Analysing Cost Data From Cluster Randomised Trials

Kathy Johnston
Health Economics Research Centre
University of Oxford
Institute of Health Sciences
Oxford OX3 7LF

Email:kathy.johnston@ihs.ox.ac.uk

Paper presented at the Health Economists' Study Group, University of Oxford,

January, 2001

Work in progress. Please do not quote.

Acknowledgements

I am grateful for contributions from Rebecca Turner, Pat Yudkin and Mike Moher.

Summary

This paper discusses issues in the analysis of cost data from cluster randomised trials. The key issues are whether to adjust for baseline costs and how to adjust for clustering. In analysing clinical data from cluster randomised trials, the analysis is performed at either the cluster-level or patient level. The former requires aggregating data to the level of the cluster; the latter requires adjusting for the effects of clustering using multi-level modelling. This paper adopts both approaches in analysing cost data from a cluster randomised trial. The trial compares three methods of promoting secondary prevention of coronary heart disease in primary care. Costs are estimated for 1906 patients aged 55-75 years with established coronary heart disease from 21 general practices. The cluster-level results greatly underestimates the variability in costs and have low precision. The patient-level results suggest that patient variation in cost far outweighs practice variation in cost, partly as a result of skewed patient-level data. Further refinements to the model are required to confirm this finding.

Introduction

Cluster randomised trials

Cluster randomised trials are increasingly being used in health services research, partly as a result of a need to evaluate organisational interventions and implementation trials i.e. those designed to change professional practice (Mason et al. 1999). In such trials, groups of patients, rather than individual patients, are randomised (Donner, 1998). Randomisation by individual is inappropriate (Ukoumunne et al. 1999a) because the outcome of each patient can no longer be assumed to be independent of that for any other patient (Campbell and Grimshaw, 1998). This violates one of the assumptions underlying conventional statistical methods (Bland and Kerry, 1997).

The lack of independence between patients has implications for the sample size of cluster randomised trials and statistical power of the cluster randomised trial is reduced in comparison to similarly sized patient randomised trials (Campbell et al. 2000). A statistical measure of independence is the intra-cluster correlation coefficient (ICC). The sample size estimates have to be inflated by a factor $1 +(n-1)p$ where n is the average cluster size and p is the ICC. The impact of this inflation factor on the resulting sample size will depend on clustering effect which may be high if, for example, management of patients is very similar within clusters but very different between clusters. Standard sample size calculations underestimate the number of patients required because they allow for variations within clusters but not between clusters (Ukoumunne et al. 1999b)(Kerry and Bland, 1998a)(Kerry and Bland, 1998b)Ukoumunne et al. 1999b).

The lack of independence between patients in cluster randomised trials also has implications for the analysis of data. Ignoring clustering will generally cause standard errors to be underestimated (ML manual). In the clinical literature there are essentially two approaches to analysing the data. The first approach is to aggregate data by cluster, construct a summary statistic for each cluster and then use a t-test (or weighted t-test). The second approach is to analyse individual patient data using multi-level modelling. This models the association between patients in the same cluster (Ukoumunne et al. 1999b) and estimates the intervention effects controlling for both individual and cluster level characteristics (Ukoumunne et al. 1999b)(Kerry and Bland, 1998a).

Economic evaluation alongside trials

In recent years, one notable trend in economic evaluation of health care interventions has been the number of economic evaluations conducted alongside, or as part of, clinical trials. Clearly, one of the main advantages of using clinical trials as a framework for economic evaluation is that they provide the opportunity to collect and analyse patient-specific resource use (and hence cost) data. This has led to a body of methodological research into how to address uncertainty in economic evaluation (Briggs and Gray, 1999) and, in particular, how to perform statistical analysis in the presence of skewed cost data (Barber and Thompson, 1998; Thompson and Barber, 2000; Briggs and Gray A., 1998).

There is now a substantial number of economic evaluations alongside clinical trials but few of these are alongside cluster randomised trials and only two have been identified (Baxter et al. 1999); (Lord et al. 1999). These two studies adopted different approaches to analysis. Baxter et al. (1999) separated costs into those incurred by the practice and those incurred by the other parts of the NHS. They then reported the costs according to randomisation group but no statistical analysis of costs was performed and point estimates of cost were presented. Lord et al. (1999) estimated costs one year before and one year after the intervention. Costs were analysed using patient-level data using linear regression to adjust for baseline differences in cost and clustering at the practice level. A fixed effects model was used and consequently there was separation of variance in cost into that arising because of between practice variation and that arising because of within practice variation.

Although there are currently few published economic evaluations alongside cluster trials, in the future this number is expected to grow given the current increase in the number of studies evaluating general practice based interventions and interventions designed to change professional practice. Although the issues in the analysis of clinical data from cluster randomised trials have been set out clearly, there has been no consideration of whether the recommended approaches remain relevant for the analysis of cost data. Consequently, the methodological issues in the analysis of economic evaluations of such trials need to be investigated. In particular, of interest is whether the issues are similar to, or different from, the issues in the analysis of economic data from trials where the unit of randomisation is the patient. For example, some of the issues in the analysis of cluster randomised trials may be a subset of those raised by multi-centre trials. Centre variation in cost between centres (or clusters) may arise because of differences in, for example, organisation and practice. To date,

the literature on the economic evaluation of multi-centre trials has focused on whether centre-specific costs should be attached to centre-specific resource use (Raikou et al. 2000; Willke et al. 1998). These studies have not used a multi-level modelling to explore centre differences although other studies have applied multi-level modelling to cost data (Carey, 2000)(Van-der and Cairns, 1998).

Aims of paper

The aims of this paper are to:

- Identify methodological issues in the analysis of cost data from cluster randomised trials
- Perform initial analysis of cost data from a cluster randomised trial
- Assess whether the approach taken to analyse clinical data from cluster randomised trials is relevant for cost data.

Methods

ASSIST trial

Design

The aim of the ASSessment of Implementation STrategies (ASSIST) trial is to compare three methods of promoting secondary prevention of coronary heart disease (CHD) in primary care in Warwickshire. All three interventions included a baseline audit of the practice notes of all patients aged 55-75 years, to identify patients with established CHD. The interventions were:

- Audit and feedback (A&F) (control group): Following the baseline audit, practices were told the number of patients in their practice identified as having CHD. They received summary audit results, and anonymised data from other practices in the study as a basis for comparison and were asked to provide usual patient care.
- Recall to general practitioner (RGP): In addition to the information given in the A&F group, these practices were given the names of patients who were identified as having CHD and given support in setting up a register and recall system for regular review of patients by their general practitioner.
- Recall to nurse clinic (RN): In addition to the information given in the RGP group, these practices were also given the names of patients identified with CHD and given ongoing

support in setting up a register and recall system for a nurse-run clinic for systematic review of patients.

The interventions ran for 18 months and at the end of the intervention a final audit determined adequate assessment. [1]

Sample size

The sample size for the clinical outcomes was designed to test differences in adequate assessment. Both the number of patients and the number of clusters (general practices) was considered. The proportion of patients expected to be adequately assessed at baseline was estimated as about 25%, based on the results of a previous study. Allowing for a small improvement during the 18 months of the trial, it was expected that 35% of control group patients would be adequately assessed at the end of the trial. The trial was designed to detect an increase in this proportion to 55% in RGP (effect size 0.40) and 75% in RN. To detect a difference between 35% and 55%, with 80% power (two-tailed α=0.05) and assuming individual randomisation, 96 patients were needed in each group. This sample size would have 83% power to detect a difference between 55% and 75%.

In relation to cluster size, published epidemiological data suggested that a practice of 10,000 patients might expect about 500 of them to have CHD, of whom roughly half would be aged 75 years or under. With an average practice size in Warwickshire of about 7,000 patients, a mean cluster size (CHD patients per practice) of about 175 was estimated. A cluster inflation factor was applied and sensitivity analyses were performed assuming an intra-cluster correlation coefficient between 0.05 and 0.0617 and a cluster size between 100 and 200. Under any of these assumptions seven practices in each trial arm would give the study power of at least 82% (maximum 90%).

The final sample size was a total number of patients recruited was 1906: A&F, 559 patients; RGP, 682 patients; and RN 665 patients.[2] Seven clusters per trial arm were recruited giving a total of 21 general practices in the trial. The highest number of patients in a practice was 222.

---

[1] Adequate assessment is tightly defined to include recent assessment and recording of three risk factors: blood pressure, smoking habit, and serum cholesterol.

[2] In the end, taking only a sample of CHD patients from a practice was not an option since the identification of all patients with CHD was one of the incentives offered to practices to encourage them to join the trial. Consequently, all patients with established CHD were included.

Analysis of costs

Cost estimates

Heath service costs associated with the interventions as they would be incurred in routine practice were estimated; consequently research costs, such as travel costs, were omitted. Resource use data was collected at both the baseline and final audits. The costs were separated into practice and other NHS costs.[3] Costs were estimated for the 18 months before and after the intervention.

Practice costs included the costs of: baseline audit and feedback; setting up the intervention; GP and nurse consultations; and performing final audit to determine adequate assessment. The costs of the baseline audit were estimated by recording the time taken by nurses to audit the notes of all patients aged 55-75 for each practice. The costs of feeding back the results of the audit were estimated by recording the time taken to present the results. The costs of setting up the recall to GP and recall to nurse interventions included the costs of training and the costs of providing additional capacity (additional GP and nurse sessions). The costs of consultations were estimated by recording the number of GP and nurse consultations over the 18 months of the intervention at the final audit of patient notes. The costs of performing the final audit of patient notes (to determine whether patients were adequately assessed) were estimated by dividing the time taken per patient note reviewed at the baseline audit.

Other NHS costs included the costs of: inpatient admissions; tests and investigations (full blood counts, liver function tests, cholesterol tests, glucose tests); outpatient consultations; and drugs. The number of inpatient admissions was costed assuming an average length of stay per type of episode was using trust financial returns data. The drugs (and dose of each drug) each patient was being prescribed was recorded at baseline and final audits. The dosage of each drug was costed based on 1999 net prices (BNF, 2000). An additional cost to the NHS for dispensing, professional fees and allowances was included for each prescription, assuming four prescriptions per year. An 18 month follow up drug cost was calculated and assumed that for the first 3 months of the intervention the patient was on the drugs recorded at baseline and for the remaining 15 months the intervention the patient was on the drugs

[3] Detailed description of costs is available from the author.

recorded at final audit. An 18 month baseline drug cost was estimated by assuming that the patient was on the drugs recorded at baseline for the previous 18 months.

Attribution of costs

A factor sometimes overlooked in deciding how to measure resource use is attribution, that is, whether resource use (or cost) is attributable to an intervention (Johnston et al. 1999). Establishing whether or not resource use is related is often a subjective decision (Hurley et al. 1995). In clinical trials, however, attribution of costs can be determined by randomisation and, consequently, only costs from the point of randomisation are included. This approach aims to adjust for any pre-randomisation differences in cost. Some economic evaluations have, however, adjusted for baseline costs. For example, in an economic evaluation of a cluster randomised trial, Lord et al. (1999) adjusted for baseline costs (costs in the year before the intervention was implemented). Attribution of costs post-randomisation is an issue even in a randomised trial. For example, some resource use, for example, inpatient admissions may be unrelated to the intervention.[4] Thus, the issue of attribution of costs within trials may not be as straightforward as it first appears.

In terms of pre-randomisation differences, many clinical trials adjust for differences between patients at baseline when they analyse clinical data. This practice has recently been questioned by (Assmann et al. 2000)) who argue that significance tests of variables at baseline should not be performed and that any baseline differences arise as a result of flawed randomisation. If baseline adjustments are to be made, they are two approaches to doing so. The first is to subtract the baseline value from the follow up value. The second is to adjust for covariates using analysis of covariance. The latter approach is more precise and can also increase the statistical power of a study (Norman and Streiner, 1994).

In this paper, the following approach to baseline adjustment of costs is adopted. The primary approach was to adjust for baseline differences in the costs of conducting baseline audit and feedback of patients since this was not attributable to the intervention. Henceforth this is referred to as baselinecost1. A secondary approach was adopted and adjusts for baseline differences in the costs of all baseline costs (the costs of audit, inpatient admissions, drugs, GP and nurse consultations). Henceforth this is referred to as baselinecost2. Before

[4] For example an inpatient admission for colorectal cancer in an intervention examining hip replacement should be defined as unrelated.

adjustments were made for clustering and baseline, summary statistics on costs were estimated without adjustments.

Cluster-level analysis

In the cluster-level analysis, cost data for each individual was aggregated to the level of the practice. This resulted in 21 observations on cost. The costs of the three interventions were then compared using a weighted analysis of variance, with the weights being the number of patients in each practice. Two separate analyses were conducted, one adjusting for baselinecost1 and the other adjusting for baselinecost2. Baseline adjustments were conducted using a weighted analysis of covariance with baseline cost entered as a covariate. The analysis was carried out using general linear modeling in SPSS.

Patient-level analysis

Multi-level modelling was used to perform patient-level analysis. There are two levels in the model: level 2 (practice) and level 1 (patient). The multi-level model adjusts for the effects of clustering. The aim of the model is to adjust for the effect of baseline cost before the effect of the interventions is tested. Consequently the covariates included in the model were the baseline cost (either baselinecost1 or baselinecost2) and two dummy variables for the three interventions (coded as 1 if true, 0 otherwise).

At this stage in the work, two models are built, differing according to how random effects are specified. In future work, other random effects will be examined. Model 1, the simplest model, is a variance components model. Model 2 allows for random effects across practices.

Model 1 (variance components) is specified as follows:

$$y_{ij} \sim N(XB, S)$$

$$y_{ij} = \beta_{0ij} X_o + \beta_1 X_{1ij} + \beta_2 X_{2j} + \beta_3 X_{3j}$$

$$\beta_{0ij} = \beta_0 + u_{0j} + e_{0ij}$$

$$[u_{0j}] \sim N(0, S_u) : S_u = [\sigma^2_{u0}]$$

$$[e_{0j}] \sim N(0, S_e) : S_e = [\sigma^2_{e0}]$$

Where *XB* is the fixed part of the model; Ω the covariance matrix (random part of the model); X_0 is a constant; X_1 is baseline cost; X_2 is a dummy variable for RGP; and X_3 is a dummy variable for RN. The variables u_{0j} and e_{0ij} indicate that it is a multi-level model. Both u_{0j} and e_{0ij} are random quantities (with means equal to zero) and represent the random part of the model, with u indicates random departures from level 2; e indicates random departures at level 1. The ICC is the level 2 variance as a proportion of the sum of the level 1 and 2 variances.

Model 2 is specified as follows:

$$y_{ij} \sim N(XB, \Omega)$$

$$y_{ij} = \beta_{0ij} X_o + \beta_{1j} X_{1ij} + \beta_2 X_{2j} + \beta_3 X_{3j}$$

$$\beta_{0ij} = \beta_0 + u_{0j} + e_{0ij}$$
$$\beta_{1j} = \beta_1 + u_{1j}$$

$$[u_{0j}] \sim N(0, \Omega_u) : \Omega_u = [\sigma^2_{u0}]$$
$$[u_{1j}] \sim N(0, \Omega_u) : \Omega_u = [\sigma_{u10} \ \sigma^2_{u1}]$$

$$[e_{0j}] \sim N(0, \Omega_e) : \Omega_e = [\sigma^2 e_0]$$

Where *XB* is the fixed part of the model; Ω the covariance matrix (random part of the model); X_0 is a constant; X_1 is baseline cost; X_2 is a dummy variable for RGP; and X_3 is a dummy variable for RN. All β parameters are the fixed part of the model. The terms u0*j* and u1*j* are random departures at the practice level from β_0 and β_1 and follow a multivariate normal distribution with mean 0 and covariance matrix Ω_u. In this model there are two random variables at level 2 so Ω_u is a 2 by 2 covariance matrix.

Two estimation procedures were used. The first was iterative generalised least squares (IGLS). This involves iterating until two consecutive estimates for each parameter are sufficiently close together and hence convergence has been achieved. IGLS assumes that u and e are normally distributed. If, however, u and e are not randomly distributed, then although the estimates of β will be consistent and efficient, the estimates of variances will not be efficient. Consequently, a second estimator was used - a restricted maximum likelihood

estimation or restricted iterative least squares (RIGLS). This approach leads to unbiased estimates (Rice and Jones, 1997). In future work, additional estimators will be used. The models produce loglikelihood statistics to allow testing of whether a change in specification improves that fit of the model. The analysis was carried out using Mlwin 1.1 software.

Results

Summary statistics on costs (no adjustments for clustering or baseline)

Table 1 presents mean costs, based on patient-level data, without adjusting for clustering or baseline costs. Mean costs are presented by cost category for both follow up costs and baseline costs, where appropriate. Average total cost (follow up costs only) is estimated. In terms of cost category, RN has the highest set up cost and highest follow up nurse cost. RGP has the highest inpatient stay cost. Overall, RGP has the highest average total cost but all average total costs have very large standard deviations compared to their means, indicating that the data is skewed (partly as a result of some patients experiencing inpatient stays).

Table 1: Mean costs per patient by intervention

Cost category	A&F Mean cost £	RGP Mean cost £	RN Mean cost £
Practice costs			
Audit and feedback	39.07	33.82	29.62
Setting up intervention	0	14.05	31.63
GP consultations			
Baseline	32.91	29.77	26.60
Follow up	20.95	30.81	23.39
Nurse consultations			
Baseline	14.86	3.33	3.83
Follow up	4.24	1.16	17.73
Final audit	2.49	2.37	2.26
Other NHS costs			
Drugs			
Baseline	277.92	218.85	245.49
Follow up	409.93	378.81	383.36
Inpatient stays			
Baseline	495.77	320.93	549.35
Follow up	180.28	314.01	248.63
Tests/investigations	26.94	30.94	49.23
Outpatient consultations	49.13	74.58	71.05
Average total cost *(sd)**	639.93 *(1056.87)*	846.71 *(1056.88)*	827.28 *(1056.88)*

* Follow up costs only

Cluster-level analysis

Table 2 presents the mean cost per patient after adjusting for the effects of clustering and baseline cost using a cluster-level analysis. Although the means are broadly similar to those reported in table 1, the standard deviations are markedly lower. This implies that cluster-level analysis ignores variation in cost between patients.

Table 2: Mean costs per patient by intervention (adjusting for clustering and baseline)

Baseline adjustment	A&F Mean cost £ (sd)	RGP Mean cost £ (sd)	RN Mean cost £ (sd)
Baseline cost 1	688.24 *(234.41)*	846.44 *(211.30)*	831.94 *(216.58)*
Baseline cost 2	684.09 *(235.93)*	863.80 *(220.48)*	818.03 *(216.45)*

Table 3 presents the mean cost differences and 95% confidence intervals between interventions. The only statistically significant cost difference is between RGP and A&F. Despite the lower variation in cost compared to unadjusted costs, the confidence intervals are very wide. This arises because the sample size is only 7 per intervention and thus large standard errors are observed resulting in low precision and lack of statistical significance. The confidence interval for the mean differences between RN and RGP, adjusting for baseline cost 1, implies that the results are not inconsistent with either an additional cost of £153.59 for RN compared to RGP or a saving of £123.78 for RN compared to RGP.

Table 3 Mean cost differences between interventions

Baseline adjustment	RGP (difference from A&F) Mean difference £	RN (difference from A&F) Mean difference £	RN (difference from RGP) Mean difference £
Baseline cost 1	+158.62	+143.71	-14.91
95% CI for difference	*13.31 to 303.93**	*-3.30 to 290.63*	*-123.78 to 153.59*
Baseline cost 2	+179.71	+133.93	-45.77
95% CI for difference	*27.71 to 331.71**	*-11.63 to 279.50*	*-191.29 to 99.75*

* mean difference is significant at the 0.05 level

Table 4 Patient-level analysis of costs using multi-level modelling

Explanatory variables	Model 1 Baseline cost1 *IGLS* *$ (s.e.)*	Model 1 Baseline cost1 *RIGLS* *$ (s.e.)*	Model 2 Baseline cost1 *IGLS* *$ (s.e.)*	Model 2 Baseline cost1 *RIGLS* *$ (s.e.)*	Model 1 Baseline cost2 *IGLS* *$ (s.e.)*	Model 1 Baseline cost2 *RIGLS* *$ (s.e.)*	Model 2 Baseline cost2 *IGLS* *$ (s.e.)*	Model 2 Baseline cost2 *RIGLS* *$ (s.e.)*
Fixed effects								
Constant	653.18 *(53.33)*	653.66 *(57.05)*	653.18 *(53.33)*	653.18 *(53.38)*	543.56 *(46.50)*	543.53 *(48.82)*	548.77 *(47.51)*	549.28 *(52.06)*
Baseline cost	0.57 *(0.79)*	0.55 *(0.81)*	0.57 *(0.78)*	0.57 *(0.78)*	0.16 *(0.02)*	0.16 *(0.02)*	0.17 *(0.04)*	0.17 *(0.04)*
RGP	174.20 *(59.08)*	173.40 *(63.37)*	174.20 *(59.08)*	174.20 *(59.13)*	204.92 *(58.25)*	203.79 *(61.73)*	195.05 *(55.90)*	192.56 *(62.67)*
RN	157.12 *(59.73)*	160.59 *(64.29)*	157.12 *(59.73)*	157.12 *(59.79)*	171.90 *(58.48)*	174.35 *(62.25)*	154.35 *(55.875)*	158.85 *(63.34)*
Random effects								
Level 2:								
$*^{2:}u_{0j}$	0	1554	0	0	0	1247	3455	6520
$*^{2:}u_{1j}$	n/a	n/a	0	0	n/a	n/a	-8.28; 0.016	-9.24; 0.017
Level 1:	1067020	1068039	1067020	1069259	1036464	1037624	1017893	1017385
$*^{2:}e_{0ij}$								
*2loglikelihood	31865	31865	31865	31865	31809	31809	31793	31793

n/a is not applicable

Patient-level analysis

Table 4 presents the results of patient-level analysis of costs using multi-level modelling. The first part of the table presents the results of model 1 and model 2 using both IGLS and RIGLS, adjusting for baseline cost1. The second part of the table adjusts for baseline cost 2.

In terms of the fixed part of the model, the coefficients on RGP and RN can be interpreted as estimates of incremental costs of the intervention, that is, the cost of each intervention minus the costs of the control group (A&F). Adjusting for baseline cost1, the magnitude of the coefficients is similar for models 1 and 2 using either IGLS or RIGLS. Adjusting for baseline cost2, the magnitude of the coefficients on the intervention effects are slightly different (for example under model 1 using IGLS, the coefficient on RGP is £204.92 but is £195.05 under model 2 using IGLS).

In terms of the random part of the model, the estimates of u0*j* (between practice variance in cost) are very small. For all the IGLS models, the between practice variation in cost is zero. This confirms that IGLS produces inefficient estimates of variances. Using the RIGLS estimates of u0*j* produces ICCs of approximately 0.001 (i.e. 0.1% of the variation in cost arises as a result of variation between practices). This implies that practice variation in cost is very small. The estimates of e0*j* suggest that the variation between patients is very high. This would be expected given the cost data before adjustment for clustering or baseline found very large variation in cost between patients (partly as a result of inpatient costs).

Discussion

There are a number of methodological issues in the analysis of cost data from cluster randomised trials, specifically, whether to adjust for baseline costs and how to account for clustering. In terms of adjusting for clustering, one of the aims of this paper was to establish whether, in cluster randomised trials, the methods used in the analysis of clinical data were relevant for cost data.

Cluster-level analysis greatly underestimated the variation in cost and results in low precision because of the aggregation of costs to practice level. It is unclear how to interpret the results

of cluster-level analysis since any causal relationship should include individual patients but information about these is discarded. The problem is that the relationship between aggregate level variables may not be the same as that between the same variables at individual level.

Patient-level analysis did address patient variation and found that the majority of variation in cost (99.99%) can be explained by variation in cost between individuals and that there is little variation in cost between centres. This suggests that there may be little need to adjust for clustering (centre variation) given the greater variation in patient costs. If baseline costs are to be adjusted for, a fixed effects model could be used. If baseline costs are not to be adjusted for, the usual approach to analysing cost data in clinical trials could be adopted.

Before the finding that centre variation in cost is low in cluster-randomised trials is confirmed, further analysis should be undertaken. In particular, this should focus on additional estimators. Both the estimators used, IGLS and RIGLS, assume errors are randomly distributed. Since the data are skewed, future analysis will explore use of alternative estimators such as non-parametric bootstrapping. Given, however, the extent of patient variation, it is unlikely that an alternative estimator would uncover additional between-practice variation in cost. As well as applying a different estimator, future work will explore in more detail the statistical significance of results and conduct formal specification tests of the models.

Skewed cost data arose in this study as a result of some patients having inpatient admissions. Given that the interventions were targeted at patients with existing coronary heart disease it is unsurprising that some patients experienced inpatient admissions. The sample size of the trial was not powered to detect differences in the number of inpatient events (or inpatient costs). Given that skewed cost data is common in many clinical trials, future trials (both individually randomised and cluster randomised) should determine in advance the power of the study in the presence of inpatient admissions.

It is useful to consider whether the results from the initial analyses reported here would apply to other cluster randomised trials. In the ASSIST trial, with 21 general practices and 1906 patients, there are a small number of relatively large clusters. The number of practices may have been too small for a random effects model. This could be judged by comparing the results of the random effects model with that of a fixed effects model. An alternative

explanation of a low clustering effect is that each practice had the same trial protocol. A higher clustering effect might have been observed had the intervention been directed at the practitioner, rather than patient, level.

The results of this study from initial analyses suggest that patient-variation in cost is far greater than centre variation. Multi-level modelling could be usefully applied to other cluster randomised trials (and multi-centre trials) in order to investigate the relative magnitude of patient to centre variation in cost.

References

Assmann, S.F., Pocock, S.J., Enos, L.E. and Kasten, L.E. (2000) Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet* 355, 1064-1069.

Barber, J.A. and Thompson, S.G. (1998) Analysis and interpretation of cost data in randomised controlled trials: Review of published studies. *British Medical Journal* 317, 1195-1200.

Baxter, K., Peters, T.J., Somerset, M. and Wilkinson, C. (1999) Anxiety amongst women with mild dyskaryosis: Costs of an educational intervention. *Family Practice* 16, 353-359.

Bland, J.M. and Kerry, S.M. (1997) Trials randomised in clusters. *British Medical Journal* 315, 600

Briggs, A. and Gray A. (1998) The distribution of health care costs and their statistical analysis for economic evaluation. *Journal of Health Services Research and Policy* 3, 233-245.

Briggs, A.H. and Gray, A.M. (1999) Handling uncertainty when performing economic evaluation of healthcare interventions. *Health Technology Assessment* 3, 2

Campbell, M.K. and Grimshaw, J.M. (1998) Cluster randomised trials: Time for improvement. *British Medical Journal* 317, 1171-1172.

Campbell, M.K., Mollison, J., Steen, N., Grimshaw, J.M. and Eccles, M. (2000) Analysis of cluster randomized trials in primary care: A practical approach. *Family Practice* 17, 192-196.

Carey, K. (2000) A multilevel modelling approach to analysis of patient costs under managed care. *Health Economics* 9, 435-446.

Donner, A. (1998) Some aspects of the design and analysis of cluster randomization trials. *Appl.Statist.* 47, 95-113.

Hurley, S.F., Bond, L.M., Carlin, J.B., Evans, D.B. and Kaldor, J.M. (1995) A method for estimating baseline health care costs. *Journal of Epidemiology and Community Health* 49, 525-531.

Johnston, K., Buxton, M.J., Jones, D.R. and Fitzpatrick, R. (1999) Assessing the costs of health technologies in clinical trials. *Health Technology Assessment* 3, 6

Kerry, S.M. and Bland, J.M. (1998a) Statistics notes: Sample size in cluster randomisation. *British Medical Journal* 316, 549

Kerry, S.M. and Bland, J.M. (1998b) Trials which randomize practices II: Sample size. *Family Practice* 15, 84-87.

Lord, J., Victor, C., Littlejohns, P., Ross, F.M. and Axford, J.S. (1999) Economic evaluation of a primary care-based education programme for patients with osteoarthritis of the knee. *Health Technology Assessment* 3, 23

Mason, J., Wood, J. and Freemantle, N. (1999) Designing evaluations of interventions to change professional practice. *Journal of Health Services Research and Policy* 4, 106-111.

Norman, G.R. and Streiner, D.L. (1994) *Biostatistics: The bare essentials*, St Louis, Missouri: Mosby.

Raikou, M., Briggs, A., Gray A. and McGuire, A. (2000) Centre-specific or average unit costs in multi-centre studies? Some theory and simulation. *Health Economics* 9, 191-198.

Rice, N. and Jones, A. (1997) Multilevel models and health economics. *Health Economics* 6, 561-575.

Thompson, S.G. and Barber, J.A. (2000) How should cost data in pragmatic randomised trials be analysed? *British Medical Journal* 320, 1197-1200.

Ukoumunne, O.C., Gulliford, M.C., Chinn, S., Sterne, J.C. and Burney, P.J. (1999a) Methods for evaluating area-wide and organisation-based interventions in health and health care: A systematic review. *Health Technology Assessment* 3, 5

Ukoumunne, O.C., Gulliford, M.C., Chinn, S., Sterne, J.C., Burney, P.J. and Donner, A. (1999b) Methods in health service research. Evaluation of health interventions at area and organisation level. *British Medical Journal* 319, 376-379.

Van der Pol, M. and Cairns, J.A. (1998) The efficient organization of blood donation. *Health Economics* 7, 455-463.

Willke, R.J., Glick, H.A., Polsky, D. and Schulman, K. (1998) Estimating country-specific cost-effectiveness from multi-national clinical trials. *Health Economics* 7, 481-493.

Yudkin, P. and Moher, M. (2000) Putting theory into practice: A cluster randomised trial with a small number of clusters. *Statistics in Medicine* forthcoming,