Gynaecologic surgery from uncertainty to science: evidence-based surgery is no passing fad

N.P. Johnson\textsuperscript{1,4}, T. Selman\textsuperscript{2}, J. Zamora\textsuperscript{3} and K.S. Khan\textsuperscript{2}

\textsuperscript{1}Department of Obstetrics and Gynaecology, University of Auckland, National Women’s Health, Level 12, Auckland Hospital, Auckland, New Zealand; \textsuperscript{2}Department of Reproductive and Child Health, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK; \textsuperscript{3}Clinical Biostatistics Unit, Hospital Ramon y Cajal, Madrid, Spain

\textsuperscript{4}Correspondence address. Tel: +64-9-3737599 ext 89493; Fax: +64-9-3035969; E-mail: n.johnson@auckland.ac.nz

BACKGROUND: The randomized controlled trial (RCT) is the least biased measure of the effectiveness of interventions, including surgical interventions. The aim was to review the available evidence base in gynaecologic surgery, to assess what progress has been made and to determine gaps in the evidence for clinical decision-making. METHODS: Systematic reviews involving gynaecological surgery interventions were extracted from the Cochrane Database of Systematic Reviews (Issue 2, 2007) and data were extracted for key primary outcomes from each of the randomized trials in the reviews. The reviews were categorized as to whether they had provided evidence of effectiveness for pre-defined outcomes of most relevance to patients. RESULTS: Of 371 reviews or protocols published on the Cochrane Database of Systematic Reviews (Issue 2, 2007), only 30 were completed reviews assessing surgical interventions. Seven reviews concluded there was evidence of a significant effect (whether beneficial or harmful) of the interventions studied for pre-defined primary outcomes; 11 reviews concluded there was some evidence of significant effects for primary outcomes along with some gaps for primary outcomes; 12 reviews concluded insufficient evidence of effectiveness. Common themes of unique methodological challenges and pitfalls with trials of surgical interventions were apparent. CONCLUSIONS: Cochrane reviews have gone a long way to establishing a sound evidence base in gynaecologic surgery: some gaps in the evidence have been eliminated and others highlighted. In general, gynaecology has been a specialty where surgical interventions have been well exposed to the scrutiny of RCTs compared with other surgical specialties.

Keywords: Cochrane; gynaecology; randomized controlled trials; surgery; systematic review

Introduction

Gynaecology, following the lead of perinatology, is a specialty in which interventions have been exposed to the scrutiny of the randomized controlled trial (RCT) more readily than in many other specialties (Johnson \textit{et al}., 2003). Indeed, subfertility was one of the first fields where the need to base practice on evidence from robust randomized trials was highlighted, rather than to conduct a ‘cookery based’ approach to practice (Vandekerckhove \textit{et al}., 1993). Has a similar phenomenon occurred in surgical specialties, including ‘gynaecologic surgery from uncertainty to science’? Traditionally in surgical specialties, a non-evidence-based approach to practice has been prevalent. To call this ‘butchery based’ might be emotive, although no inference should be taken that the surgeons involved were not skilful, but there has been a tendency to adopt the latest surgical technique because it seems rational (or worse, because it demonstrates the technical skill of the surgeon) rather than because it fulfils the stringent criteria for effectiveness that we now demand for non-surgical interventions. Our view, for what it is worth, is that gynaecologists have been less guilty of this phenomenon than other surgical specialists. This is probably linked to the origins of evidence-based practice in the specialty of obstetrics and gynaecology (Johnson \textit{et al}., 2003), and our related study reports in detail on the evolution of methodological quality of the randomized trials to which the present study relates (Selman \textit{et al}., 2008).

The Cochrane Database of Systematic Reviews provides up-to-date evidence on health care, where bias is minimized. In addition, systematic reviews can also identify ‘gaps’ where there is insufficient or no evidence, or where the quality of evidence is insufficient. (Johnson \textit{et al}., 2003). Elaborate grading systems for evidence exist, although these have largely been designed for formulation of practice guideline recommendations (Guyatt \textit{et al}., 2006).
The aim of this study was to assess qualitatively to what extent the Cochrane review groups whose scope covers gynaecologic surgical interventions have been able to provide evidence coverage of this field and where gaps in the evidence remain.

Materials and Methods

The relevant gynaecology review groups (Gynaecological Cancer, Incontinence, Fertility Regulation, Menstrual Disorders and Subfertility Groups) with reviews on the Cochrane Database of Systematic Reviews (Issue 2, 2007) were searched for titles suggesting assessment of a surgical intervention. After confirmation that trials did indeed assess a gynaecologic surgical technique, each review was subgrouped as oncology, urogynaecology, fertility regulation, menstrual disorders, subfertility or other benign gynaecologic surgery.

We had defined \textit{a priori} a hierarchy of primary outcomes by consensus among the authors of this paper, to be extracted from each systematic review (Selman \textit{et al}, 2008), based on what we considered to be outcomes of relevance to patients, as follows.

(i) For oncology reviews: disease eradication;

(ii) For incontinence reviews: subjective ‘cure’ or dryness (objective ‘cure’ if not available) in reviews of incontinence; for prolapse reviews: subjective ‘cure’ or prevention of recurrent prolapse (objective if not available);

(iii) For fertility regulation reviews: uncomplicated termination of pregnancy or attainment of sterilization;

(iv) For menstrual disorders reviews: subjective ‘cure’ of heavy menstrual bleeding (HMB) or satisfaction with treatment (objective reduction of HMB if not available);

(v) For subfertility reviews: live birth (or clinical pregnancy if data on live birth were not available);

(vi) For benign gynaecologic surgery, reviews were assessed and a primary outcome allocated \textit{a priori} as follows:

1. return to normal activities (in the absence of extensive quality of life data) for hysterectomy;
2. live birth (or pregnancy) or pain (or recurrent adhesions if not available) for adhesion prevention agents;
3. elimination of ectopic pregnancy;
4. pain for endometriosis surgery (including endometriomas) and neuroablation.

If such outcomes were not available, the primary outcomes selected by review authors were considered and a primary outcome agreed upon by consensus among authors of this paper.

The following data were collected for each review:

(i) date of the most recent search for trials;

(ii) the number of trials and participants contributing to the primary outcomes;

(iii) whether there was evidence of a significant difference (whether beneficial or harmful) of the interventions studied for pre-specified primary outcomes from meta-analysis (evidence category allocation ‘E’ for ‘evidence of an effect’), or simply insufficient evidence of effectiveness or harm (evidence category allocation ‘G’ for ‘gap in evidence’). The term ‘relative effectiveness’ was used when two interventions were compared and the term ‘effectiveness’ was used when the treatment was compared with either placebo or no treatment. Evidence category ‘E&G’ was allocated when there was some evidence of significant effects along with some gaps for primary outcomes.

(iv) Methodological difficulties highlighted by the reviewers were also noted.

Results

Thirty completed reviews assessing surgical interventions, from a total of 371 reviews and protocols in gynaecologic oncology, urogynaecology, fertility regulation, menstrual disorders and subfertility, were published on the Cochrane Library. These comprised:

(i) two of 84 reviews ($n = 47$) and protocols ($n = 37$) from the Gynaecological Cancer Group (Ansink and van der Velden, 2000), only one of which included any RCTs (Martin-Hirsch \textit{et al}, 2000);

(ii) six of 69 reviews ($n = 51$) and protocols ($n = 18$) from the Incontinence Group (Glazener and Cooper, 2001; Moehrer \textit{et al}, 2002; Maher \textit{et al}, 2004; Bezerra \textit{et al}, 2005; Lapitan \textit{et al}, 2005);

(iii) four of 53 reviews ($n = 46$) and protocols ($n = 7$) from the Fertility Regulation Group (Kulier \textit{et al}, 2001; 2004; Nardin \textit{et al}, 2003; Say \textit{et al}, 2005);


The findings of these reviews of gynaecologic surgery and our extracted conclusions are summarized in Tables I–VI. Overall, from the 30 gynaecologic surgery reviews, there was evidence of effectiveness (or relative effectiveness) of the interventions from primary trial data or from meta-analysis of trial data for our \textit{a priori} defined primary outcomes in seven reviews where clear answers to the clinical questions posed were found (evidence category E) and in a further 11 reviews where evidence of effectiveness for some comparisons or outcomes were found in conjunction with some comparisons or outcomes for which there was insufficient evidence (evidence category E&G). There was insufficient evidence (or a ‘gap’ in evidence) of effectiveness in 12 reviews (evidence category G).

Table VII shows a practical categorization for the various interventions assessed in the 30 gynaecologic surgery reviews. On the basis of this system, 19 recommendations would be that the interventions are likely to show benefit, with 21 recommendations that interventions are likely to be harmful or ineffective (with some duplication from the previous 19 recommendations in the case of interventions whose relative efficacy was found to be inferior to another intervention) and only 13 interventions where the interventions are of unknown benefit.
For Tables I–VI, figures in brackets refer to individual comparisons.
Category E, ‘Evidence’ of effectiveness from meta-analysis; Category G, insufficient evidence (or ‘gap’ in the evidence); Category E&G, combination of evidence of effectiveness for one primary outcome and a gap in the evidence for another primary outcome; E, ‘Evidence’ of effectiveness from RCTs; G, ‘Gap’ in evidence from RCTs; LLETZ, large loop excision of the transformation zone; RCT, randomised controlled trial.

### Table I. Cochrane reviews of gynaecologic cancer surgery.

<table>
<thead>
<tr>
<th>Review title</th>
<th>Included RCTs</th>
<th>Last update</th>
<th>Conclusions for primary outcomes (evidence category E, G or E&amp;G)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surgical interventions for early squamous cell carcinoma of the vulva (Ansink and van der Velden, 2000)</td>
<td>0</td>
<td>October 1999</td>
<td>Two observational studies only (G)</td>
</tr>
<tr>
<td>Surgery for cervical intraepithelial neoplasia (Martin-Hirsch et al., 2000)</td>
<td>28</td>
<td>July 1999</td>
<td>No significant differences in disease eradication for seven techniques (knife cone biopsy, laser conization, LLETZ, laser ablation, cryotherapy: single and double freeze, radical diathermy) (G)</td>
</tr>
</tbody>
</table>

Discussion

Systematic reviews of gynaecologic surgical interventions are under-represented in Cochrane review groups whose scope covers the breadth of gynaecology. This probably reflects the methodological difficulties with conducting randomized trials of surgical interventions. The surprisingly small number of surgical RCTs was most striking in gynaecologic cancer treatment, in which a strong evidence base for techniques of surgical removal of cancer might be expected. This may be due in part to a reluctance of patients with cancer to submit to randomization, difficulties with obtaining ethical approval where a patient with cancer could be randomized to not having the cancer surgically removed, and a discomfort among gynaecologic oncology surgeons to admit to being in equipoise. We have disproved the widely held belief that a clear answer to a clinical dilemma is seldom the case from RCTs or meta-analysis of RCT results, in our finding that a reasonable proportion (60%) of all Cochrane reviews of gynaecologic surgical interventions were able to find evidence of effectiveness or superior relative effectiveness of these interventions for at least some primary outcomes—this figure was 32% for a similar analysis in Cochrane subfertility reviews in 2003 (Johnson et al., 2003). Some RCTs assessing surgical interventions will have been excluded from this study because they have not yet been included in Cochrane reviews, but such evidence may remain difficult for the average clinician to reach (as there is currently much more relevant literature than an individual clinician can reasonably expect to absorb).

The approach we have used to summarize the level of evidence provided for an intervention (Table VII) can be utilized for guiding practice in gynaecologic surgery. For strong evidence, where the benefits clearly outweigh the risks, patients will commonly make the same choices and surgeons can recommend the intervention with confidence. For weak evidence, patient choice will differ and here surgeons must communicate the evidence with a particular emphasis on patients’ individual needs. An example of one such recently published guideline is for laparoscopic uterosacral nerve ablation for the treatment of chronic pelvic pain (http://guidance.nice.org.uk/IPG234). However, our objective was to evaluate the evidence rather than to generate guidelines, this doubtless being the role of authoritative bodies, such as specialty colleges and departments of health.

Pitfalls with systematic reviews have been highlighted (Farquar and Vail, 2006). In common with trials and systematic reviews of medical interventions, clinical trials and systematic reviews of RCTs in gynaecologic surgery are prone to difficulties with study quality, funding bias, publication bias, reliance on outcomes of little help in clinical decision-making, analysis errors and incorrect use of evidence statements in conclusions. However, surgical trials and systematic reviews have their own unique pitfalls, including a lower threshold for limitations to completion of trials, more limitations to broad applicability of trial results, surgical reputation conflict of interest, in addition to design problems, such as the performance bias resulting from inability to employ blinding (especially concerning subjective outcomes).

There are undoubtedly more confounding variables in surgical RCTs than in the more straightforward A versus B comparison that RCTs address in assessing the effectiveness of medical interventions, including variation in expertise of surgeons with different operations leading to an almost unavoidable confounding surgeon effect. For example, in the largest RCT of laparoscopic versus abdominal or vaginal hysterectomy (Garry et al., 2004), surgeons were required to have performed only 20 laparoscopic hysterectomies prior to participation in the trial; a similar problem occurred in the first RCT comparing laparoscopic versus open colposuspension (Burton, 1999), where there was a requirement for surgeons to have performed only 15 laparoscopic colposuspensions (an operation requiring the highest level of laparoscopic surgical expertise) prior to participation, although they would probably have vast experience of the more traditional open surgical approach. It is recognized that the learning curve for these complex laparoscopic surgical procedures may be exceptionally long, where the ‘first couple of hundred’ advanced laparoscopic procedures appear to be more hazardous than ‘the next thousand’ cases in any surgeon’s series (Johnson, 2006). Conversely, surgeons participating in trials may be enthusiasts or innovators, so they may not be representative of all surgeons. Major adverse events in surgery are fortunately rare, but RCTs are therefore not often large enough to detect rare adverse events. For example, the ureteric injury rate in the systematic review of RCTs of women randomized to laparoscopic hysterectomy was 1 in 88, compared with 1 in 512 women randomized to abdominal hysterectomy, but this difference fell short of statistical significance (Johnson et al., 2005).
These difficulties with surgical RCTs have led some to suggest that RCTs have little or no place in the evaluation of surgical interventions (Black, 1999), implying that factors other than RCT evidence, particularly training and expertise among surgeons, patient preferences or even cost of treatments, are more important in determining surgical approach. It would, however, in our opinion, present some danger to introduce new surgical procedures without RCT scrutiny of these new operations versus the current gold standard treatments, as the RCT is the most reliable indicator of the effectiveness of an intervention, whether medical or surgical. Of course, RCT evidence must be integrated with individual surgical expertise in evidence-based surgical practice. The other types of study design (such as case series, even with very large numbers, cohort studies or case-controlled studies), so often relied upon to assess surgical interventions, are prone to unacceptable bias.

A number of prerequisites for surgical RCTs will vastly improve the evidence base in the future. First, pragmatic trials with non-restrictive entry criteria will improve recruitment and generalizability of trial results. Second, only when we collaborate in large multi-centre RCTs of gynaecologic surgical interventions will we have sufficient evidence to assess the impact of prolapse surgery on continence issues, although addition of tension free vaginal tape to endopelvic fascia placation is effective in reducing postoperative stress incontinence (E&G).

### Table II. Cochrane reviews of urogynaecologic surgery.

<table>
<thead>
<tr>
<th>Review title</th>
<th>Included RCTs</th>
<th>No. of participants</th>
<th>Last update</th>
<th>Conclusions for primary outcomes (evidence category E, G, or E&amp;G)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anterior vaginal repair for urinary incontinence in women (Glazener and Cooper 2001)</td>
<td>9</td>
<td>932</td>
<td>January 2001</td>
<td>Anterior vaginal repair is less effective than open abdominal retropubic suspension operations. Has insufficient evidence of effectiveness versus physical therapy. No RCTs examined anterior vaginal repair versus suburethral sling operations. Laparoscopic colposuspension.</td>
</tr>
<tr>
<td>Bladder neck needle suspension for urinary incontinence in women (Glazener and Cooper, 2004)</td>
<td>8</td>
<td>784</td>
<td>April 2004</td>
<td>Needle suspension is less effective than open abdominal retropubic suspension. Has insufficient evidence of effectiveness versus suburethral slings. No RCTs examined needle suspension versus conservative management. Periurethral injections. Sham surgery. Laparoscopic surgery (E&amp;G).</td>
</tr>
<tr>
<td>Traditional suburethral sling operations for urinary incontinence in women (Bezerra et al., 2005)</td>
<td>13</td>
<td>760</td>
<td>July 2005</td>
<td>No significant differences were found between slings and open abdominal retropubic suspension. Needle suspension. Nor among different types of suburethral slings, but slings are more effective than anticholinergic medication. No RCTs compared slings with laparoscopic retropubic suspension, anterior repair, peri-urethral injections or artificial sphincters (G). Laparoscopic colposuspension versus open colposuspension. Gives similar subjective impression of cure at 18 months versus open colposuspension. Is less effective at ‘cure’ on urodynamic outcomes. Laparoscopic colposuspension versus newer self-fixing sling devices. Gives similar subjective cure rates. Two paravaginal sutures are more effective than a single suture in improving subjective cure rates (E&amp;G). Abdominal sacral colpopexy is more effective than vaginal sacrospinous colpopexy in reducing recurrent vault prolapse and dyspareunia. Supplementation with Vicryl mesh overlay is effective in reducing recurrent cystocele at anterior vaginal repair. Vaginal repair is more effective than transanal repair for posterior vaginal wall prolapse. Insufficient evidence to assess impact of prolapse surgery on continence issues, although addition of tension free vaginal tape to endopelvic fascia placation is effective in reducing postoperative stress incontinence (E&amp;G).</td>
</tr>
<tr>
<td>Laparoscopic colposuspension for urinary incontinence in women (Moehret et al., 2002)</td>
<td>21</td>
<td></td>
<td>July 2006</td>
<td>Laparoscopic colposuspension versus open colposuspension. Gives similar subjective cure rates. Is less effective at ‘cure’ on urodynamic outcomes. Laparoscopic colposuspension versus newer self-fixing sling devices. Gives similar subjective cure rates. Two paravaginal sutures are more effective than a single suture in improving subjective cure rates (E&amp;G). Abdominal sacral colpopexy is more effective than vaginal sacrospinous colpopexy in reducing recurrent vault prolapse and dyspareunia. Supplementation with Vicryl mesh overlay is effective in reducing recurrent cystocele at anterior vaginal repair. Vaginal repair is more effective than transanal repair for posterior vaginal wall prolapse. Insufficient evidence to assess impact of prolapse surgery on continence issues, although addition of tension free vaginal tape to endopelvic fascia placation is effective in reducing postoperative stress incontinence (E&amp;G).</td>
</tr>
<tr>
<td>Surgical management of pelvic organ prolapse in women (Maher et al., 2004)</td>
<td>14</td>
<td>1004</td>
<td>October 2004</td>
<td>Surgical management of pelvic organ prolapse in women (Maher et al., 2004).</td>
</tr>
</tbody>
</table>
**Table III.** Cochrane reviews of fertility regulation surgery.

<table>
<thead>
<tr>
<th>Review title</th>
<th>Included RCTs</th>
<th>No. of participants</th>
<th>Last update</th>
<th>Conclusions for primary outcomes (evidence category E, G or E&amp;G)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surgical methods for first trimester termination of pregnancy (Kulier et al., 2001)</td>
<td>3</td>
<td>767</td>
<td>October 2001</td>
<td>No significant differences for vacuum aspiration versus dilatation and curettage (G)</td>
</tr>
<tr>
<td>Medical versus surgical methods for first trimester termination of pregnancy (Say et al., 2005)</td>
<td>6</td>
<td>1002</td>
<td>October 2002</td>
<td>Abortion is completed more often and is less painful for suction termination of pregnancy versus prostaglandins (E)</td>
</tr>
<tr>
<td>Techniques for the interruption of tubal patency for female sterilization (Nardin et al., 2003)</td>
<td>9</td>
<td>4453</td>
<td>October 2002</td>
<td>No differences in failure rates between ring versus clip, Pomeroy versus electrocoagulation (G)</td>
</tr>
<tr>
<td>Minilaparotomy and endoscopic techniques for tubal sterilization (Kulier et al., 2004)</td>
<td>6</td>
<td>2601</td>
<td>July 2004</td>
<td>Morbidity significantly higher for culdoscopy versus both minilaparotomy and laparoscopic approaches</td>
</tr>
<tr>
<td>Surgical methods for first trimester termination of pregnancy (Lethaby et al., 2005)</td>
<td>19</td>
<td>3285</td>
<td>October 2001</td>
<td>No significant differences for vacuum aspiration versus dilation and curettage (G)</td>
</tr>
<tr>
<td>Endometrial resection and ablation versus hysterectomy for heavy menstrual bleeding (Lethaby et al., 2000)</td>
<td>5</td>
<td>752</td>
<td>April 1999</td>
<td>Hysterectomy has significantly improved effectiveness and satisfaction versus endometrial ablation, at a cost of increased adverse events (E)</td>
</tr>
<tr>
<td>Surgery versus medical therapy for heavy menstrual bleeding (Majoribanks et al., 2006)</td>
<td>8</td>
<td>821</td>
<td>April 2006</td>
<td>Surgery, especially hysterectomy, is more effective than medical treatments at reducing menstrual loss at 1 year, but there are no significant differences in quality of life improvement between surgery and the levonorgestrel intrauterine system (E&amp;G)</td>
</tr>
</tbody>
</table>

**Table IV.** Cochrane reviews of surgery for menstrual disorders.

<table>
<thead>
<tr>
<th>Review title</th>
<th>Included RCTs</th>
<th>No. of participants</th>
<th>Last update</th>
<th>Conclusions for primary outcomes (evidence category E, G or E&amp;G)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endometrial destruction techniques for heavy menstrual bleeding (Lethaby et al., 2005)</td>
<td>19</td>
<td>3285</td>
<td>October 2001</td>
<td>Insufficient evidence of differences in effectiveness between different endometrial destruction techniques (G)</td>
</tr>
<tr>
<td>Endometrial resection and ablation versus hysterectomy for heavy menstrual bleeding (Lethaby et al., 2000)</td>
<td>5</td>
<td>752</td>
<td>April 1999</td>
<td>Hysterectomy has significantly improved effectiveness and satisfaction versus endometrial ablation, at a cost of increased adverse events (E)</td>
</tr>
<tr>
<td>Minilaparotomy and endoscopic techniques for tubal sterilization (Kulier et al., 2004)</td>
<td>6</td>
<td>2601</td>
<td>July 2004</td>
<td>Morbidity significantly higher for culdoscopy versus both minilaparotomy and laparoscopic approaches</td>
</tr>
</tbody>
</table>

**Table V.** Cochrane reviews of subfertility surgery.

<table>
<thead>
<tr>
<th>Review title</th>
<th>Included RCTs</th>
<th>No. of participants</th>
<th>Last update</th>
<th>Conclusions for primary outcomes (evidence category E, G or E&amp;G)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Techniques for pelvic surgery in subfertility (Ahmad et al., 2006)</td>
<td>7</td>
<td>437</td>
<td>April 2006</td>
<td>No evidence of superior relative effectiveness of carbon dioxide laser versus standard techniques in adhesiolysis and salpingostomy (G)</td>
</tr>
<tr>
<td>Laparoscopic surgery for subfertility associated with endometriosis (Jacobson et al., 2002)</td>
<td>2</td>
<td>437</td>
<td>October 2002</td>
<td>Laparoscopic surgery significantly improves the odds of live birth plus ongoing pregnancy (E)</td>
</tr>
<tr>
<td>Surgical treatment for tubal disease in women due to undergo IVF (Johnson et al., 2004)</td>
<td>3</td>
<td>295</td>
<td>July 2004</td>
<td>Laparoscopic salpingectomy for hydrosalpinges prior to IVF significantly improves the odds of pregnancy and live birth (E)</td>
</tr>
<tr>
<td>Laparoscopic drilling by diathermy or laser for ovulation induction in anovulatory polycystic ovary syndrome (Farquhar et al., 2005)</td>
<td>6</td>
<td>313</td>
<td>July 2005</td>
<td>No significant difference in pregnancy rates between laparoscopic ovarian drilling (6–12 months follow-up) and gonadotrophin injections (3–6 cycles); significantly fewer multiple pregnancies with ovarian drilling (G)</td>
</tr>
<tr>
<td>Postoperative procedures for improving fertility following pelvic reproductive surgery (Johnson and Watson, 1999)</td>
<td>5</td>
<td>608</td>
<td>April 2007</td>
<td>No evidence of effectiveness for postoperative hydrotubation (G)</td>
</tr>
<tr>
<td>Techniques for surgical retrieval of sperm prior to ICSI for azoospermia (Van Peperstraten et al., 2006)*</td>
<td>2</td>
<td>98</td>
<td>April 2005</td>
<td>Insufficient evidence to recommend any particular technique for surgical retrieval of sperm over another (G)</td>
</tr>
</tbody>
</table>

*Surgical procedures performed on men by fertility specialists.
power to find modest improvements that may add up to clinically meaningful improved quality of life, more babies born to infertile couples or even lives saved. Such an approach enhances generalizability and speeds recruitment, thus avoiding the problem seen with laparoscopic colposuspension, an operative intervention that became
almost obsolete before the evidence base was established, owing to the rapid progress that is typical of most surgical fields. A third key is the pursuit of important long-term outcomes, which are rarely reported in surgical RCTs. For example, no long-term outcomes were reported in 23 RCTs of surgical approach to hysterectomy (Johnson et al., 2005), including satisfaction, quality of life or long-term urinary, bowel or fistula complications. A further difficulty is the masking of long-term outcomes in the pursuit of less relevant surrogate outcomes—the many RCTs on adhesion prevention usually include a second look laparoscopy to assess adhesions, often combining this with adhesiolysis, thereby confounding the true primary outcomes of interest, either subsequent pregnancy outcome or pain outcomes (Farquhar et al., 2000). Fourth, recognition and acknowledgement of the limitations of evidence-based surgery by its proponents will go a long way to ensuring it is embraced by the majority of thoughtful practicing gynaecologic surgeons.

In conclusion, gynaecologic surgery has moved from a non-scientific to an evidence-based approach, but gaps in the evidence supporting gynaecologic surgical interventions have been highlighted. Notwithstanding the difficulties with methodology and conduct of RCTs of surgical interventions, we
need to organize well-powered multi-centre trials of surgical interventions—in doing so, we will move completely away from a non-scientific to a scientific basis for surgery.

Author’s Contribution
I, N.P.J., 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, T.S., contributed to the concept of the review, took part in the literature search, contributed first draft and amendments. I have approved the final version.
I, J.Z., 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.
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 November 15, 2007.

839