelling variability in cost-effectiveness analysis. The proposed framework could easily be adapted to facilitate an estimate of the cost-effectiveness of the alternative forms of management by country, centre or even by clinician. The benefits from the proposed shift in the analytical paradigm would include studying cost-effectiveness results by patient sub-group (e.g. by age, gender and baseline co-morbidites) and by location-related factors such as country (in the case of for international trials), centre and clinician (in the case of multi centre trials).

The present work is an extension of the above-defined regression-based approach, in which we make use of the multi-level framework in the economic analysis to allow for a *centre effect* (i.e. hospital) within the model. For a discussion and application of multilevel models see Rice and Leyland (1996), and for their use in health economics see Rice and Jones (1997).

Cost-effectiveness has been expressed in terms of a series of cost-effectiveness acceptability curves specific to patient sub-groups and centres. The aim of this paper is threefold (a) to demonstrate the superiority of multi level modelling in estimating incremental net benefits compared to standard Ordinary Least Square (OLS) by allowing for the known multi level structure in the data; (b) to present how to conduct *sub-group economic analysis* at patient and hospital level using multi level modelling; and finally (c) to show how multi level regression models can be used to assess the stability of cost-effectiveness estimates across geographical locations.

Methodological Background

There is a growing number of economic evaluations make use of patient-level data - with a high prevalence of economic analyses alongside multi centre and or multi national clinical trials. This has prompted the publication of several methodological contributions, and it could be argued that some form of convergence in a number of subjects is starting to emerge. Examples of subjects where health service researchers are beginning to agree are (a) the methods for the correct quantification of sampling uncertainty; (b) how to present this information to decision-makers; and (c) appropriate methods to estimate medical costs in presence of censored data.

Unlike the above areas of research, one of the subjects where major contributions are needed is the investigation of methods to assess the extent to which the results of studies are generalisable across different settings and patient groups (Drummond MF, *et al.*, 1992; O'Brien BJ, 1997; Mason J, 1997; Phelps, 1997; Bryan S, Brown J, 1998; Johnston K *et al*, 1998). Economic assessment involves the measurement and valuation of the resource use and consequences of alternative interventions to assess their relative value for money (Gold MR, *et al*., 1996; Drummond MF, *et al.*, 1997). Given the nature of the problem, we can expect that the estimated results in terms of cost and health outcomes are likely to be affected by the *setting* (e.g. geographical location, type of health care facility) from which the data were collected. In this context, examples of factors that may limit generalisability *by location* of the results of economic evaluations are (Drummond MF, *et al*., 1992; O'Brien BJ, 1997):

- *demography and epidemiology of disease;*
- *clinical practice and conventions;*
- *incentives and regulations facing health care providers;*
- *absolute or relative price levels;*
- *clinician and or patient individual preferences;*
- *opportunity cost of resources;*
- *difference in costing methodologies;*

Although evidence that these differences can be important exists, little analytical research has been produced to explore the causal relationships between the factors outlined above and the absolute difference or direction of change in estimates of cost and effect in studies undertaken in different settings. Most of the methods literature has focused on the problems of generalisability of economic studies undertaken in parallel with RCTs (Drummond MF, Davies LM, 1991; Drummond MF, Davies LM, 1994; Adams ME, *et al.* 1992; Coyle D, *et al.* 1998). In particular, a number of papers have considered the question of *how to analyse* international trials (Willke *et al.* 1998; Schulman K, *et al*. 1996; Jonnson B, *et al.* 1997). These include multi-national trials where potential differences between countries in clinical practice and unit costs limit the applicability of results to any one country; or trials undertaken in one country where attempts are made to transfer their results to one or more other locations

outside the trial. Approaches that have been proposed for the analysis of these studies include

- the application of regression analysis to estimate costs and effects whilst controlling for between-country variation (Willke *et al.* 1998); and
- the use of tests for heterogeneity to check for outcome-centre interactions, where absence of interactions is considered justification for pooling data (Cook J, *et al.* 1997).

Each of these approaches has its advantages and limitations, and systematic comparison of methods to assess the extent to which they generate convergent results is outside the scope of the present work.

The various methods used in international trials may be equally applicable to multi-centre trials in the NHS. Standard methods of trial-based economic analysis calculate total costs per patient, simply multiplying the volume of resources used by that patient by the unit costs for those resources. However it is obvious that health care providers (hospital, primary care trusts, etc.) have the possibility to alter the mix of resources used in the health care process as a response to variations in the relative prices of these resources. In this case the use of *pooled* resource use and *average* unit costs across different settings can be misleading, in practice reducing the variability in the results, and therefore masking the *substitution effect*.

Hoch and colleagues (Hoch JS, et al., in press) have recently suggested a reformulation of the cost-effectiveness problem within a standard regression-type framework. The authors demonstrate the versatility of this approach and suggest more potential applications using an example in which standard OLS regression is applied. However, although the present paper builds on their work, we argue that Hoch *et* al. fail to recognise the relevance of the hierarchical nature of the data in a RCT.

In a multi centre (multi national) randomised trial individuals within a centre are more similar than those from different centres. In pragmatic clinical trials for instance, patients in different hospitals may be managed in slightly different ways.

In this case the assumption underpinning the use of the standard OLS regression - that the observations are *independently distributed* - may not hold. Failure to acknowledge this feature of the data in the analysis would result in incorrect p-values and confidence intervals (because of the correlation among patients within the same centre), biased estimates of standard errors and therefore misleading conclusions. Without an adjustment for clustering the standard OLS would produce results that may seem statistically significant when they are not.

The use of mixed models offers a range of potential advantages such as an increase in the precision of the estimates and the opportunity to make wider

inference (Brown and Prescott, 1999). In this paper we shall use the terms mixed models and multi level models interchangeably. It is important here to stress that even when the difference between treatment effects can be assumed to be identical in all centres, a mixed model can improve the precision of the treatment effect estimates by accounting for the hierarchical nature of the data. Therefore, the emphasis here is on estimation rather than hypothesis testing.

In this work we applied a multi-level regression model of patients nested within hospitals using data from a recently conducted cost-utility analysis undertaken alongside an UK multi-centre randomised clinical trial.

Clinical Context

The EVALUATE trial is a recently conducted multi centre RCT comparing laparoscopic-assisted hysterectomy and standard hysterectomy (vaginal or abdominal). The study enrolled 1380 women which were first allocated by their clinician to one of the two forms of standard treatment on the basis of their clinical condition. Once allocated to one of the two standard procedures, the randomisation process between laparoscopic-assisted hysterectomy and the standard procedure took place. The time horizon for the initial economic analysis was a median of 52 weeks (range 3-52). Data were collected at baseline, at 6-weeks after discharge, at 4 and 12 months. All relevant resource use was costed adopting a UK National Health Service perspective. Health benefits were expressed in terms of quality-adjusted life-years (QALYs). Given the design of the trial, in effect two separate comparisons were undertaken depending on which standard procedure patients were initially allocated to: (a) laparoscopic-assisted hysterectomy versus vaginal hysterectomy; and (b) laparoscopic-assisted hysterectomy versus abdominal hysterectomy.

To control for the *learning curve* effect when performing the procedures, each surgeon was required to have completed each type of operation at least 25 times. At the patient-level, the only inclusion criterion was that a woman should have had gynaecological symptoms which directed her surgeon to conclude that a hysterectomy was indicated, and that she consented to be included. Patients were excluded from the study if (a) confirmed or suspected malignant disease of the genital tract was present; (b) if they had 2nd or 3rd degree prolapse; or (c) if a uterine mass greater than the size of a 12-week pregnancy was detected.

Resource use and health outcomes

Using detailed case record forms, resource utilisation data were collected for each woman in the study using a patient record form. These were completed by clinical staff in trial centres, and complemented by postal questionnaires completed and returned by patients. The study collected information on a number of different resource use items (e.g. length of stay in hospital, duration of surgical procedure, time in recovery room, use drugs, anaesthetics, consumables, post-operative complications, etc). All women were invited to

return to hospital for a follow-up visit approximately 6 weeks after their operation. During this visit, case record forms were used to collect data on the incidence of any complications, including whether further surgery or blood transfusion was required. Patients were also asked to complete similar questionnaires at 4 months and 12 months after their operation. UK unit costs at 1999-2000 prices were used to value the resource use measured in trial (CIPFA, 2000; BNF N.39, 2000; Netten and Dennet, 2000; Booth, 2000; Gough, 2000). The health outcomes for the economic analysis were quantified in terms of quality-adjusted life-years (QALYs). QALYs were calculated, for each woman in the trial, on the basis of their responses to the EQ-5D at baseline, and at each of the follow up assessments.

Methods

For the present exercise we use the data only from the comparison laparoscopic assisted hysterectomy *versus* abdominal hysterectomy in the EVALUATE trial. To avoid censored cost data, we only include 6-week follow-up resource use and QALYs. A total of 32 hospitals in England, Scotland, Wales, and South Africa participated in the trial. However, due to the unavailability of centre-specific unit cost data for the Welsh, Scottish, and South African hospitals it was decided to focus this work on the English centres. In addition, limited availability of data on hospital characteristics reduced the number of centres used in this application to 18, for a total of 540 patients (Laparoscopic assisted hysterectomy n=361, and Abdominal hysterectomy n=179). The multilevel model included both patient and hospital-specific covariates - where available. The results from the standard (pooled) statistical analysis are reported in Table 1a, while Figure 1 shows the cost-effectiveness acceptability curve for the above comparison. For the present analysis STATA 7.0 software was used.

Multi level modelling versus OLS

The first objective of this paper is to show the superiority of multi level modelling over a standard OLS regression approach when the purpose of the economic analysis (alongside a multi centre or multi national RCT) is to estimate the differential net benefit associated with the new intervention. To achieve this, and the other objectives set out in the introduction section, we compare the results obtained from the two regression frameworks (i.e. OLS and multi level modelling) over a range of different model specifications. We start with a simple model where the variable Net Monetary Benefit (NMB) is regressed against no other explanatory variable but the treatment arm. The OLS specification is represented by equation 1.1, whereas equation 1.2 describes the same model under a two-level hierarchical structure of patients nested within hospitals.

$$NMB_i = \alpha + \delta t + \varepsilon_i \qquad (1.1)$$

$$NMB_{ij} = \alpha + \delta t + u_{ij} \qquad (1.2)$$

For consistency with the notation used by Hoch and colleagues, we identify t as the treatment dummy (0 = standard treatment, 1= new intervention). The index i (= 1,…..p) is the patient identifier, whereas j (=1,…...k) is the hospital identifier. The Net Monetary Benefit for a generic patient in the trial is defined as:

$$NMB = QALY \cdot \lambda - Cost \tag{2}$$

where λ represents the decision maker willingness to pay for an additional QALY. In the two-level random intercept model specification described from equation 1.2, the error term can be decomposed as

$$u_{ij} = v_j + e_{ij} \tag{3}$$

in which v_j represents the intercept shift of the j^{th} hospital on the average NMB (i.e. shift from α), and is assumed to be normally distributed with mean zero and constant variance. Equation 1.2 can therefore be rewritten as:

$$NMB_{ij} = \alpha + \delta t + v_j + e_{ij} \tag{1.3}$$

The coefficient δ represents the incremental net monetary benefit between the two arms of the trial, which in the OLS regression case is identical to that obtained in a traditional statistical "pooled" analysis. A comparison of the results obtained from the OLS regression (see Table 1.2) with those resulting from a traditional statistical analysis (see Table 1.1), shows the equivalence of the two approaches. Results are reported for three different λ values (i.e. λ=0, λ=30,000, and $\lambda \to \infty$). It is perhaps worth stressing that the case where λ=0 is equivalent to use the negative of the cost vector as dependent variable, while the case where $\lambda \to \infty$ is equivalent to use only effects as dependent variable.

A closer look at Table 1.2 confirms that - as we expected - because the hierarchical nature of the data is now considered in the analysis, standard errors of the estimated coefficients under the multi level specification are larger than those obtained in the OLS model. The multi level regression in Table 1.2 shows how this specification can help first, to account for the role of the *centre effect*, and second to disentangle the variability in the results attributable to level-1 variation (i.e. patient level) from that associated with level-2 variation (i.e. at hospital level). Fig. 1 shows the cost effectiveness acceptability curve for these results reported in Table 1.2.

Multi level modelling versus OLS for patient sub-group analysis

Let's now discuss the second area of application where mixed models can prove to be superior to standard OLS regression: sub-group analysis. For this purpose, equations 1.1 and 1.3 are reformulated with the addition of patient-level covariates and treatment-covariate interactions in the model (see equations 4.1 and 4.2.). Table 2.1 reports the results obtained from including

the patient-specific characteristic *age* in the two model specifications (for simplicity we use only one patient characteristic to illustrate the method).

$$NMB_i = \alpha + \delta t_{ij} + \sum \beta x_i + \sum t_{ij}\gamma x_i + \varepsilon_i \quad (4.1)$$

$$NMB_{ij} = \alpha + \delta t_{ij} + \sum \beta x_{ij} + \sum t_{ij}\gamma x_{ij} + v_j + e_{ij} \quad (4.2)$$

In the above equations, x_{ij} is the patient characteristic age, which varies at both hospital and patient level, and β is its regression coefficient. Similarly $t_{ij}x_{ij}$ represent the treatment-age interaction, and γ its coefficient. For sub-group analysis, we are interested in quantifying the marginal variation associated with the treatment for specific categories of patients in the trial. The magnitude and significance of the coefficient γ will therefore provide an indication of how cost-effectiveness of treatment is expected to vary at the margin.

Figure 2.1 shows a cost-effectiveness acceptability curve describing the results of the economic sub group analysis using the median age in the sample (i.e. 41.2 years) as cut off for the two age sub groups.

Hospital sub group analysis
Following the same logic we could also potentially perform a hospital subgroup analysis (see equations 4.3 and 4.4). Table 2.2 reports the results of the analysis using the variable *hospital average inpatient stay* across all specialties and its interaction with the treatment as explanatory variables, and Fig. 2.2 shows the cost-effectiveness acceptability curves using an average hospital length of stay of 3.2 days as cut off.

$$NMB_{ij} = \alpha + \delta t_{ij} + \sum \theta z_j + \sum \phi t_{ij} z_j + \varepsilon_{ij} \quad (4.3)$$

$$NMB_{ij} = \alpha + \delta t_{ij} + \sum \theta z_j + \sum \phi t_{ij} z_j + v_j + e_{ij} \quad (4.4)$$

Variable z_j is the hospital characteristic *hospital average inpatient stay*, which varies only at hospital level, and θ is its regression coefficient. Similarly $t_{ij}z_j$ represent the treatment-hospital average length of stay, and ϕ its regression coefficient. Again we are interested in quantifying the marginal variation associated with the treatment for specific categories of hospitals, and the coefficient ϕ will provide an indication of how cost-effectiveness of treatment is expected to vary at the margin.

However, it is important to stress that a hospital sub-group analysis conducted without accounting for the case-mix of the patients treated in each hospital can be highly misleading.

Multi-level random effect model
In the light of this, the full model specification is obtained from equations 4.1 and 4.2, adding centre-specific variables and the interaction terms between

these and the treatment arm (see equations 5.1 and 5.2). Results are reported in Table 3.

$$NMB_{ij} = \alpha + \delta t_{ij} + \sum \beta x_{ij} + \sum t_{ij}\gamma x_{ij} + \sum \theta z_j + \sum t_{ij}\phi z_j + \varepsilon_{ij} \quad (5.1)$$

$$NMB_{ij} = \alpha + \delta t_{ij} + \sum \beta x_{ij} + \sum t_{ij}\gamma x_{ij} + \sum \theta z_j + \sum t_{ij}\phi z_j + v_j + e_{ij} \quad (5.2)$$

The third objective of the present work was to demonstrate how multi level regression models could be used to assess the stability of cost-effectiveness estimates across geographical locations. Random effect models assume that the *centre effect* is characterized by a distribution, that is that differences among centres vary randomly. In this study the validity of adopting a random effect intercept specification for the mixed model was verified by conducting a Hausman specification test on each of the models in equations 1.3, 4.2, 4.4, and 5.2.

The benefit from using a valid random effect model rather than a fixed effect specification is that inference drawn from the former can be applied to the full population of centres, whereas any conclusion derived from a fixed effect model is valid only within the sample of the centres that were part of the study. In addition in a random effect model estimates are closer to the overall mean than if they were fitted as fixed (Brown and Prescott, 1999).

For the above reasons the use of multi level models is particularly useful to assess the issue of generalisability of the results *by location*.

Results from Tables 2.1 and 2.2 show that when we move from a simple model with no other explanatory variable but treatment arm (i.e. equation 1.3) to a model where we take into account patient specific variables (i.e. equation 4.2), to a model including both patient and hospital variables we are successful in explaining part of the variability in the model. The proportion of variability due to hospital level in the case of λ=0, for instance, is reduced from a 12.53% in the model described by equation 1.3, to 6.29% in the full model (equation 5.2). The implication of this is that the possibility to explain the variability at hospital level allows an increase in the precision of the estimates and the opportunity to make wider inference.

Conclusions and Discussion

This paper shows that multi level regression techniques can be used to extend the application of traditional OLS techniques in a context were a natural hierarchy in the data structure is observed. In addition, the benefits from using multi level modelling is that accounting for centre-specific as well as patient-level covariates helps to explain variability across patient sub-groups and by geographical location.

In this work we demonstrated (a) the superiority of the multi level models in estimating incremental net benefits compared to standard OLS in the context of

economic evaluations conducted alongside multi centre RCTs; (b) that multi level modelling can be employed to conduct *sub-group economic analysis* at patient and hospital level; and finally (c) that multi level regression models can be used to assess the stability of cost-effectiveness estimates across geographical locations.

However, a number of issues still need to be addressed in this context. First, problems with the quality of the data about hospital characteristics can produce spurious results. Heterogeneity across accountancy methods in the estimation of unit costs by hospital is one possible element of risk. In addition, it is still not clear which hospital variables we need to use in the model, as different variables can attempt to explain different elements in the hospital production function (e.g. economies of scale, economies of scope, efficiency, etc).

In this study, not all centres have provided their unit costs. To solve this problem we applied a linear prediction method from the available hospital specific unit costs, which is far from perfect and, provided we are interested in the mean unit cost, it relies on the assumption that the observed hospitals are representative of those for which data were unavailable.

Second, we have applied a two-level random intercept model to the dataset. This is equivalent to saying that the slope coefficients in regression equation associated to the patient variables are identical across hospitals. One way to extend this formulation is to build a random intercept and slope model.

Finally, there may be concerns around the policy implications of this kind of analysis. Providing the decision-maker with information regarding which sub-groups of patients is likely to benefit most is definitely valuable, and to identify in which sub-groups of patients the health technology is most cost-effective can be useful. But, what could be the impact of saying: “this intervention is cost-effective in a specific hospital in the north of the country but it is not in London”?

What is sure is that *the marriage of health econometrics and medical statistics* within economic evaluation opens the door to new forms of analysis and at the same time poses new (and perhaps reiterate old) methodological and theoretical challenges to health economist.

References

Adams ME, McCall NT, Gray ET et al. Economic analysis in randomized controlled trials. Medical Care 1992;30:231-243.

Booth C. Review Body for Nursing Staff, Midwives, Health Visitors and Professions Allied to Medicine 17th Report (Cm 4563). London: HMSO, 2000.

Brown H, Prescott R. Applied mixed models in medicine, Chichester: John Wiley & Sons, Ltd.,1999.

Bryan S, Brown J. Extrapolation of cost-effectiveness information to local settings. Journal of Health Services Research and Policy 1998;3:108-112.

Cook J, Drummond M, Glick H et al . Analysing economic data from multi-national clinical trials. Issues and recommendations. Paper presented at the Health Economics Research Group meeting, July 1997.

Coyle D, Davies L, Drummond M. Trials and tribulations - emerging issues in designing economic evaluations alongside clinical trials. International Journal of Technology Assessment in Health Care. 1998;14:135-144.

Drummond MF, Bloom BS, Carrin G et al. Issues in the cross-national assessment of health technology. International Journal of Technology Assessment in Health Care 1992;8:671-682.

Drummond MF, Davies L. Economic analysis alongside clinical trials. International Journal of Technology Assessment in Health Care 1991;7:561-573.

Drummond MF, Davies L. Economic evaluation of lenograstim for prophylaxis of chemotherapy induced neutropenia in patients with small cell lung cancer. Pharamacoeconomics 1994;6 (suppl 2):44-52.

Drummond MF, O'Brien BJ, Stoddart GL et al. *Methods for the Economic Evaluation of Health Care Programmes*. New York: Oxford University Press, 1997.

Gold MR, Siegel JE, Russell LB et al. *Cost-Effectiveness Analysis in Health and Medicine*. New York: Oxford University Press, 1996.

Gough CB. Review Body on Doctors and Dentists Remuneration 29th Report (Cm 4562). London: HMSO, 2000.

Hoch JS, Briggs AH, Willan A. Something old, something new, something borrowed, something BLUE: a framework for the marriage of health econometrics and cost-effectiveness analysis. *Health Economics*, in press.

Johnston K, Gerard K, Brown J. Generalizing costs from trials. Analyzing center selection bias in a breast screening trial. International Journal of Technology Assessment in Health Care 1998;14:494-504.

Jonnson B, Weinstein M. Economic evaluation alongside multinational clinical trials: study considerations for GUSTO IIb. International Journal of Technology Assessment in Health Care 1997;13:49-58.

Koopmanschap MA, Touw KCR, Rutten FFH. Analysis of costs and cost-effectiveness in multinational trials, *Health Policy* 2001, **58**: 175-186

Localio AR, Berlin JA, Ten Have TR, Kimmel SE. Adjustment for center in multicenter studies: an overview, *Annals of Internal Medicine* 2001, **135**: 112-123.

Mason J. The generalisability of pharmacoeconomic studies. Pharmacoeconomics 1997; 11:503-514.

Netten A, Dennett J, Knight J. Unit Costs of Health and Social Care. Canterbury: PSSRU, University of Kent, 2000.

O'Brien BJ. A tale of two (or more) cities: geographic transferability of pharmacoeconomic data. The American Journal of Managed Care 1997;3:S33-S39.

Phelps CE. Good technologies gone bad. How and why the cost-effectiveness of a medical intervention changes for different populations. Medical decision Making 1997;17:107-117.

Rice N, Jones AM. Multilevel models and health economics, *Health Economics* 1997, **6**: 561-575.

Rice N, Leyland A. Multilevel models: applications to health data, *J Health Serv Res Policy* 1996, **1** (3): 154-164.

Schulman K, Glick H, Buxton M et al. The economic evaluation of the FIRST study: design of a prospective analysis alongside a multinational Phase III clinical trial. Controlled Clinical Trials 1996;17:304-315.

Sculpher MJ, Bryan S, Dwyer N et al. An economic evaluation of transcervical endometrial resection versus abdominal hysterectomy for the treatment of menorrhagia. British Journal of Obstretrics and Gynaecology 1993;100:244-252.

The Chartered Institute of Public Finance and Accountancy (CIPFA). The Health Service Database 2000. Croydon: CIPFA, 2000.

Willke RJ, Glick H, Polsky D et al. Estimating country-specific cost-effectiveness from multinational clinical trials. Health Economics 1998;7:481-493.

Table 1.1: Results from the standard statistical analysis (*N=540)*

Group variable	Mean	S.D.	S.E.
Overall analysis			
Abdominal Arm (n=179)			
Cost	1399.25	705.47	52.72
QALYs	.0870	.0205	.0015
Laparoscopic Hysterectomy Arm (n=361)			
Cost	1478.053	818.38	43.07
QALYs	.0903	.0217	.0011

Table 1.2: Comparison of OLS and multilevel estimates (*N=540)*

	OLS analysis*			Multilevel analysis		
	Coefficient	*SE*	*t-value*	*Coefficient*	*SE*	*t-value*
Net monetary benefits with λ = 0 (i.e. –Costs)						
Constant term	-1399.25	52.68	-26.56	-1514.07	92.88	-16.3
Treatment dummy	-78.80	68.06	-1.16	-88.29	68.43	-1.3
σ_v (Level 1)				282.68		
σ_e (Level 2)				746.58		
Fraction due to level 1 variability				0.1253		
Net monetary benefit with λ=30,000						
Constant term	1212.49	70.65	17.16	1146.27	95.50	12
Treatment dummy	18.10	90.07	0.2	15.92	92.36	0.17
σ_v (Level 1)				204.96		
σ_e (Level 2)				1009.26		
Fraction due to level 1 variability				0.0396		
Net monetary benefit with $\lambda\rightarrow\infty$ (=Effect)						
Constant term	0.0870	0.0015	56.87	0.0872	0.0020	43.31
Treatment dummy	0.0032	0.0019	1.69	0.0033	0.0019	1.76
σ_v (Level 1)				0.0044		
σ_e (Level 2)				0.0210		
Fraction due to level 1 variability				0.0427		

* Using robust standard errors

Table 2.1: Comparison of OLS and multilevel estimates for the model with patient covariates (*N=540*)

	OLS analysis*			Multilevel analysis		
	Coefficient	*SE*	*t-value*	*Coefficient*	*SE*	*t-value*
Net monetary benefit with λ=0 (i.e. –Cost)						
Constant term	-1413.67	81.707	-17.30	-1537.84	103.21	-14.90
Treatment dummy	-9.59	93.47	-0.10	-13.25	90.76	-0.15
Age	35.87	95.58	0.38	59.71	114.89	0.52
Treatment-Age interaction	-150.84	129.92	-1.16	-166.52	139.50	-1.19
σ_v (Level 1)				280.60		
σ_e (Level 2)				746.38		
Fraction of variance due to level 1 variability				0.1238		
Net monetary benefit with λ=30,000						
Constant term	1162.89	103.34	11.25	1101.98	113.19	9.74
Treatment dummy	73.33	125.14	0.59	71.18	122.60	0.58
Age	123.30	133.91	0.92	115.81	154.93	0.75
Treatment-Age interaction	-135.12	175.21	-0.77	-134.06	188.16	-0.71
σ_v (Level 1)				199.37		
σ_e (Level 2)				1010.48		
Fraction of variance due to level 1 variability				0.0374		
Net monetary benefit with λ→∞ (=Effect)						
Constant term	0.0858	0.0020	42.04	0.0858	0.0020	41.69
Treatment dummy	0.0027	0.0027	1.02	0.0027	0.0025	1.07
Age	0.0029	0.0030	0.95	0.0029	0.0032	0.90
Treatment-Age interaction	0.0005	0.0038	0.14	0.0005	0.0039	0.13
σ_v (Level 1)				N/A		
σ_e (Level 2)				0.021027		
Fraction of variance due to level 1 variability				N/A		

* Using robust standard errors

Table 2.2: Comparison of OLS and multilevel estimates for the model with hospital covariates (*N=540*)

	OLS analysis*			Multilevel analysis		
	Coefficient	*SE*	*t-value*	*Coefficient*	*SE*	*t-value*
Net monetary benefit with λ=0 (i.e. –Cost)						
Constant term	-1432.75	85.09	-16.84	-1500.74	121.58	-12.34
Treatment dummy	-74.55	107.44	-0.69	-83.23	95.08	-0.88
Hospital average LOS	69.74	103.97	0.67	-46.90	196.67	-0.24
Treatment dummy x Hospital average LOS	-8.69	134.67	-0.06	-10.94	137.11	-0.08
σ_v (Level 1)				294.14		
σ_e (Level 2)				747.28		
Fraction of variance due to level 1 variability				0.1341		
Net monetary benefit with λ=30,000						
Constant term	1170.37	107.10	10.93	1113.16	134.17	8.30
Treatment dummy	46.76	136.43	0.34	45.13	128.12	0.35
Hospital average LOS	87.66	140.41	0.62	41.61	208.05	0.20
Treatment dummy x Hospital average LOS	-59.57	179.05	-0.33	-61.29	184.80	-0.33
σ_v (Level 1)				247.65		
σ_e (Level 2)				1010.12		
Fraction of variance due to level 1 variability				0.0566		
Net monetary benefit with λ→∞ (=Effect)						
Constant term	0.0867	0.0021	40.39	0.0867	0.0025	34.68
Treatment dummy	0.0040	0.0027	1.48	0.0041	0.0026	1.56
Hospital average LOS	0.0005	0.0030	0.19	0.0009	0.0037	0.26
Treatment dummy x Hospital average LOS	-0.0017	0.0038	-0.44	-0.0017	0.0038	-0.45
σ_v (Level 1)				0.0033		
σ_e (Level 2)				0.0210		
Fraction of variance due to level 1 variability				0.0245		

* Using robust standard errors

Table 3: Comparison of OLS and multilevel estimates for the full model (*N=540*)

	OLS analysis*			Multilevel analysis		
	Coefficient	*SE*	*t-value*	*Coefficient*	*SE*	*t-value*
Net monetary benefit with λ=0 (ie. –Costs)						
Constant term	-1447.80	113.44	-12.76	-1507.49	112.78	-13.37
Treatment dummy	-23.84	134.12	-0.18	-30.55	118.79	-0.26
Age	36.83	96.29	0.38	52.23	115.97	0.45
Hospital average LOS	70.23	104.74	0.67	-5.05	159.64	-0.03
Treatment dummy x Age	-113.88	166.30	-0.68	-117.61	160.67	-0.73
Treatment dummy x Hospital average LOS	33.68	137.26	0.25	34.02	158.30	0.21
Treatment x Age x Hospital average LOS	-82.75	175.01	-0.47	-84.78	160.28	-0.53
σ_v (Level 1)				193.80		
σ_e (Level 2)				747.73		
Fraction of variance due to level 1 variability				0.0629		
Net monetary benefit with λ=30,000						
Constant term	1119.49	138.47	8.08	1090.69	135.11	8.07
Treatment dummy	93.50	175.17	0.53	90.92	159.64	0.57
Age	124.53	134.77	0.92	117.38	155.67	0.75
Hospital average LOS	89.31	141.01	0.63	60.75	178.65	0.34
Treatment dummy x Age	-115.58	218.41	-0.53	-112.49	215.62	-0.52
Treatment dummy x Hospital average LOS	-39.40	198.29	-0.2	-40.36	212.83	-0.19
Treatment x Age x Hospital average LOS	-44.92	225.02	-0.2	-43.72	215.27	-0.2
σ_v (Level 1)				152.55		
σ_e (Level 2)				1012.31		
Fraction of variance due to level 1 variability				0.0222		
Net monetary benefit with λ→∞ (=Effect)						
Constant term	0.0855	0.0025	33.36	0.0855	0.0025	33.08
Treatment dummy	0.0039	0.0037	1.04	0.0039	0.0033	1.17
Age	0.0029	0.0030	0.95	0.0029	0.0032	0.89
Hospital average LOS	0.0006	0.0030	0.21	0.0006	0.0032	0.20
Treatment dummy x Age	-0.00004	0.0044	-0.01	-0.00006	0.0045	-0.01
Treatment dummy x Hospital average LOS	-0.0024	0.0046	-0.52	-0.0024	0.0044	-0.55
Treatment x Age x Hospital average LOS	0.0012	0.0044	0.28	0.0012	0.0045	0.28
σ_v (Level 1)				0.00036		
σ_e (Level 2)				0.02105		
Fraction of variance due to level 1 variability				0.00029		

* Using robust standard errors

Figure 1: Cost-effectiveness acceptability curve from regression analysis pooled results (laparoscopic *vs.* Abdominal)

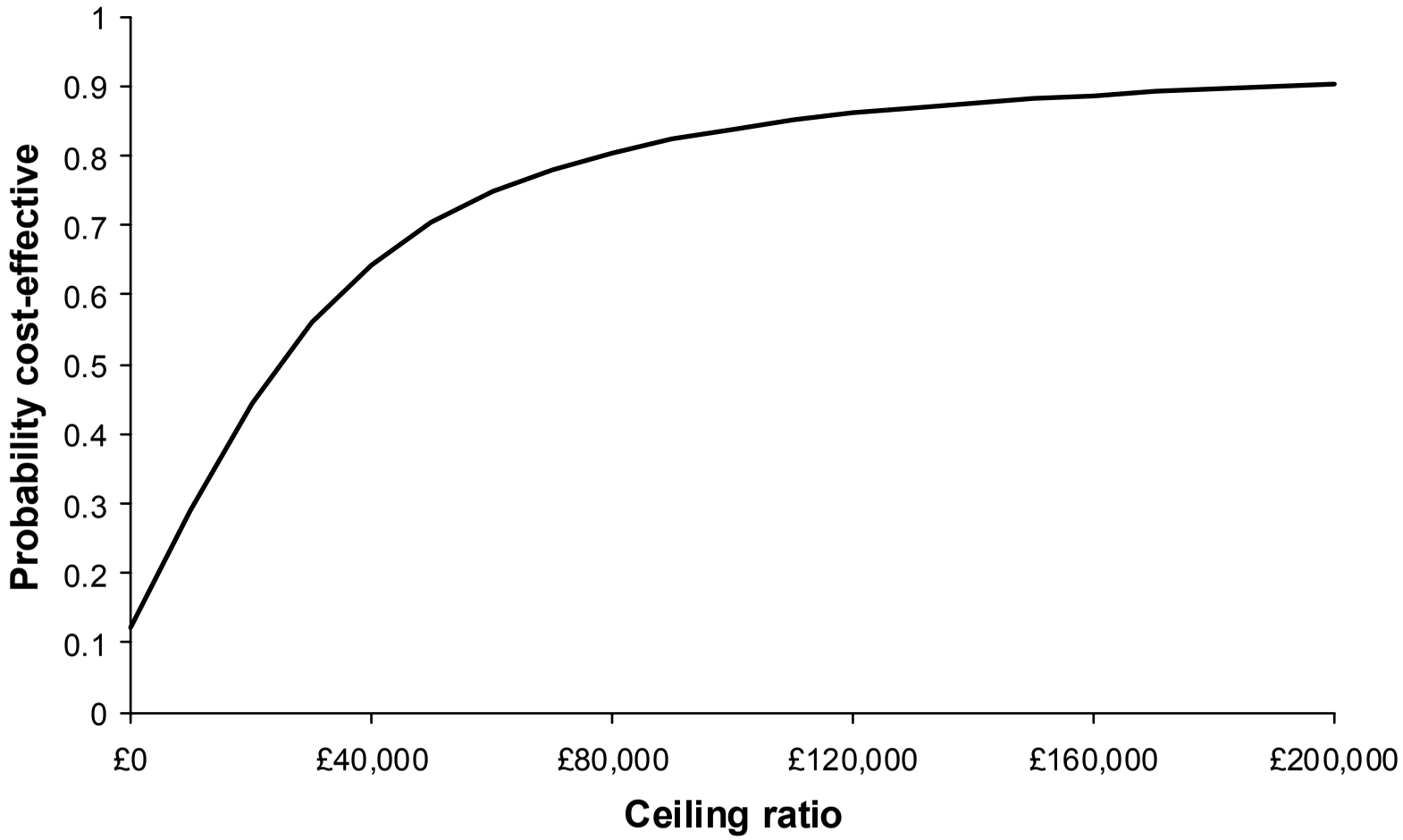

Figure 2.1 : Cost-effectiveness acceptability curve from multi level regression analysis, using patients' *age* and treatment-age interaction as additional covariates

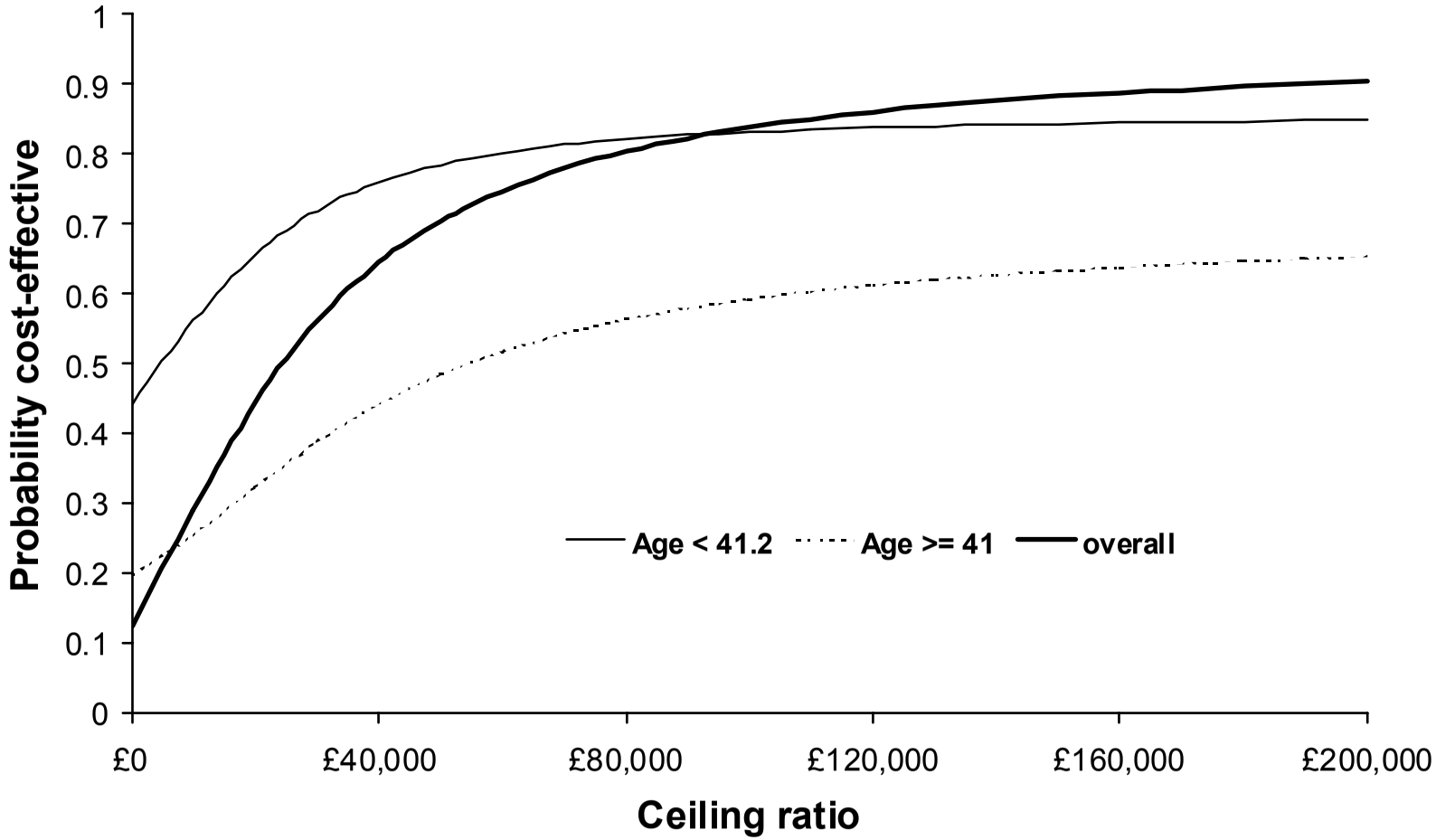

Figure 2.2 : Cost-effectiveness acceptability curve from multi level regression analysis using hospital average inpatient length of stay (across all specialties) and treatment-hospital average in-patient stay interaction as additional covariates

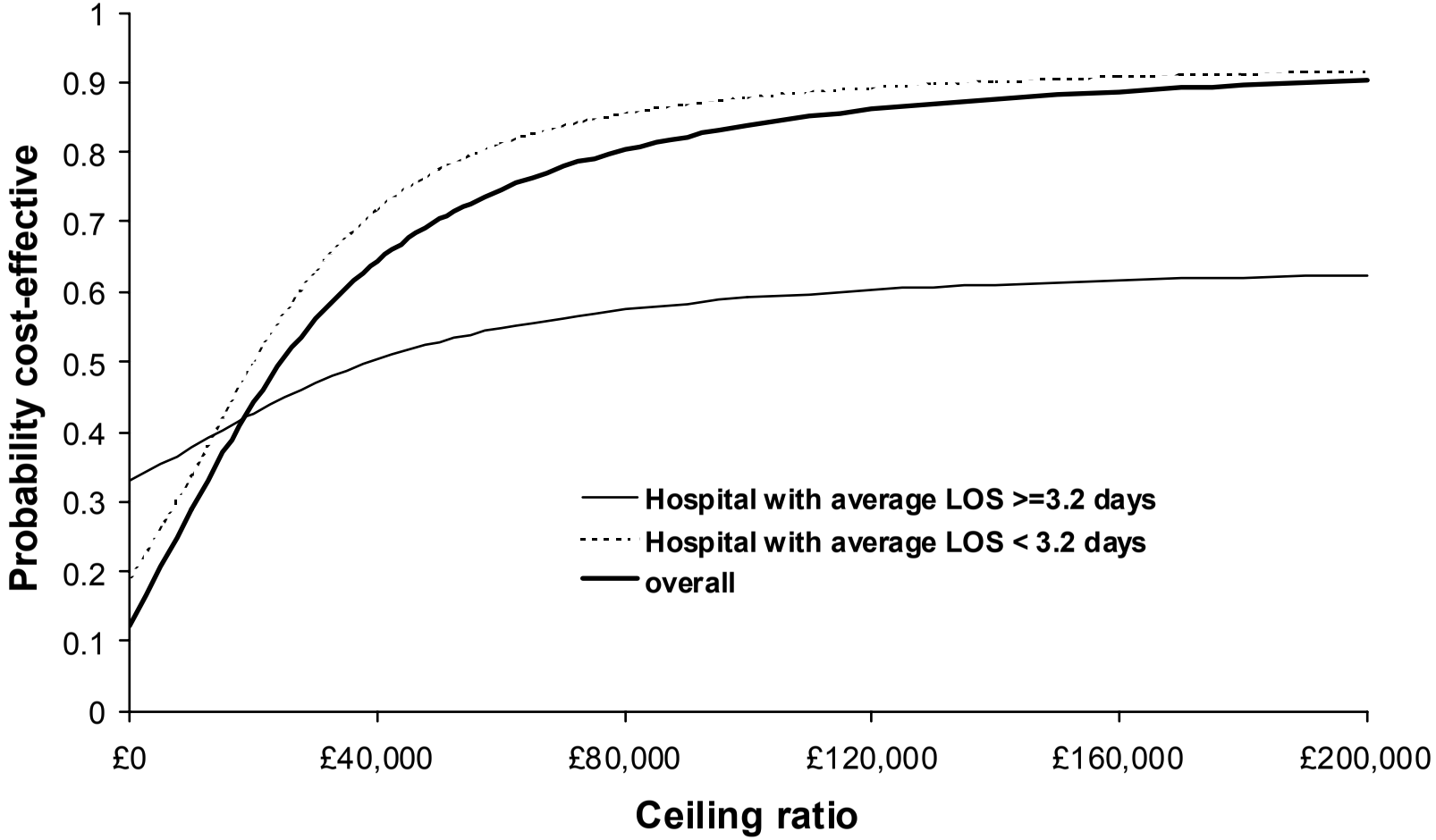

Figure 3: Cost-effectiveness acceptability curve from multi level regression analysis using the full model