Gynaecologic surgery from uncertainty to science:
evolution of randomized control trials

Tara J. Selman1,4, Neil P. Johnson2, Javier Zamora3 and Khalid S. Khan1

1Department of Reproductive and Child Health, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK; 2Department of Obstetrics and Gynaecology, University of Auckland, National Women’s Health, Auckland Hospital, Auckland, New Zealand; 3Clinical Biostatistics Units, Hospital Ramón y Cajal, Spain
4Correspondence address. E-mail: taraselman@blueyonder.co.uk

BACKGROUND: It is now accepted that both medical and surgical practice should be based on reliable and sound clinical evidence. However, randomized control trials comparing surgical interventions have been associated with many problems. The aim of this review is to assess if there has been progress made in establishing the evidence base for surgical interventions in gynaecology. METHODS: Relevant reviews were identified from Cochrane Database of Systematic Reviews (Issue 3, 2006) and data from individual randomized control trials extracted. Chi-squared test was used to compare quality pre- and post-Consolidated Standards of Reporting Trials (CONSORT) statement. Meta-regression analyses were performed to test the hypothesis that effect size decreased over time. Further multiple linear regression analyses were used to test the hypothesis that precision increased over time and finally a logistic regression model was used to estimate whether treatment effects differed between trials with and without allocation concealment. RESULTS: Twenty-three relevant reviews were identified, including 94 trials. The proportion of studies reporting allocation concealment significantly increased after the introduction of the CONSORT statement (P = 0.002). There was a trend towards improvement in precision over time. Similarly, there was a reduction in size of treatment effect over time (log of the ratio of odds ratios per year 0.96; 95% confidence interval 0.93–0.99, P = 0.04). CONCLUSIONS: Gynaecologic surgical practice appears to be benefiting from improvement in its research base in a subject where practitioners do not participate readily in randomized evaluation.

Keywords: Cochrane; evidence based; gynaecology; randomized controlled trials; surgery

Introduction

It is now widely accepted that healthcare practice, whether medical or surgical, should be based on reliable and sound clinical evidence. This is a far cry from what has been a largely non-evidence-based approach to practice in surgical specialties. There has been a tendency to adopt the latest surgical technique because it seems rational, or to advance surgical expertise, rather than because it fulfils the stringent criteria for effectiveness that we demand for medical interventions. Randomized controlled trials (RCTs) comparing surgical interventions have been associated with many problems, not only surrounding the logistic difficulties of establishing such a trial, but also in preventing sources of bias, for example, difficulties with blinding surgeons and patients risking performance or measurement bias. Such problems have resulted in general surgical specialities producing a remarkably small proportion (only 3.4% of publication in the leading surgical trials) of RCTs, less than half of these actually comparing differing types of surgical interventions (Wente et al., 2001).

Gynaecologic interventions have been exposed to the scrutiny of the RCTs more readily than in many other surgical specialities (Johnson et al., 2003, 2007). Indeed, subfertility was one of the first fields where the need to move from ‘cookery based practice to science’ was recognized (Vandekerckhove et al., 1993) and this has since become well established in infertility treatment. Has gynaecological surgery moved forward in the same way? With increasing general emphasis on evidence-based practice and the introduction of methodological initiatives such as the Consolidated Standards of Reporting Trials (CONSORT) statement (Begg et al., 1996) in particular, it would be expected that methodology of gynaecologic surgical trials should have improved over time, resulting in improvements in both reliability (precision) and validity of studies. With more improvements in quality, one can also expect to see a reduction in the size of effect as biasing factors that inflate effects are avoided (Schulz et al., 1995). Did RCTs evolve this way in gynaecologic surgery? The aim of this study was to assess if there has...
been progress made in establishing the evidence base for surgical interventions in gynaecology.

Materials and Methods

We developed a protocol to undertake a methodological review of all trials of gynaecological surgery included in Cochrane reviews. We assessed if surgical trials improved in methods of allocation concealment and precision over time. We hypothesized that as research in this speciality matures over time, the observed effect of surgical interventions will be reduced in size. Further, we tested if bias due to lack of adequate allocation concealment was responsible for exaggerated effect size.

Identification of systematic reviews

The Cochrane Database of Systematic Reviews (Issue 3, 2006) was used to identify relevant review groups in gynaecology (Gynaecological Cancer, Incontinence, Menstrual Disorders and Subfertility and Fertility Regulation). From these groups, titles of potentially relevant reviews of surgical interventions were obtained. Two reviewers (T.J.S. and N.P.J.) assessed all potentially relevant reviews to determine eligibility and data were extracted from individual RCTs using a predesigned proforma.

Data extraction

The Cochrane review group co-coordinators were requested to supply source files for the most up-to-date version of each of the selected reviews. From these files, data were extracted for the individual RCTs included in each review. The year of publication was recorded. If the trial was not published in peer reviewed literature, the year that the trial was first referenced in grey literature was used, and in the absence of this information, the year that the review authors allocated to the trial was recorded. A measure of the validity of the studies was made using the adequacy of allocation concealment as judged by the Cochrane reviewers. Allocation concealment was used because it is possible to achieve this in surgical trials even when blinding is not possible and this item is associated with the greatest magnitude of bias (Schulz et al., 1995). The number of participants in the study was recorded along with the effect size for the primary outcome. We defined a priori a hierarchy of primary outcomes (Johnson et al., 2007) to be extracted from each trial by consensus among the authors of this paper, based on what we considered to be outcomes of relevance to patients (Johnson et al., 2007) If such outcomes were not available, the primary outcomes used by Cochrane reviewers were agreed upon by consensus among authors of this paper. We allocated each study to one of the five clinical topics: urogynaecology, menstrual disorders (including surgery for benign gynaecologic conditions), infertility, fertility regulation and cancer. These subgroups provided a control variable in the regression models. For the measurement of effect size, we calculated the odds ratio (OR) from 2×2 tables or by converting the standardized mean difference into OR if required (Chinn, 2000). The numbers of conversions required constituted <5% of all comparisons within selected trials. Where the effect was not a beneficial outcome, we inverted the treatment effect in order that OR greater than 1.0 indicated benefit of surgical intervention. If the data to estimate effect size were not available from the Cochrane files, the trial was excluded from analysis (10 studies).

Statistical analysis

To test the hypothesis that quality of studies improved over time, we grouped studies according to a time cut-off of 1996 (year of the CONSORT statement publication), excluding those studies from 1996 to 1998 to allow reasonable time for CONSORT to be implemented. Chi-squared test was used to examine if the percentage of studies with adequate concealment of treatment allocation was higher after this time compared with before. For the analysis we use A (adequate allocation concealment) against the other concealment categories i.e. adequate allocation concealment against unclear, inadequate or concealment not used. We performed meta-regression analyses (Song et al., 2001) to test the hypothesis that effect size decreased over time. We used the log OR as the dependent variable and year of publication as independent variable. The analysis was weighted by inverse of log OR variance and dummy variables were included for review topic to control for the different effect in different surgical fields. We performed multiple linear regression analyses to test the hypothesis that precision increased over time using inverse of log OR variance as the dependent variable and year of publication as independent variable, adding dummy variables for review topic to control for the different sample sizes in different surgical fields.

We used logistic regression (Schulz et al., 1995; Moher et al., 1998; McAuley et al., 2000) model to estimate whether treatment effects differed between trials with and without allocation concealment. In the logistic models, we only included studies within reviews with the same intervention and primary outcome to avoid obvious problems with heterogeneity and, also, only meta-analyses that contained trials with and without the characteristic of interest, i.e. allocation concealment. To ensure that the OR is estimable using logistic regression, studies were only included in the analyses if no zeros were observed in the corresponding 2×2 tables. This model allowed the probability of the outcome event to vary according to treatment or control group, trial, methodological quality characteristic (i.e. allocation concealment). A term representing meta-analysis was also included in the model. The effect of allocation concealment on treatment effect was given by the treatment–concealment interaction term. This estimated the log of the ratio of ORs (ROR) in trials with and without allocation concealment. RORs>1.0 implied greater (more beneficial) treatment effects in the trials with adequate allocation concealment than in those without. We used robust SEs as this approach corrects the SE of the regression coefficients, taking into account the correlation with meta-analysis and produces conservative estimates of 95% confidence interval (CI) width.

All analyses were performed using STATA version 8.2. A value of P<0.05 was considered significant.

Results

In total, 23 relevant gynaecologic surgical reviews were identified that included 110 meta-analyses. We selected only one meta-analysis per review (that with the largest number of trials) in order to avoid using duplicated patients data from the same review. There were 94 trials (12 298 women) from which hypotheses concerning the effect of time on validity, precision, and effect size could be tested. The publication dates ranged from 1974 to 2005 (median 1998). The analyses concerning bias due to lack of allocation concealment included 13 reviews with 60 trials (8485 women). For this analysis, we excluded those meta-analyses that did not include trials with both characteristics (concealed and not concealed trials). We then further excluded 10 meta-analyses because all the trials in those were not concealed (Fig. 1).

Allocation concealment was reported in 42 out of the 94 selected studies. Before the CONSORT statement 10 out of
39 (25.6%) had allocation concealment compared with 21 out of 35 (60%) post-CONSORT ($P = 0.002$).

Figure 2 and Table I show that there was a similar trend towards improvement in precision over time and similarly in sample size (Fig. 3). Similarly, there was a reduction in size of treatment effect over time (ROR per year 0.96; 95% CI 0.93–0.99, $P = 0.04$) (Fig. 3 and Table I). This meant that trials published in the years after the first trial on a topic tended to be larger and gave less optimistic treatment effect. Regarding allocation concealment, consistent with current understanding of mechanisms of allocation concealment bias, estimated treatment effects were more beneficial in trials without allocation concealment, although this association did not reach statistical significance (ROR = 1.19; 95% CI 0.86–1.66; $P = 0.278$).

**Figure 1:** Study selection process for evaluation of quality and precision of gynaecologic surgical trials.

**Figure 2:** The trends in precision (reliability) and effect size over time among gynaecologic surgical trials.

**Table I.** Coefficients of the meta-regression models (with 95% confidence intervals) assessing the impact of publication year on precision [model 1 dependent variable was inverse of variance of log odds ratio, log (OR)] or on effect size [model 2 independent variable was log (OR)]

<table>
<thead>
<tr>
<th>Coefficient 95% confidence intervals</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: Precision improvement over time $^*$</td>
<td>0.12</td>
</tr>
<tr>
<td>Model 2: Effect size reduction over time $^*$</td>
<td>$-0.04$</td>
</tr>
</tbody>
</table>

$^*$Adjusted by review topic.  
$^*$Weighted by inverse of variance of log (OR).
Discussion

Our study shows an improvement in quality of surgical trials in gynaecology over time as the CONSORT guidelines boosted allocation concealment. The trials got larger in size (more reliable or more precise) over time. There was also a general trend towards a reduction in effect size as a topic matured over time.

The low percentage of systematic reviews that assess surgical interventions among the four collaborative review groups whose scope covers gynaecologic surgery reflects the difficulties with performing good surgical RCTs. Of particular concern is that there are only two systematic reviews in Gynaecologic Cancers Group’s database that include trials assessing surgical treatments (only one of which had RCTs meeting the entry requirements for our review). Although this doubtless, in part, reflects simply an absence of Cochrane systematic reviews in this surgical field, it highlights the difficulty with conducting trials in cancer surgery where patients (and their surgeons) may be reluctant to consent to randomizations even in the face of genuine equipoise. Thus, it would be argued that plugging gaps in the evidence base in the field of gynaecologic cancer surgery is of particularly high priority (Johnson et al., 2007).

Our study used robust systematic review and statistical techniques that have previously been validated (Schulz et al., 1995, Moher et al., 1998). We may be criticized because of the reliance on Cochrane reviews for data extraction alone. There has been concern that authors of Cochrane reviews may incorrectly classify allocation concealment. However, Cochrane reviews go through a rigorous peer review process and their quality has been shown to be better than non-Cochrane reviews. Hence, we are confident that our findings are robust. We have also limited our study to use allocation concealment as the sole measure of study quality, as this was the most consistently extracted quality item by Cochrane reviewers. However, there are other measures, such as blinding, that would make further interesting research.

It is evident from our review that gynaecologic surgical practice is benefiting from improvement in its research base, a trend which has been highlighted by the Cochrane Fertility Regulation Group, who also showed an improvement in RCT quality with the introduction of QUDAS, a checklist for quality items that should be included in studies of diagnostic test accuracy (Helmhorsst et al., 2006). This improvement may be described metaphorically as moving from ‘butchery’ to science in a subject where practitioners do not participate readily in randomized evaluation. It is also sobering to note that the improvements in research methodology are associated with less optimism about effects of gynaecologic surgery.

Conflict of interest: All authors state that they have nothing to declare.

Author’s Contribution

I, T.J.S, contributed to the conception of the review, performed the literature search, took part in the analysis and completion of the first draft and subsequent amendments. I have approved the final version and am guarantor.

I, Javier Zamora contributed to the concept of the review, lead the statistical analysis, contributed to the first draft and amendments. I have approved the final version.

I, N.P.J., contributed to the conception of the review, the literature search, took part in the analysis and contributed to the first draft and amendments. I have approved the final version.
I, K.S.K, contributed to the conception of the review, took part in the analysis and contributed to the first draft and amendments. I have approved the final version.

References


Submitted on September 1, 2007; resubmitted on November 1, 2007; accepted on December 4, 2007