

Availability of evidence and comparative effectiveness for surgical versus drug interventions: an overview of systematic reviews

Emmanuel A. Zavalis^{1,2*}, Anaïs Rameau^{3*}, Anirudh Saraswathula^{4*}, Joachim Vist¹, Ewoud Schuit^{5,6}, John P. A. Ioannidis^{2,7}

*co-first authors

1 Department of Learning Informatics Management and Ethics, Karolinska Institutet, Stockholm, Sweden

2 Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, CA, USA

3 Sean Parker Institute for the Voice, Department of Otolaryngology–Head and Neck Surgery, Weill Cornell Medical College, New York, NY, USA

4 Department of Otolaryngology–Head and Neck Surgery, Johns Hopkins University School of Medicine, Baltimore, MD, USA

5 Julius Center, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

6 Cochrane Netherland, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

7 Stanford Prevention Research Center, Department of Medicine, and Department of Epidemiology and Population Health, Stanford University School of Medicine, Stanford, CA, USA

Corresponding author:

John P.A. Ioannidis, MD, DSc

Medical School Office Building, 1265 Welch Rd, Room X306, Stanford, CA 94305, USA

E-mail: jioannid@stanford.edu

Word count: 2894

Abstract

Objectives. To examine the prevalence of comparisons of surgery to drug regimens, the strength of evidence of such comparisons, and whether surgery or the drug intervention was favored.

Design. Systematic review of systematic reviews (umbrella review)

Data sources. Cochrane Database of Systematic Reviews (CDSR)

Eligibility criteria and synthesis of results. Using the search term “surg*” in CDSR, we retrieved systematic reviews of surgical interventions. Abstracts were subsequently screened to find systematic reviews that aimed to compare surgical to drug interventions; and then, among them, those that included any randomized controlled trials (RCTs) for such comparisons. Trial results data were extracted manually and synthesized into random-effects meta-analyses.

Results. Overall, 188 systematic reviews intended to compare surgery versus drugs. Only 41 included data from at least one RCT (total, 165 RCTs with data) and covered a total of 103 different outcomes of various comparisons of surgery versus drugs. A GRADE assessment was performed by the Cochrane reviewers for 87 (83%) outcomes in the reviews, indicating the strength of evidence was high in 4 outcomes (4%), moderate in 22 (21%), low in 27 (26%) and very low in 33 (32%). Based on 95% confidence intervals, the surgical intervention was favored in 38/103 (37%), and the drugs were favored in 13/103 (13%) outcomes. Of the outcomes with high GRADE rating, only one showed conclusive superiority (sphincterotomy was better than medical therapy for anal fissure). Of the 22 outcomes with moderate GRADE rating, 6 (27%) were inconclusive, 14 (64%) were in favor of surgery, and 2 (9%) were in favor of drugs.

Conclusions. Though the relative merits of surgical versus drug interventions are important to know for many diseases, high strength randomized evidence is rare. More randomized trials comparing surgery to drug interventions are needed.

Protocol registration. <https://osf.io/p9x3j>

Funding. The work of John Ioannidis has been funded by an unrestricted gift from Sue and Bob O'Donnell. Anaïs Rameau is supported by a Paul B. Beeson Emerging Leaders Career Development Award in Aging (K76 AG079040) from the National Institute on Aging and by the Bridge2AI award (OT2 OD032720) from the NIH Common Fund. Anirudh Saraswathula is supported by the National Institute on Deafness and Other Communication Disorders training grant 2T32DC000027.

Financial disclosure. Anaïs Rameau is a medical advisor for Perceptron Health, Inc.

Summary boxes

Section 1: What is already known on this topic

- Many conditions and diseases can be managed either with surgery or with drugs.

Comparative effectiveness of different treatment options is important to know for shared decision-making.

Section 2: What this study adds

- An assessment of the entire Cochrane Database of Systematic Reviews found 188 reviews that intended to assess surgical versus drug interventions, but only 41 had at least one randomized trial.
- Only four of the 103 assessed outcomes had high strength of evidence according to GRADE assessments.
- More evidence is needed to compare the relative merits of surgical and drug interventions and sequestration of these major modes of interventions should be overcome in clinical trial agendas.

Introduction

Many diseases are treated or managed with surgery. Some may also be addressed by pharmaceutical interventions, and studying the effectiveness of these different interventions is important in optimizing shared decision-making for patients and physicians. However, the amount and certainty of the evidence we hold in healthcare is limited[1], and this situation might be worse for surgical interventions due to serious challenges in running placebo-controlled or comparative effectiveness trials[2]. Challenges to controlled trials include unique patient anatomy, operator dependent variables such as the skill or experience of the surgeon[3–5], and the difficulty of successful blinding[6]. Due to these challenges, randomized controlled trials (RCTs) in surgery are less common than in non-surgical medical specialties. Although there have been calls to strengthen the quality of the evidence in surgery[2,7,8], these challenges have resulted in relatively few RCTs assessing surgical interventions, particularly in comparison to medical treatments.

A summary of the existing body of evidence on surgical versus medical interventions across diseases does not exist in the literature. To synthesize this existing body of evidence is of paramount importance to evidence-based care and informed decisions in the clinic where surgery or drugs are available interventions. To find RCTs comparing surgical vs. pharmaceutical interventions, we set up to conduct an umbrella review (an overview of systematic reviews) [9,10] by searching the Cochrane Database of Systematic Reviews for reviews considering comparisons of surgery to drugs, analyze the strength of the evidence and evaluate results of these comparisons. Finally, we explored whether results favoring surgery were more likely to be published in the surgical literature.

Methods

Search strategy and selection criteria

We queried the Cochrane Database of Systematic Reviews using the term “surg*” in “Title/Abstract/Keywords” on April 25, 2022. Inclusion criteria for reviews were consideration of RCTs and comparing a surgical to a drug intervention.

A surgical intervention was defined as a procedural technique aiming to change anatomy to treat or alleviate a pathology or symptom (including dermatological excisions). We excluded endoscopic and endovascular procedures since many of them are performed by medical rather than surgical specialists. A drug intervention was defined as a treatment that utilized a non-supplement and non-vitamin, pharmaceutical agent. Dental procedures, radiation treatment, as well as comparisons of surgery vs. no treatment or only placebo were excluded from our study. Cochrane reviews that intended to compare surgical and pharmaceutical interventions were considered even in cases where the review was unsuccessful in finding any such comparisons.

As many surgical procedures also require drug regimens (e.g., pre-operatively or as background treatment), we allowed comparisons where the surgical arm including a drug intervention was compared to a drug intervention as well. Comparisons of surgery to surgery plus drugs were not eligible, as both arms used surgery.

The articles' abstracts were reviewed by EAZ, and JV who coded the reviews independently for eligibility and then sought to reach a consensus. Remaining differences were mediated by JPAI.

Main outcomes

The main outcomes assessed in this umbrella review were the number of Cochrane systematic reviews that considered comparisons of surgical and drug interventions, and the number of systematic reviews that found any eligible RCTs comparing a surgical and a drug arm. The strength of evidence of the existing comparison was also treated as a main outcome, as were the direction of effects in the review assessments, both in the original Cochrane analysis and our standardized re-analysis).

Data extraction

EAZ extracted data for the included systematic reviews. In the cases where no data were available from any RCT, the review was tabulated into the category of reviews that found no evidence from RCTs on comparisons of surgical versus drug treatments.

The systematic reviews were classified into their corresponding surgical specialty field: cardiac surgery, dermatology, general surgery, neurosurgery, obstetrics and gynaecology, ophthalmology, orthopaedic surgery, otolaryngology, plastic surgery, thoracic surgery, urology and vascular surgery.

Whenever data were available from at least one RCT comparing a surgical to a drug arm, we identified the primary outcome(s) of the systematic review for the eligible comparison(s) by examining the methods section of the systematic review. Data, in the form of contingency tables or means, standard deviations and number of participants in each arms, from individual RCTs were then collected from Cochrane eligible reviews. The primary outcomes were classified as mortality, composite, and non-mortality. We also collected information on GRADE assessments for the eligible comparisons and outcomes and the summary effect size as well as the 95% confidence interval of the effect for the eligible comparison outcomes.

Meta-analysis

Cochrane reviewers may have used different statistical models in each topic to combine the results of RCTs in meta-analyses. To achieve standardization, we recalculated the summary effect size and heterogeneity for each topic using a random effects model following the Hartung-Knapp-Sidik-Jonkman approach[11,12] so that all outcomes/topics would be analyzed with the same statistical methods. The modified Haldane-Anscombe continuity correction was used, i.e. when studies had no event in either the surgical or the drug arm we added 0.5 to the entire contingency table of the specific study[13].

The analysis of the data was performed using R version 4.1.3 (2022-03-10), with assessment of statistical significance being using a threshold for α of 0.005, as previously proposed[14]. The Wilson approach was used for confidence intervals (99.5%) created for the primary outcomes.

Additions to the protocol

The original pre-registered protocol can be found in <https://osf.io/p9x3j>. Some additions were made during the process of conducting this umbrella review. For each systematic review, we noted the search date to understand how old they might be. We also extracted the year of

publication of each RCT to capture how recent the evidence was. Finally, we extracted the specialty orientation of the journal, in which the RCT was published, using the categories “mostly surgical”, “general”, and “mostly non-surgical”. The category “mostly surgical” includes those journals that have "surgery" in their title, those that have the name of a surgical specialty in their title, and those affiliated with a surgical society. The category "general" pertains to journals that cover all of medicine and its specialties, surgical and non-surgical. The category "mostly non-surgical" includes all the remaining journals. We assessed whether the direction of effects (favoring surgery or favoring drug) was associated with the type of journal, hypothesizing that RCTs published in mostly surgical journals may be more likely than other journals to favor surgery.

Results

Search results

The selection flowchart for Cochrane systematic reviews is represented in Figure 1. The search strategy retrieved 2495 articles from the Cochrane Database of Systematic Reviews. Among them, 440 were excluded by an automated search for withdrawn reviews and of studies with no mention of the word surgery and any of its variations in the abstract. Further manual inspection of titles and abstracts resulted in 223 Cochrane reviews being potentially eligible. Upon full-text evaluation, 35 were excluded: in 5 reviews, the surgical and drug treatments were not in separate arms and hence they were not an eligible head-to-head comparison[15–19]; in 7 reviews, there was no surgical intervention arm[20–26]; in 17 reviews, there was no drug intervention [27–32,32–42]); 2 reviews were excluded for evaluating an endoscopic intervention [43,44]; 3 reviews were excluded for evaluating an endovascular intervention [45–47]; and finally 1 review was excluded for being an umbrella review[48].

Therefore, 188 Cochrane reviews were found to meet the inclusion criteria (Supplement 1). Of those, 147 Cochrane reviews aimed to investigate surgical versus drug interventions but were unable to find any RCTs meeting their selection criteria. The remaining 41 reviews contained data for at least one RCT in at least one head-to-head comparison of a surgical versus a drug intervention arm (22% (99.5% CI 14 to 31%)).

The 188 reviews covered all major surgical specialties (Supplementary Table 1), with the most commonly represented specialties being general surgery (n=35), obstetrics and gynecology (n=31), ophthalmology (n=25), orthopedic surgery (n=23) and otolaryngology (n=23). No significant difference was found across specialties in the proportion of reviews that contained data from at least one RCT for a surgery versus drug comparison (Fisher's exact p=0.62).

Comparative treatment effect for surgery versus drug comparisons

The 41 eligible reviews with data included 103 comparisons of surgery versus drug treatments with data on various primary outcomes (Table 1), and they included data from a total of 165 RCTs with a total of 295 primary outcome assessments. For the 165 trials, the median

publication year was 2005 and the interquartile range was 1994 to 2016. The median search date year of the eligible reviews was 2016 (interquartile range, 2010 to 2022).

Based on the 95% confidence interval of the summary estimate obtained by the Cochrane review authors, surgery was more effective in 36 of the 103 outcomes of various comparisons (35% (99.5% CI 23 to 49%)), and drugs were more effective in 15 (15% (99.5% CI 6 to 26%)). Fifty two (50% (99.5 CI% 37 to 64%)) outcomes were inconclusive. The respective numbers were 1/12 (8%), 1/12 (8%), and 10/12 (83%) for mortality outcomes; 3/11 (27%), 3/11 (27%) and 5/11 (46%) for composite outcomes; and 32/80 (40%), 11/80 (14%), and 37/80 (46%) for non-mortality outcomes.

When we standardized the meta-analyses to use the same random effects method for all analyses, surgery was favored in 28/103 outcomes (32%), drugs were favored in 9 (10%) outcomes and 66 (58%) outcomes were inconclusive. The respective numbers were 1/12 (8%), 0/12 (0%), and 11/12 (92%) for mortality outcomes; 3/11 (18%), 2/11 (27%) and 6/11 (55%) for composite outcomes; and 24/80 (30%) 7/80 (9%), and 49/80 (61%) for non-mortality outcomes.

Table 2 shows the topics for which the surgical intervention was found to be more effective and Table 3 shows those where the drug arm was found to be more effective, all according to the Cochrane authors' analysis. Supplementary Table 2 does the same for the topics for which the comparisons were inconclusive.

Strength of evidence according to GRADE

GRADE assessment of the strength of the evidence showed high rating for 4 outcomes (4%), moderate for 22 (21%), low for 27 (26%), and very low for 33 (32%). No GRADE assessment was performed for 17 (17%) outcomes.

According to GRADE assessments, only cardiac surgery, obstetrics and gynecology and general surgery interventions had high GRADE ratings. Otolaryngology and dermatology had many moderate ratings. Almost all other GRADE ratings were low or very low (Table 4).

Of the four outcomes with high GRADE rating, sphincterotomy for anal fissure showed superiority over medical treatment while the other three comparisons were inconclusive. Of the 22 outcomes with moderate GRADE rating, 6 (27%) were inconclusive, 14 (64%) were in favor of surgery, and 2 (9%) were in favor of the drug regimen according to the calculations of the Cochrane authors (14 (64%), were inconclusive, 7 (32%) favored the surgical arm and 1 (5%) were in favor of the drug regimen according to our standard random-effects calculations).

Results of RCTs according to journal of publication

Of the 165 eligible RCTs (295 outcome assessments), 73 RCTs (133 assessments) were published in mostly surgical journals, 38 RCTs (69 assessments) in general journals, and 54 RCTs (93 assessments) in mostly non-surgical journals. Based on 95% confidence intervals for the assessments of RCTs published in mostly surgical journals, 40/133 (30%) were in favor of surgery, 14/133 (11%) were in favor of drugs, and 79/133 (59%) were inconclusive. The respective numbers for the assessments of RCTs published in general journals were 27/69 (39%), 5/69 (7%), and 37/69 (53%); and for the assessments of RCTs published in mostly non-surgical journals they were 22/93 (24%), 15/93 (16%), and 56 (60%), respectively. The proportion of RCTs favoring surgery was not significantly higher in mostly surgical journals (30%) compared to other journals (39% and 24% for general and non-surgical journals respectively) ($p=0.18$ by Fisher's exact test).

Discussion

Main findings

In a subset of Cochrane reviews that aimed to compare surgery to drugs we found that only 1 in 5 systematic reviews that had shown interest in such comparisons eventually found data from any RCTs for comparisons of the two modes of interventions. Furthermore, the majority of the comparisons where RCTs of surgery versus drugs had inconclusive results, and also had low or very low strength of the evidence on GRADE assessments. Anal fissure was the only disease in our sample that had high GRADE evidence and a direction of effect indicating that one intervention (sphincterotomy) was more effective. Consequently, in the vast majority of cases where surgical and pharmaceutical interventions are available for treatment, an evidence-based decision in the clinic is difficult. Our secondary post hoc analysis of the type of journal where the eligible RCTs were published showed that results published in surgical journals were not necessarily more prone to favor the surgical arm of an RCT over the pharmaceutical arm.

Strengths

This study covers the entire Cochrane database which is considered a high-quality comprehensive collection of systematic reviews. Cochrane reviews tend to address questions typically asked in routine clinical practice and underpin many clinical guideline recommendations, making this sample all the more relevant to everyday practice [49]. Another strength of this study is that all surgical specialties were included. This is, therefore, to our knowledge the first project aiming to assess the extent of comparative evidence for surgery versus pharmacotherapy for a diverse spectrum of diseases.

Limitations

Our analysis has several limitations. First, our pre-defined inclusion criteria excluded non-pharmacological medical interventions. Several comparisons may be found in the literature where surgery is compared against non-surgical non-pharmacological medical interventions, such as CPAP or radiotherapy. We also excluded endovascular and endoscopic procedures since they may be performed by surgical and medical specialists. These eligibility choices aimed to achieve some homogeneity in a project that is by definition already very heterogeneous. The use of an algorithm to filter out papers with no mention of the word surgery as well as the search

strategy itself may have led to us missing reviews that discuss a particular surgical procedure but never explicitly mention the word surgery but merely the name of the intervention.

Second, we focused exclusively on RCTs, but other types of evidence, e.g., non-randomized controlled trials, or uncontrolled clinical trials may also exist and sometimes their results may be compelling enough to deem a randomized study unnecessary. Such unquestionable superiority in the absence of randomized evidence is however unlikely [50]. Efforts such as IDEAL [8] have laid out much of the groundwork for performing RCTs in surgical research, yet a dearth of RCTs in the surgical realm of research persists to this day.

Third, only one database (Cochrane Database of Systematic Reviews) was used for this study and we did not examine non-Cochrane meta-analyses published as journal articles. While the database aims to be all-inclusive, there are still some topics in medical and surgical care that have not been covered by Cochrane reviews.

Fourth, it is possible that within the same disease, subgroups of patients may be eligible only for medical or only for surgical treatment, or that one or the other approach is much better only for specific subgroups. With the dearth of evidence we found for the overall analysis, identification of such subgroup effects would be unlikely and error-prone.

Context of these findings

Sequestration between different disciplines and specialties[51] may lead to isolation of specialists which use different tools, and this may lead to a lack of comparisons of the treatments that each specialty uses. Each specialty may have its own community, journals, meetings, and research agenda, limiting communication between different specialists even though they may be dealing with the same disease from different angles and with different therapeutic sets. This lack of communication may also be due to differences in mentorship and the trend of sub-specialization in medical training separating clinicians and their practices even further [52], or to differing incentive structures.

Prior literature comparing surgical and medical interventions has assessed specific treatments, such as that for basal cell carcinoma[51], and demonstrated that sequestration was prominent. Despite a large number of trials, almost all of them compared medical interventions among themselves, or surgicalaggressive interventions among themselves, rather than comparing between these two groups of treatment even though both groups of treatment could have been used. Our work shows that this issue of sequestration is widespread in surgical vs. pharmaceutical interventions.

Implications

This study suggests that comparisons of pharmaceutical and surgical interventions are infrequent. Even accepting the difficulties in performing RCTs involving surgical interventions, our results still indicate a need for more comparative effectiveness research and for improved communication between surgical and medical specialties to bridge this gap in evidence. There are, of course, barriers to this. Head-to-head comparisons of treatments are often disfavored by manufacturers leery of jeopardizing their product against that of a competitor [53,54], and incentives unfortunately exist for both surgical and medical practitioners to promote treatments they are able to offer. Moving forward, both medical and surgical professional societies should collaborate to design fair and unbiased trials, and funders should also keep such research on their radars to try and overcome these structural obstacles.

Future research

Future clinical research should try to expand the scope, volume, and methodological rigor of comparative evidence on surgical versus medical interventions. This work should involve both surgical and medical specialists and should also incorporate patient preferences. Long-term patient-centered outcomes, including both benefits and harms should become available to put surgical and medical practices into proper perspective.

References

- 1 Howick J, Koletsi D, Ioannidis JPA, *et al.* Most healthcare interventions tested in Cochrane Reviews are not effective according to high quality evidence: a systematic review and meta-analysis. *J Clin Epidemiol* 2022;:S0895-4356(22)00100-7. doi:10.1016/j.jclinepi.2022.04.017
- 2 Ergina PL, Cook JA, Blazeby JM, *et al.* Challenges in evaluating surgical innovation. *The Lancet* 2009;**374**:1097–104. doi:10.1016/S0140-6736(09)61086-2
- 3 Maruthappu M, Gilbert BJ, El-Harasis MA, *et al.* The influence of volume and experience on individual surgical performance: a systematic review. *Ann Surg* 2015;**261**:642–7. doi:10.1097/SLA.0000000000000852
- 4 Martling A, Cedermark B, Johansson H, *et al.* The surgeon as a prognostic factor after the introduction of total mesorectal excision in the treatment of rectal cancer. *British Journal of Surgery* 2002;**89**:1008–13. doi:10.1046/j.1365-2168.2002.02151.x
- 5 McGrath PD. Relation Between Operator and Hospital Volume and Outcomes Following Percutaneous Coronary Interventions in the Era of the Coronary Stent. *JAMA* 2000;**284**:3139. doi:10.1001/jama.284.24.3139
- 6 Gelijns AC, Ascheim DD, Parides MK, *et al.* Randomized trials in surgery. *Surgery* 2009;**145**:581–7. doi:10.1016/j.surg.2009.04.003
- 7 Barkun JS, Aronson JK, Feldman LS, *et al.* Evaluation and stages of surgical innovations. *Lancet* 2009;**374**:1089–96. doi:10.1016/S0140-6736(09)61083-7
- 8 McCulloch P, Altman DG, Campbell WB, *et al.* No surgical innovation without evaluation: the IDEAL recommendations. *Lancet* 2009;**374**:1105–12. doi:10.1016/S0140-6736(09)61116-8
- 9 Janiaud P, Agarwal A, Tzoulaki I, *et al.* Validity of observational evidence on putative risk and protective factors: appraisal of 3744 meta-analyses on 57 topics. *BMC Med* 2021;**19**:157. doi:10.1186/s12916-021-02020-6
- 10 Belbasis L, Bellou V, Ioannidis JPA. Conducting umbrella reviews. *bmjmed* 2022;**1**:e000071. doi:10.1136/bmjmed-2021-000071
- 11 Hartung J, Knapp G. A refined method for the meta-analysis of controlled clinical trials with binary outcome. *Stat Med* 2001;**20**:3875–89. doi:10.1002/sim.1009
- 12 Int'Hout J, Ioannidis JPA, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Med Res Methodol* 2014;**14**:25. doi:10.1186/1471-2288-14-25

- 13 Weber F, Knapp G, Ickstadt K, *et al.* Zero-cell corrections in random-effects meta-analyses. *Res Syn Meth* 2020;**11**:913–9. doi:10.1002/jrsm.1460
- 14 Ioannidis JPA. The Proposal to Lower *P* Value Thresholds to .005. *JAMA* 2018;**319**:1429. doi:10.1001/jama.2018.1536
- 15 Phelps PO, Abariga SA, Cowling BJ, *et al.* Antimetabolites as an adjunct to dacryocystorhinostomy for nasolacrimal duct obstruction. *Cochrane Database Syst Rev* 2020;**4**:CD012309. doi:10.1002/14651858.CD012309.pub2
- 16 Rees K, Beranek-Stanley M, Burke M, *et al.* Hypothermia to reduce neurological damage following coronary artery bypass surgery. *Cochrane Database Syst Rev* 2001;:CD002138. doi:10.1002/14651858.CD002138
- 17 Hoare BJ, Wallen MA, Imms C, *et al.* Botulinum toxin A as an adjunct to treatment in the management of the upper limb in children with spastic cerebral palsy (UPDATE). *Cochrane Database Syst Rev* 2010;**2010**:CD003469. doi:10.1002/14651858.CD003469.pub4
- 18 Shylasree TS, Bryant A, Howells RE. Chemoradiation for advanced primary vulval cancer. *Cochrane Database Syst Rev* 2011;**2011**:CD003752. doi:10.1002/14651858.CD003752.pub3
- 19 Lansbury L, Leonardi-Bee J, Perkins W, *et al.* Interventions for non-metastatic squamous cell carcinoma of the skin. *Cochrane Database Syst Rev* 2010;:CD007869. doi:10.1002/14651858.CD007869.pub2
- 20 Moreno G, Corbalán J, Peñaloza B, *et al.* Topical corticosteroids for treating phimosis in boys. *Cochrane Database Syst Rev* 2014;:CD008973. doi:10.1002/14651858.CD008973.pub2
- 21 Mitra S, Scrivens A, von Kursell AM, *et al.* Early treatment versus expectant management of hemodynamically significant patent ductus arteriosus for preterm infants. *Cochrane Database Syst Rev* 2020;**12**:CD013278. doi:10.1002/14651858.CD013278.pub2
- 22 Flower A, Liu JP, Lewith G, *et al.* Chinese herbal medicine for endometriosis. *Cochrane Database Syst Rev* 2012;:CD006568. doi:10.1002/14651858.CD006568.pub3
- 23 Görk AS, Ehrenkranz RA, Bracken MB. Continuous infusion versus intermittent bolus doses of indomethacin for patent ductus arteriosus closure in symptomatic preterm infants. *Cochrane Database Syst Rev* 2008;**2008**:CD006071. doi:10.1002/14651858.CD006071.pub2
- 24 Zhao J-G, Wang J, Huang W-J, *et al.* Interventions for treating simple bone cysts in the long bones of children. *Cochrane Database Syst Rev* 2017;**2**:CD010847. doi:10.1002/14651858.CD010847.pub3
- 25 Bettany-Saltikov J, Weiss H-R, Chockalingam N, *et al.* Surgical versus non-surgical interventions in people with adolescent idiopathic scoliosis. *Cochrane Database Syst Rev* 2015;:CD010663. doi:10.1002/14651858.CD010663.pub2

- 26 Kang D, Han J, Neuberger MM, *et al.* Transurethral radiofrequency collagen denaturation for the treatment of women with urinary incontinence. *Cochrane Database Syst Rev* 2015;:CD010217. doi:10.1002/14651858.CD010217.pub2
- 27 Colquitt JL, Pickett K, Loveman E, *et al.* Surgery for weight loss in adults. *Cochrane Database Syst Rev* 2014;:CD003641. doi:10.1002/14651858.CD003641.pub4
- 28 Palmer JS, Monk AP, Hopewell S, *et al.* Surgical interventions for symptomatic mild to moderate knee osteoarthritis. *Cochrane Database Syst Rev* 2019;7:CD012128. doi:10.1002/14651858.CD012128.pub2
- 29 Laupattarakasem W, Laopaiboon M, Laupattarakasem P, *et al.* Arthroscopic debridement for knee osteoarthritis. *Cochrane Database Syst Rev* 2008;:CD005118. doi:10.1002/14651858.CD005118.pub2
- 30 Liu JP, Yang H, Xia Y, *et al.* Herbal preparations for uterine fibroids. *Cochrane Database Syst Rev* 2013;:CD005292. doi:10.1002/14651858.CD005292.pub3
- 31 Cirocchi R, Trastulli S, Pressi E, *et al.* Non-operative management versus operative management in high-grade blunt hepatic injury. *Cochrane Database Syst Rev* 2015;2015:CD010989. doi:10.1002/14651858.CD010989.pub2
- 32 Page MJ, Massy Westropp N, O'Connor D, *et al.* Splinting for carpal tunnel syndrome. *Cochrane Database Syst Rev* 2012;2012:CD010003. doi:10.1002/14651858.CD010003
- 33 Nikolaidis I, Fouyas IP, Sandercock PA, *et al.* Surgery for cervical radiculopathy or myelopathy. *Cochrane Database Syst Rev* 2010;2010:CD001466. doi:10.1002/14651858.CD001466.pub3
- 34 Cheuk DKL, Wong V, Wraige E, *et al.* Surgery for scoliosis in Duchenne muscular dystrophy. *Cochrane Database Syst Rev* 2015;2015:CD005375. doi:10.1002/14651858.CD005375.pub4
- 35 Weitz M, Portz S, Laube GF, *et al.* Surgery versus non-surgical management for unilateral ureteric-pelvic junction obstruction in newborns and infants less than two years of age. *Cochrane Database Syst Rev* 2016;7:CD010716. doi:10.1002/14651858.CD010716.pub2
- 36 Abraham A, Kumar S, Chaudhry S, *et al.* Surgical interventions for diaphyseal fractures of the radius and ulna in children. *Cochrane Database Syst Rev* 2011;:CD007907. doi:10.1002/14651858.CD007907.pub2
- 37 Chang MY, Coleman AL, Tseng VL, *et al.* Surgical interventions for vertical strabismus in superior oblique palsy. *Cochrane Database Syst Rev* 2017;2017:CD012447. doi:10.1002/14651858.CD012447.pub2
- 38 Monk AP, Davies LJ, Hopewell S, *et al.* Surgical versus conservative interventions for treating anterior cruciate ligament injuries. *Cochrane Database Syst Rev* 2016;4:CD011166. doi:10.1002/14651858.CD011166.pub2

- 39 Gosler MW, Testroote M, Morrenhof JW, *et al.* Surgical versus non-surgical interventions for treating humeral shaft fractures in adults. *Cochrane Database Syst Rev* 2012;**1**:CD008832. doi:10.1002/14651858.CD008832.pub2
- 40 Smith TO, Donell S, Song F, *et al.* Surgical versus non-surgical interventions for treating patellar dislocation. *Cochrane Database Syst Rev* 2015;:CD008106. doi:10.1002/14651858.CD008106.pub3
- 41 Abudou M, Chen X, Kong X, *et al.* Surgical versus non-surgical treatment for thoracolumbar burst fractures without neurological deficit. *Cochrane Database Syst Rev* 2013;:CD005079. doi:10.1002/14651858.CD005079.pub3
- 42 Cataneo AJM, Cataneo DC, de Oliveira FHS, *et al.* Surgical versus nonsurgical interventions for flail chest. *Cochrane Database Syst Rev* 2015;**2015**:CD009919. doi:10.1002/14651858.CD009919.pub2
- 43 Khalil HS, Nunez DA. Functional endoscopic sinus surgery for chronic rhinosinusitis. *Cochrane Database Syst Rev* 2006;:CD004458. doi:10.1002/14651858.CD004458.pub2
- 44 Tóth M, Shah A, Hu K, *et al.* Endoscopic cyclophotocoagulation (ECP) for open angle glaucoma and primary angle closure. *Cochrane Database Syst Rev* 2019;**2**:CD012741. doi:10.1002/14651858.CD012741.pub2
- 45 Fakhry F, Fokkenrood HJ, Spronk S, *et al.* Endovascular revascularisation versus conservative management for intermittent claudication. *Cochrane Database Syst Rev* 2018;**3**:CD010512. doi:10.1002/14651858.CD010512.pub2
- 46 Araujo ST, Moreno DH, Cacione DG. Percutaneous thrombectomy or ultrasound-accelerated thrombolysis for initial management of acute limb ischaemia. *Cochrane Database Syst Rev* 2022;**1**:CD013486. doi:10.1002/14651858.CD013486.pub2
- 47 Jenks S, Yeoh SE, Conway BR. Balloon angioplasty, with and without stenting, versus medical therapy for hypertensive patients with renal artery stenosis. *Cochrane Database Syst Rev* 2014;**2014**:CD002944. doi:10.1002/14651858.CD002944.pub2
- 48 O'Connell NE, Wand BM, McAuley J, *et al.* Interventions for treating pain and disability in adults with complex regional pain syndrome. *Cochrane Database Syst Rev* 2013;**2013**:CD009416. doi:10.1002/14651858.CD009416.pub2
- 49 Korfitsen CB, Mikkelsen M-LK, Ussing A, *et al.* Usefulness of Cochrane Reviews in Clinical Guideline Development-A Survey of 585 Recommendations. *Int J Environ Res Public Health* 2022;**19**:685. doi:10.3390/ijerph19020685
- 50 Hozo I, Djulbegovic B, Parish AJ, *et al.* Identification of threshold for large (dramatic) effects that would obviate randomized trials is not possible. *Journal of Clinical Epidemiology* 2022;**145**:101–11. doi:10.1016/j.jclinepi.2022.01.016

- 51 Kim DD, Tang JY, Ioannidis JPA. Network geometry shows evidence sequestration for medical vs. surgical practices: treatments for basal cell carcinoma. *Journal of Clinical Epidemiology* 2014;**67**:391–400. doi:10.1016/j.jclinepi.2013.10.015
- 52 Hirshman BR, Tang JA, Jones LA, *et al.* Impact of medical academic genealogy on publication patterns: An analysis of the literature for surgical resection in brain tumor patients: Medical Academic Genealogy. *Ann Neurol* 2016;**79**:169–77. doi:10.1002/ana.24569
- 53 Flacco ME, Manzoli L, Boccia S, *et al.* Head-to-head randomized trials are mostly industry sponsored and almost always favor the industry sponsor. *Journal of Clinical Epidemiology* 2015;**68**:811–20. doi:10.1016/j.jclinepi.2014.12.016
- 54 Lathyris DN, Patsopoulos NA, Salanti G, *et al.* Industry sponsorship and selection of comparators in randomized clinical trials. *European Journal of Clinical Investigation* 2010;**40**:172–82. doi:10.1111/j.1365-2362.2009.02240.x

Authors' contribution

The draft of the manuscript and the analysis was initially outlined and drafted by EAZ and JPAI with input from all authors in iterations. All authors had access to data, and approved the project design and critically reviewed, contributed to the final version of this work, and approve it.

Transparency statement

Authors declare manuscript is an honest, accurate, and transparent account of the study being reported; no important aspects of the study have been omitted; and that any discrepancies from the study as originally planned (and, if relevant, registered) have been explained in the published protocol.

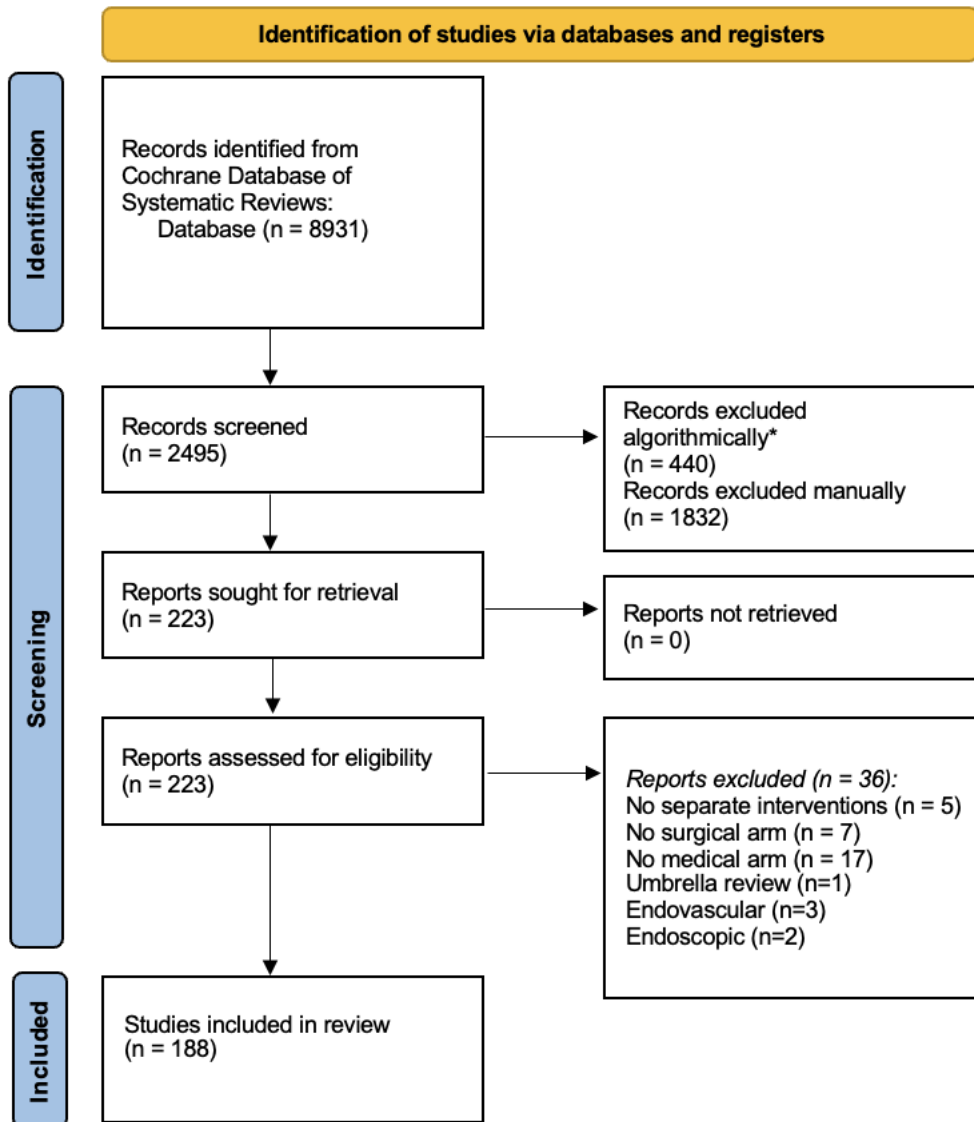
Ethics approval

No ethics approval was required

Data availability

The PRISMA checklists, code, and data as well as the protocol are found in the Open Science Framework repository and is found in <https://osf.io/p9x3j>

Figure 1. PRISMA study selection flowchart



*Searched title and abstract for withdrawals, and for abstracts without the word surgery or any of its variations to exclude them

Table 1. Eligible comparisons of surgical versus medical interventions

Surgical arm	Drug arm	Disease	Outcomes (studies/N)
Cardiac surgery			
Transmyocardial lazer revascularization	Continued medication	Refractory angina	Angina reduction (7/1053) Overall mortality (7/1053) Postoperative mortality (30 d) (6/967)
Surgical closure	IV indomethacin	Patent ductus arteriosus	Death before discharge (1/154)
Dermatology			
Surgical excision	Imiquimod	BCC	Recurrence (3 y) (1/501) Recurrence (5 y) (1/501) Observer-rated good/excellent cosmetic outcome (1/501) Patient-rated good/excellent cosmetic outcome (1/501)
Surgical excision	MAL-PDT	BCC	Recurrence (3 y) (1/68) Observer-rated good/excellent cosmetic outcome (2/351) Patient-rated good/excellent cosmetic outcome (2/351)
Surgical excision	ALA-PDT	BCC	Recurrence (3 y) (1/173) Recurrence (5 y) (1/173)
General surgery			
Lateral internal sphincterotomy	Medical therapy (mainly GTN Isosorbide dinitrate and Botox)	Anal fissure	NON-Healing (persistence or recurrence) 2 mo. (15/979)
Pancreatic resection	Chemoradiotherapy	Pancreatic cancer	Overall mortality (5 y) (2/98)
Oesophagectomy	Chemoradiotherapy and/or radiotherapy	Oesophageal cancer	Short-term mortality (5/689) Long-term mortality (3/511) Serious adverse event (3 months) (1/80) Short-term health-related QOL (1/165) Medium-term health-related QOL (1/62)
Laparoscopic fundoplication	Protein pump inhibitors	GERD	Health-related QOL: <1 y (3/605) 1-5 y (3/323) GORD-specific QOL: <1 y (4/1160) 1-5 y (3/994) Serious adverse events (2/637)
Surgery	Tamoxifen	Primary breast cancer	Overall survival (3/495)
Neurosurgery			
Decompressive surgery	Prednisolone	Leprosy	Change in motor or sensory score after one year (1/57) Proportion of ulnar nerves with: sensory improvement after one year (1/62), and motor improvement after one year (1/62)
Epilepsy surgery	Continued antiepileptic drugs	Epilepsy	Proportion free from seizures (1 y) (2/196) Proportion free from all seizures including auras (1 y) (1/80)
Decompressive craniectomy	Medical treatment (including barbiturates)	High ICP in closed TBI	Neurological unfavourable outcome 6 mo (3/571) Mortality 6 mo (3/571)
Surgical decompression	Osmotic agents, blood pressure	Cerebral oedema in acute ischaemic	Death at the end of follow-up (3/134)

Surgical arm	Drug arm	Disease	Outcomes (studies/N)
Surgical decompression	control, and glucose control Dexamethasone, antihypertensives and intermittent diuresis	stroke Primary supratentorial intracerebral haemorrhage	Death or dependence at end of follow up (9/1994)
Obstetrics and gynaecology			
Suction aspiration	Vaginal suppositories or im inj. of 9-methylene-PGE2	Abortion	Abortion not completed with intended method (2/472) Ongoing pregnancy (2/472) Pelvic infection (1/419)
Suction aspiration	Misoprostol	Abortion	Complete miscarriage (22/5285) Composite outcome of death or serious complication (9/2146)
Suction aspiration	Vaginal or oral misoprostol	Abortion	Complete miscarriage (15/3862) Surgical evacuation (13/3070) Death or serious complication (5/1248)
Suction aspiration	Misoprostol and mifepristone	Abortion	Complete miscarriage (2/716) Composite outcome of death or serious complication (1/618)
Dilatation and curettage	Misoprostol	Abortion	Complete miscarriage (1/107) Composite outcome of death or serious
Dilation and evacuation	Misoprostol	Abortion	Combined major and minor complications (1/94)
Laparoscopic ovarian drilling	Medical ovulation induction	Infertility due to PCOS	Live birth (9/1015) Multiple pregnancy (14/1161)
Laparoscopic ovarian drilling	Letrozele	Infertility due to PCOS	Live birth (3/548) Rate of ovarian hyperstimulation syndrome (1/250)
Laparoscopic ovarian drilling	Gonadotropins	PCOS	Menstrual regularity at 6 mo. (1/35) Improvement in androgenic symptoms 6 mo. (1/126)
Laparoscopic ovarian drilling	Metformin, clomiphene	PCOS	Menstrual regularity at 6 mo. (2/332)
Laparoscopic ovarian drilling	Letrozele	PCOS	Menstrual regularity at 6 mo. (1/260)
Laparoscopic ovarian drilling	Metformin, letrozele	PCOS	Menstrual regularity at 6 mo. (1/156)
Laparoscopic ovarian drilling	Metformin	PCOS	Menstrual regularity at 6 mo. (2/236) Improvement in androgenic symptoms 6 mo. (1/50)
Transcervical resection of endometrium using rollerball coagulation	Hormone therapy or antifibrinolytic	Heavy menstrual bleeding	Control of bleeding (cure or improvement to acceptable level) 4 mo. (1/186) Control of bleeding (cure or improvement to acceptable level) 2 y (1/173) Control of bleeding (cure or improvement to acceptable level) 5 y (1/140) Overall satisfaction with treatment 4 mo. (1/186) Overall satisfaction with treatment 2 y (1/173) Overall satisfaction with treatment 5 y (1/141) Adverse events at 4 months (1/186)
Ophthalmology			
Amniotic membrane transplantation and medication	Lubrication, antibiotics and pressure lowering medication	Acute ocular burns	Epithelial defect 21 d post-injury, and Visual acuity at final follow-up (1/68)
Laser surgery	intravitreal anti-VEGF	Pathological myopia	Change in best corrected visual acuity (1/36) Proportion of participants with a gain of 3+ lines in BCVA at 1 y (1/36)
iStent	Latanoprost/timolol	Open angle glaucoma	Proportion of participants who were drop-free 6-18 mo (2/285)
Argon laser trabeculoplasty	IOP reducing medication	Open angle glaucoma	Failure to control IOP (3/735) Visual field progression (2/624)

Surgical arm	Drug arm	Disease	Outcomes (studies/N)
Surgical correction	Botulinum toxin	Strabismus	Optic neuropathy progression (2/264) Improved ocular alignment > 10 dioptres, children (2/102), adults (1/30)
Orthopaedic surgery			
Open section of the carpal ligament	NSAID and splinting or corticosteroid injections	Carpal tunnel syndrome	Improvement in clinical symptoms at three months of follow-up (2/245)
Open surgery	Corticosteroid injection	Trigger finger	Resolution of triggering (2/270)
Decompressive surgery with or without fusion	Epidural steroid injection	Lumbar spinal stenosis	Oswestry Disability index 6 w (1/38) Pain intensity (VAS) 6 w (1/38) Zurich claudication questionnaire (symptom evaluation) 6 w (1/38)
Open unilateral sympathectomy (L2-4)	IV prostanoid iloprost	Critical limb ischaemia	Complete ulcer healing w/o rest pain or major amputation (24 w) (1/162)
Surgical rotator cuff repair	Non-operative treatment including corticosteroid injection and exercise	Rotator cuff tear	Pain (VAS) 12 mo (1/56)
Arthroscopic surgery	Sclerosing injection	Jumper's knee	Knee pain (0-100, 12 mo.) (1/50) Participant global assessment of success (1-100, 12 mo.) (1/50) Withdrawal rate (1/52)
Otolaryngology			
Surgical orbital decompression	IV Methylprednisolone 1x3 followed by oral prednisolone	Thyroid eye disease	Proportion of successes compared to the proportion of treatment failures as defined by the study authors based on the use of composite outcome scores (1/15)
Grommets (ventilation tubes)	Antibiotic prophylaxis	Recurrent acute otitis media	Proportion of patients who have no AOM recurrences (6 mo.) (2/96)
Tonsillectomy or adenotonsillectomy	Watchful waiting with or without analgesics and antibiotics	Tonsillitis	Episodes of sore throat of any severity (children) (5/795) Episodes of moderately or severely sore throat (children) (4/564) Sore throat days (children) (5/776) Episodes of sore throat of any severity (adults) (2/156) Sore throat days (adults) (2/156)
Thoracic surgery			
Open thoracotomy	Thoracostomy drainage (with fibrinolytics)	Pleural empyema	Mortality (1/30)
VATS	Thoracostomy drainage (with fibrinolytics)	Pleural empyema	Mortality (7/367)
Urology			
Surgical reimplantation of ureters	Antibiotics	Primary vesicoureteric reflux	Rate of patients with symptomatic UTI (1/297)
Vascular surgery			
Carotid endarterectomy and Aspirin 325 mg daily	Aspirin 325 mg daily	Asymptomatic carotid stenosis	Perioperative stroke or death, or stroke of any territory or type during follow up (2/2103)
Aspirin and carotid surgery	Aspirin	Carotid stenosis	Any stroke or operative death (3/6090) Ipsilateral ischaemic stroke, and any operative stroke or death near occlusion (3/6090)
Saphenofemoral disconnection	Therapeutic LMWH	Superficial thrombophlebitis	Symptomatic VTE (1/60) Major bleeding (1/60)
Surgery including primary amputation	Thrombolysis (w/ rt-Pa or urokinase)	Acute limb ischaemia	Limb salvage (30 d) (3/841)

Abbreviations

BCC: basal cell carcinoma of the skin

GERD: Gastro-oesophageal reflux disease

GTN: glyceryl tri-nitrate

IOP: intra-ocular pressure

PCOS: polycystic ovarian syndrome

QOL: Quality of life

Table 2. Comparisons where the surgical treatment was superior to the drug treatment

Surgical arm	Drug arm	Disease	Outcome	Treatment effect (95% CI)	GRADE assessment
Cardiac surgery					
Transmyocardial lazer revascularization	Continued medication	Refractory angina	Angina reduction	OR=4.63 (3.43-6.25)	Low
Dermatology					
Surgical excision	Imiquimod	BCC	Recurrence (3 y)	RR=0.1 (0.03-0.31)	Moderate
			Recurrence (5 y)	RR=0.13 (0.05-0.36)	Moderate
Surgical excision	MAL-PDT	BCC	Recurrence (3 y)	RR=0.04 (0-0.61)	Low
Surgical excision	ALA-PDT	BCC	Recurrence (3 y)	RR=0.09 (0.02-0.38)	Moderate
			Recurrence (5 y)	RR=0.08 (0.02-0.34)	Moderate
General surgery					
Laparoscopic fundoplication	Protein pump inhibitors	GERD	GORD-specific QOL (<1 y)	SMD=0.58 (0.46-0.7)	Low
Lateral internal sphincterotomy	Medical therapy (mainly GTN and Botox)	Anal fissure	Non-Healing (persistence or recurrence) 2 mo.	OR=0.11 (0.06-0.23)	High
Neurosurgery					
Epilepsy surgery	Continued antiepileptic drugs	Epilepsy	Proportion (%) free from seizures (1 y)	RR=9.78 (4.73-20.2)*	Low
			Proportion free from all seizures incl. auras (1 y)	RR=15 (2.08-108.23)	Very Low
Surgical decompression	Osmotic agents, blood pressure control, and glucose control	Cerebral oedema in acute ischaemic stroke	Death at the end of follow-up	OR=0.19 (0.09-0.37)	
Surgical decompression	Dexamethasone, antihypertensives and intermittent diuresis	Primary supratentorial intracerebral haemorrhage	Death or dependence at end of follow up	OR=0.71 (0.58-0.88)	
Obstetrics and gynaecology					
Suction aspiration	Misoprostol	Abortion	Complete miscarriage	RR=1.11 (1.06-1.17)	Very Low
			Complete miscarriage	RR=1.04 (1.02-1.06)	Very Low
Dilatation and curettage	Misoprostol	Abortion	Complete miscarriage	RR=1.18 (1.1-1.27)*	Very Low
Dilatation and evacuation	Misoprostol	Abortion	Combined major and minor complications	OR=0.12 (0.03-0.46)	
Laparoscopic ovarian drilling	Medical ovulation induction	Infertility due to PCOS	Multiple pregnancy	OR=0.34 (0.18-0.66)	Moderate
Laparoscopic	Gonadotropins	PCOS	Menstrual regularity at	OR=19.2 (3.17-116)	Very Low

Surgical arm	Drug arm	Disease	Outcome	Treatment effect (95% CI)	GRADE assessment
ovarian drilling			6 mo.		
Transcervical resection of endometrium using rollerball coagulation	Hormone therapy or antifibrinolytic	Heavy menstrual bleeding	Control of bleeding (cure or improvement to acceptable level) 4 mo.	RR=2.66 (1.94-3.64)	Moderate
			Control of bleeding (cure or improvement to acceptable level) 2 y	RR=1.29 (1.06-1.57)	Low
			Overall satisfaction with treatment 4 mo.	RR=2.8 (1.96-3.99)	Moderate
			Overall satisfaction with treatment 2 y	RR=1.4 (1.13-1.74)	Moderate
			Adverse events at 4 months	RR=0.26 (0.15-0.46)	Moderate
Ophthalmology					
Surgical correction	Botulinum toxin	Strabismus	Improved ocular alignment > 10 dioptres, adults	RR=2.63 (1.18-5.9)	Low
iStent	Latanoprost/timolol	Open angle glaucoma	Proportion of participants who were drop-free 6-18 mo	RR=125 (17.8-884)	Very low
Argon laser trabeculoplasty	IOP reducing medication	Open angle glaucoma	Failure to control IOP	RR=0.8 (0.71-0.91)	
Orthopaedic surgery					
Arthroscopic surgery	Sclerosing injection	Jumper's knee	Knee pain (0-100, 12 mo.)	MD=-28.3 (-41.79- -14.81)	Low
			Participant global assessment of success (1-100, 12 mo.)	MD=33.9 (18.74- 49.06)	Low
Decompressive surgery with or without fusion	Epidural steroid injection	Lumbar spinal stenosis	Zurich claudication questionnaire (symptom evaluation) 6 w	MD=-0.6 (-0.77- -0.43)	Low
Open unilateral sympathectomy (L2-4)	IV prostanoid iloprost		Complete ulcer healing w/o rest pain or major amputation (24 w)	RR=1.76 (1.35-2.29)	Low
Otolaryngology					
Grommets (ventilation tubes)	Antibiotic prophylaxis	Recurrent acute otitis media	Proportion of patients who have no recurrences (6 mo.)	RR=1.68 (1.07-2.65)*	Very Low
Tonsillectomy or adenotonsillectomy	Watchful waiting with or without analgesics and antibiotics	Tonsillitis	Episodes of sore throat of any severity (children)	MD=-0.56 (-1.04- -0.07)*	Moderate
			Sore throat days (children)	MD=-5.13 (-8.03- -2.2)*	Moderate
			Episodes of sore throat of any severity (adults)	-MD=3.61 (-7.92- -0.7)*	Moderate
			Sore throat days (adults)	MD=-10.64 (-15.52- -5.76)*	Moderate
Vascular surgery					

Surgical arm	Drug arm	Disease	Outcome	Treatment effect (95% CI)	GRADE assessment
Aspirin and carotid surgery	Aspirin	Carotid stenosis	Any stroke or operative death	RR=0.85 (0.77-0.95)*	Moderate

*our re-analysis using a random effects meta-analysis model shows that the 95% confidence interval includes the null (results are inconclusive)

RR: risk ratio

OR: odds ratio

HR: hazard ratio

MD: mean difference

SMD: standardized mean difference

BCC: basal cell carcinoma of the skin

GERD: Gastro-oesophageal reflux disease

GNT: glyceryl trinitrate

IOP: intra-ocular pressure

PCOS: polycystic ovarian syndrome

QOL: Quality of life

Table 3. Comparisons where the drug treatment was superior to the surgical treatment

Surgical arm	Drug arm	Disease	Outcome	Treatment effect (95% CI)	GRADE assessment
Dermatology					
Surgical excision	Imiquimod	BCC	Observer-rated good/excellent cosmetic outcome	RR=0.59 (0.47-0.74)	Low
Surgical excision	MAL-PDT	BCC	Observer-rated good/excellent cosmetic outcome	RR=0.85 (0.79-0.92)*	Moderate
Surgical excision	MAL-PDT	BCC	Patient-rated good/excellent cosmetic outcome	RR=0.53 (0.44-0.65)*	Moderate
General surgery					
Oesophagectomy	Chemoradiotherapy and/or radiotherapy	Oesophageal cancer	Serious adverse event (3 months)	RR=1.73 (1.11-2.67)*	Very Low
			Short-term health-related QOL	MD=0.93 (0.24-1.62)	Very Low
Laparoscopic fundoplication	Protein pump inhibitors	GERD	Serious adverse events	RR=1.46 (1.01-2.11)	Very Low
Pancreatic resection	Chemoradiotherapy	Pancreatic cancer	Overall mortality (5 y)	HR=2.63 (1.72-4)*	Very Low
Obstetrics and gynaecology					
Laparoscopic ovarian drilling	Medical ovulation induction	Infertility due to PCOS	Live birth	OR=0.71 (0.54-0.92)	Low
Suction aspiration	Vaginal or oral misoprostol	Abortion	Surgical evacuation	RR=20 (9.1-50)	Very Low
Ophthalmology					
Laser surgery	intravitreal anti-VEGF	Pathological myopia	Change in best corrected visual acuity	MD=0.22 (0.01-0.43)*	Low
Amniotic membrane transplantation and medication	Lubrication, Antibiotics and Pressure lowering medication	Acute ocular burns	Visual acuity at final follow-up	MD=-0.83 (-1.32- -0.34)	Very Low
Orthopaedic surgery					
Decompressive surgery with or without fusion	Epidural steroid injection	Lumbar spinal stenosis	Oswestry Disability index 6 w	MD=5.7 (0.57-10.83)	Low
			Pain intensity (VAS) 6 w	MD=2.4 (1.92-2.88)	Low
Otolaryngology					
Tonsillectomy or adenotonsillectomy	Watchful waiting with or without analgesics and antibiotics	Tonsillitis	Episodes of moderately or severely sore throat (children)	MD=0.62 (0.22-1.03)*	Low
Vascular surgery					
Carotid endarterectomy and Aspirin 325 mg daily	Aspirin 325 mg daily	Asymptomatic carotid stenosis	Perioperative stroke or death, or stroke of any territory or type during follow up	RR=6.49 (2.53-16.61)	

*our re-analysis using a random effects meta-analysis model shows that the 95% confidence interval includes the null (results are inconclusive)

RR: risk ratio

OR: odds ratio

HR: hazard ratio

MD: mean difference

BCC: basal cell carcinoma of the skin

GERD: Gastro-oesophageal reflux disease

PCOS: polycystic ovarian syndrome

Table 4. GRADE assessment across specialties

Specialty	Very Low	Low	Moderate	High	None available
Cardiac surgery	0 (0)	1 (25)	0 (0)	2 (50)	1 (25)
Dermatology	0 (0)	3 (33)	6 (67)	0 (0)	0 (0)
General surgery	9 (69)	3 (23)	0 (0)	1 (8)	0 (0)
Neurosurgery	5 (50)	2 (20)	1 (10)	0 (0)	2 (20)
Obstetrics and gynecology	14 (45)	4 (13)	7 (23)	1 (3)	5 (16)
Ophthalmology	2 (20)	5 (50)	0 (0)	0 (0)	3 (30)
Orthopaedic surgery	2 (20)	6 (60)	1 (10)	0 (0)	1 (10)
Otolaryngology	1 (14)	1 (14)	4 (57)	0 (0)	1 (14)
Thoracic surgery	0 (0)	1 (50)	1 (50)	0 (0)	0 (0)
Urology	0 (0)	0 (0)	0 (0)	0 (0)	1 (100)
Vascular surgery	0 (0)	1 (17)	2 (33)	0 (0)	3 (50)

Supplementary material

Supplementary table 1. Reviews per specialty

Specialty	Total reviews	Reviews with at least one comparison (%)
Cardiac surgery	6	2 (33)
Dermatology	5	1 (20)
General surgery	35	5 (14)
Neurosurgery	12	5 (42)
Obstetrics and gynecology	31	8 (26)
Ophthalmology	25	5 (20)
Orthopaedic surgery	23	6 (26)
Otolaryngology	23	3 (13)
Thoracic surgery	9	1 (11)
Urology	7	1 (14)
Vascular surgery	12	4 (33)

Supplementary Table 2. Inconclusive comparisons between surgery and drugs

Surgical arm	Drug arm	Disease	Outcome	Treatment effect (95% CI)	GRADE assessment
Cardiac surgery					
Transmyocardial lazer revascularization	Continued medication	Refractory angina	Overall mortality	OR=1.12 (0.77-1.63)	High
			Postoperative mortality (30 d)	OR=1.19 (0.63-2.24)	High
Surgical closure	IV indomethacin	Patent ductus arteriosus	Death before discharge	RR=0.67 (0.34-1.31)	
Dermatology					
Surgical excision	Imiquimod	BCC	Patient-rated good/excellent cosmetic outcome	RR=1 (0.94-1.06)	Low
General surgery					
Surgery	Tamoxifen	Primary breast cancer	Overall survival	HR=0.98 (0.81-1.2)	Low
Laparoscopic fundoplication	Protein pump inhibitors	GERD	Health-related quality of life (<1 y)	SMD=0.14 (-0.02-0.3)+	Very Low
			Health-related QOL (1-5 y)	SMD=0.03 (-0.19-0.24)+	Very Low
			GORD-specific quality of life (1-5 y)	SMD=0.28 (-0.27-0.84)+	Very Low
Oesophagectomy	Chemoradiotherapy and/or radiotherapy	Oesophageal cancer	Short-term mortality	RR=0.39 (0.11-1.35)	Very Low
			Long-term mortality	RR=1.03 (0.92-1.14)	Low
			Medium-term health-related QOL	MD=-0.95 (-2.1-0.2)	Very Low
Neurosurgery					
Decompressive surgery	Prednisolone	Leprosy	Change in sensory score after one year	MD=0.08 (-2.45-2.61)	Very Low
			Proportion of ulnar nerves with sensory improvement after one year	RR=1.13 (0.71-1.77)	Very Low
			Change in motor score after one year	MD=0.82 (-1.34-2.98)	Very Low
			Proportion of ulnar nerves with motor improvement after one year	RR=0.91 (0.64-1.28)	Very Low
Decompressive craniectomy	Medical treatment (including barbiturates)	High ICP in closed TBI	Neurological unfavourable outcome 6 mo	RR=1 (0.71-1.4)	Low
			Mortality 6 mo	RR=0.66 (0.43-1.01)	Moderate
Obstetrics and gynaecology					
Suction aspiration	Vaginal or oral misoprostol	Abortion	Death or serious complication	RR=1 (0.04-25)	

Surgical arm	Drug arm	Disease	Outcome	Treatment effect (95% CI)	GRADE assessment
Suction aspiration	Misoprostol	Abortion	Composite outcome of death or serious complication	RR=1.53 (0.45-5.16)	Very Low
Suction aspiration	Misoprostol and mifepristone	Abortion	Complete miscarriage	RR=1.29 (0.96-1.73)	Very Low
Suction aspiration	Vaginal suppositories or inj. of 9-methylene-PGE2	Abortion	Composite outcome of death or serious complication	RR=0.14 (0.01-2.74)	Very Low
Suction aspiration	Vaginal suppositories or inj. of 9-methylene-PGE2	Abortion	Abortion not completed with intended method	OR=0.62 (0.02-16.6)	
Suction aspiration	Vaginal suppositories or inj. of 9-methylene-PGE2	Abortion	Ongoing pregnancy	OR=1.82 (0.54-6.25)	
Suction aspiration	Vaginal suppositories or inj. of 9-methylene-PGE2	Abortion	Pelvic infection	OR=0.46 (0.14-1.56)	
Dilatation and curettage	Misoprostol	Abortion	Composite outcome of death or serious complication	RR=0.79 (0.34-1.85)	Very Low
Laparoscopic ovarian drilling	Metformin, Clomiphene	PCOS	Menstrual regularity at 6 mo.	OR=1.02 (0.64-1.64)	Very Low
Laparoscopic ovarian drilling	Letrozele	PCOS	Menstrual regularity at 6 mo.	OR=1.08 (0.64-1.84)	Very Low
Laparoscopic ovarian drilling	Metformin, Letrozol	PCOS	Menstrual regularity at 6 mo.	OR=0.95 (0.49-1.81)	Very Low
Laparoscopic ovarian drilling	Metformin	PCOS	Menstrual regularity at 6 mo.	OR=1.51 (0.62-3.71)	Moderate
Laparoscopic ovarian drilling	Gonadotropins	PCOS	Improvement in androgenic symptoms 6 mo.	OR=3.02 (0.56-16.33)	Low
Laparoscopic ovarian drilling	Metformin	PCOS	Improvement in androgenic symptoms 6 mo.	OR=1 (0.42-2.37)	Low
Laparoscopic ovarian drilling	Letrozele	Infertility due to PCOS	Live birth	RR=0.72 (0.5-1.05)	Moderate
Laparoscopic ovarian drilling	Letrozele	Infertility due to PCOS	Rate of ovarian hyperstimulation syndrome	RD=0 (-0.01-0.01)	High
Transcervical resection of endometrium using rollerball coagulation	Hormone therapy or antifibrinolytic	Heavy menstrual bleeding	Control of bleeding (cure or improvement to acceptable level) 5 y	RR=1.14 (0.97-1.34)	Very Low
Transcervical resection of endometrium using rollerball coagulation	Hormone therapy or antifibrinolytic	Heavy menstrual bleeding	Overall satisfaction with treatment 5 y	RR=1.13 (0.94-1.37)	Very Low
Ophthalmology					
Amniotic membrane transplantation and medication	Lubrication, Antibiotics and Pressure lowering medication	Acute ocular burns	Epithelial defect 21 d post-injury	RR=0.71 (0.27-1.85)	Low
Argon laser trabeculoplasty	IOP reducing medication	Open angle glaucoma	Visual field progression	RR=0.7 (0.42-1.16)	
Argon laser trabeculoplasty	IOP reducing medication	Open angle glaucoma	Optic neuropathy progression	RR=0.71 (0.38-1.34)	

Surgical arm	Drug arm	Disease	Outcome	Treatment effect (95% CI)	GRADE assessment
Laser surgery	intravitreal anti-VEGF	Pathological myopia	Proportion of participants with a gain of 3+ lines in BCVA at 1 y	RR=0.32 (0.08-1.33)	Low
Surgical correction	Botulinum toxin	Strabismus	Improved ocular alignment > 10 dioptres, children	RR=1.1 (0.86-1.41)	Low
Orthopaedic surgery					
Arthroscopic surgery	Sclerosing injection	Jumper's knee	Withdrawal rate	OR=1 (0.06-16.89)	Very Low
Open surgery	Corticosteroid injection	Trigger finger	Resolution of triggering	RR=1.48 (0.79-2.76)	Very low
Open section of the carpal ligament	NSAID and splinting or corticosteroid injections	Carpal tunnel syndrome	Improvement in clinical symptoms at three months of follow-up	RR=1.09 (0.91-1.32)	
Surgical rotator cuff repair	Non-operative treatment including corticosteroid injection and exercise	Rotator cuff tear	Pain (VAS) 12 mo	MD=-0.49 (-1.02-0.05)	Moderate
Otolaryngology					
Surgical orbital decompression	IV Methylprednisolone 1x3 followed by oral prednisolone	Thyroid eye disease	Proportion of successes compared to the proportion of treatment failures as defined by the study authors based on the use of composite outcome scores	RR=0.16 (0.01-1.98)	
Open thoracotomy	Thoracostomy drainage (with fibrinolytics)	Thoracic surgery Pleural empyema	Mortality	RR=NA (NA-NA)	Moderate
VATS	Thoracostomy drainage (with fibrinolytics)	Pleural empyema	Mortality	RR=0.8 (0.04-14.89)	Low
Urology					
Surgical reimplantation of ureters	Antibiotics	Primary vesicoureteric reflux	Rate of patients with symptomatic UTI	RR=0.95 (0.67-1.35)	
Vascular surgery					
Surgery including primary amputation	Thrombolysis (w/ rt-Pa or urokinase)	Acute limb ischaemia	Limb salvage (30 d)	OR=0.89 (0.27-2.91)	Low
Saphenofemoral disconnection	Therapeutic LMWH	Superficial thrombophlebitis	Symptomatic VTE	RR=5 (0.25-100)	
Aspirin and carotid surgery	Aspirin	Carotid stenosis	Major bleeding Ipsilateral ischaemic stroke, and any operative stroke or death near occlusion	RR=NA RR=0.89 (0.6-1.32)	Moderate

Abbreviations

RR: risk ratio

OR: odds ratio
HR: hazard ratio
MD: mean difference
SMD: standardized mean difference

BCC: basal cell carcinoma of the skin
GERD: Gastro-oesophageal reflux disease
GTN: glyceryl tri-nitrate
IOP: intra-ocular pressure
PCOS: polycystic ovarian syndrome
QOL: Quality of life