Journal of Clinical Epidemiology 170 (2024) 111335

Journal of Clinical Epidemiology

ORIGINAL RESEARCH

Pilot and feasibility trials in surgery are incompletely reported according to the CONSORT checklist: a meta-research study

Tyler McKechnie[a,b,*], Tania Kazi[c], Austine Wang[c], Sophia Zhang[c], Alex Thabane[b], Keean Nanji[b,d], Phillip Staibano[b,e], Lily J. Park[a,b], Aristithes Doumouras[a,b,c], Cagla Eskicioglu[a,c], Lehana Thabane[b,f,g], Sameer Parpia[b,h], Mohit Bhandari[b,i]

[a]*Division of General Surgery, Department of Surgery, McMaster University, Hamilton, Ontario, Canada*
[b]*Department of Health Research Methodology, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada*
[c]*Michael G. DeGroote School of Medicine, McMaster University, Hamilton, Ontario, Canada*
[d]*Division of Ophthalmology, Department of Surgery, McMaster University, Hamilton, Ontario, Canada*
[e]*Division of Otolaryngology Head and Neck Surgery, Department of Surgery, McMaster University, Hamilton, Ontario, Canada*
[f]*Biostatistics Unit, St Joseph's Healthcare Hamilton, Hamilton, Ontario, Canada*
[g]*Faculty of Health Sciences, University of Johannesburg, Johannesburg, South Africa*
[h]*Department of Oncology, McMaster University, Hamilton, Ontario, Canada*
[i]*Division of Orthopedic Surgery, Department of Surgery, McMaster University, Hamilton, Ontario, Canada*

Accepted 20 March 2024; Published online 27 March 2024

Abstract

Objectives: Given the key role that pilot and feasibility (PAF) trials play in addressing the challenges of surgical trials, adequate reporting completeness is essential. Our aim was to assess completeness of reporting and evaluate if the items of the Consolidated Standards of Reporting Trials (CONSORT) extension for PAF trials have been reported in surgical PAF trials.

Study Design: This is a metaresearch study reported according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. Studies were included if they were pilot or feasibility randomized trials evaluating a surgical intervention. The primary outcome was overall adherence to the CONSORT statement extension to randomized PAF trials checklist. A Poisson regression was performed to explore the association between research factors and reporting completeness.

Setting: MEDLINE, Embase, and Cochrane Central Register of Controlled Trials were searched from January 1-December 31, 2011 and 2021.

Results: After screening 1991 citations, 38 studies from 2011 to 34 studies from 2021 were included. The mean CONSORT reporting score across all included studies was 21.5 (standard deviation 6.3). After excluding items that were not applicable to all studies, a mean of 20.1 (standard deviation 6.1) of 34 items (0.59) were reported. Studies published in 2021 (vs 2011) did not have a greater number of CONSORT items reported (incidence rate ratio [IRR] 1.01, 95% confidence interval [CI] 0.89−1.15). Studies registered in a clinical trial registry (IRR 1.29, 95% CI 1.12−1.48) and randomizing more than 50 patients (IRR 1.16, 95% CI 1.04−1.30) were associated with more CONSORT items reported.

Conclusion: The reporting completeness of surgical PAF trials is poor and has not improved after the publication of the CONSORT extension.

Keywords: Pilot and feasibility trials; Pilot projects; Feasibility studies; Surgery; Methodological review; Systematic review

Source of Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

PROSPERO Registration: CRD42023475512.

* Corresponding author. Division of General Surgery, Department of Surgery, McMaster University, 1280 Main Street West, Hamilton, Ontario, Canada, L8S 4L8.

E-mail address: tyler.mckechnie@medportal.ca (T. McKechnie).

https://doi.org/10.1016/j.jclinepi.2024.111335

Plain language summary

Pilot and feasibility studies are not reported completely as per reporting guidelines in fields of study such as nephrology and psychiatric. The reporting completeness of these types of studies in surgery has never been examined. We designed this study to evaluate previously published pilot and feasibility studies in surgery to assess their completeness of reporting. We found that only approximately half of the items in the reporting guideline checklist were routinely reported. This proportion of items reported has not improved since the publication of the reporting guideline checklist. Research factors that were associated with more complete reporting included registration of the study prospectively in a database, publication of a study protocol, and randomization of more than 50 patients. Similar to other fields of medical research, there remains work to be done to improve reporting completeness of pilot and feasibility studies.

1. Introduction

The global volume of surgical procedures has seen a consistent increase, rising from 234.2 million in 2004 to approximately 312.9 million in 2012 [1]. These numbers were substantially impacted by the Coronavirus disease 2019 (COVID-19) pandemic, but surgical case volumes have since increased back to prepandemic levels [2]. Consequently, the volume of surgical research has grown substantially. Although most of this research takes the form of observational studies, the utilization of surgical randomized controlled trials (RCTs) for generating surgical evidence has become increasingly prevalent [3]. The number of published surgical RCTs increased 50% between 1999 and 2009, from 300 to 450 [3]. Other reviews from the early 2000s have indicated a surge in published RCTs by as much as 100%−250% compared to the 1990s [4,5].

Despite the increase in surgical RCT conduct, researchers continue to face challenges in conducting high-quality RCTs in surgery. Methodological issues unique to surgical trials include standardization of surgical procedures, difficulty recruiting large numbers of patients, and the impact of surgeon preference, among others [6]. Relatively straightforward methodology in RCTs, such as blinding and recruitment, is often more complex in surgical trials [6]. Given these unique challenges, coupled with the significant amount of time and resources required to successfully complete a high-quality surgical RCT, there has been increased advocacy for conducting pilot and feasibility (PAF) trials prior to definitive RCTs [7−9].

The Consolidated Standards of Reporting Trials (CONSORT) statement, initially published in 1996, aimed to provide authors with guidance for enhancing the reporting of their RCTs [10]. With the evolution and increased interest in RCT methodology over the years, extensions to the original CONSORT statement have been developed to address specific nuances relevant to different types of RCTs. In 2016, the CONSORT extension to randomized PAF trials was introduced; a 40-item checklist applicable to any RCT wherein a future definitive RCT, or a component thereof, is scaled down to assess feasibility regardless of design (eg, factorial, cluster) [11]. While previous metaresearch studies have assessed the reporting completeness of pilot RCTs in chronic kidney disease and psychiatric illness, no such study has been conducted for surgical pilot RCTs [12,13]. These previous metaresearch studies have highlighted shortcomings in the reporting completeness of PAF trials, with approximately 50% of CONSORT items being reported [13]. Given the key role that PAF trials play in addressing the challenges of surgical trials, adequate reporting completeness is essential. Therefore, our aim was to assess completeness of reporting and evaluate if the items of the CONSORT extension for PAF trials have been reported in surgical PAF trials. We hypothesized that PAF randomized trials in surgery would be poorly reported as per the CONSORT extension to randomized PAF trials [12,13].

2. Methods

2.1. Study design

This study is a metaresearch study of reporting completeness of pilot or feasibility randomized trials for surgical interventions reported according to adapted Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines for reporting methodology research (Appendix 1) [14]. The study protocol was registered prospectively on the International Prospective Register for Systematic Reviews (CRD42023475512) and has been submitted for publication in *British Medical Journal Open* (Manuscript Number: bmjopen-2024-085293.R1). Local ethics review board approval was not required.

2.2. Eligibility criteria

Articles were eligible for inclusion if they were pilot or feasibility randomized trials evaluating a surgical intervention and were published between January 1, 2011 and December 31, 2011 or January 1, 2021 and December 31, 2021. A pilot or feasibility randomized trial was defined as any study that self-described as a pilot or feasibility trial in their title, abstract, introduction, or methods section that

What is new?

Key findings

- The reporting completeness of current surgical pilot and feasibility trials is poor.
- Reporting has not improved since the 2016 publication of the CONSORT guidelines.
- Registration in trial databases was associated with greater reporting completeness.
- Sample sizes more than 50 were associated with greater reporting completeness.

What this adds to what was known?

- This is the first study assessing reporting completeness of pilot trials in surgery.

randomized patients prospectively to trial interventions [12]. We defined a surgical intervention as any invasive intervention performed by a physician in the operating room or angiography suite requiring regional or general anesthesia. Studies described as definitive randomized trials or studies that were not explicitly defined as a pilot or feasibility randomized trial by the study authors as above were excluded. Non-English studies were included if they otherwise met inclusion criteria.

2.3. Search strategy

We searched the following databases from January 1, 2011 to December 31, 2011 and January 1, 2021 to December 31, 2021: MEDLINE, Embase, and Cochrane Central Register of Controlled Trials. The search was designed and conducted by a medical research librarian using Peer Review of Electronic Search Strategies [15] with input from study investigators. The search strategy included subject headings and keywords, including ''Pilot Projects,'' ''Feasibility Studies,'' ''Randomized Clinical Trial,'' and ''Surgery'' (Appendices 2-3).

2.4. Study selection

Four reviewers (T.M., T.K., A.W., and S.Z.) independently evaluated the systematically searched titles and abstracts using Covidence. Discrepancies that occurred during the title and abstract screening phases were resolved by inclusion of the study into the full-text screening phase. Full-text screening ensued with four reviewers (T.M., T.K., A.W., and S.Z.). At the full-text screening stage discrepancies were resolved by consensus between the reviewers. If disagreement persists, an additional reviewer was consulted (M.B.).

2.5. Data collection

Data extraction was conducted independently and in duplicate by four reviewers (T.M., T.K., A.W., and S.Z.) into a data collection form designed a priori and pilot-tested on Microsoft Excel. Any discrepancies were resolved by consensus between the reviewers. The extracted data included trial characteristics (eg, author, year of publication, journal of publication, journal impact factor, study period, number of included centers, intervention details, primary outcomes, secondary outcomes, statistical analyses, study protocol publication, study registration in a clinical trial database, number of patients randomized), journal characteristics (eg, journal impact factor, explicit mention of CONSORT endorsement on the journal website), and reporting details (eg, number of items reported in the study as per the CONSORT extension). Journal-specific items were extracted directly from the website of the given journal. For the reporting of individual items on the CONSORT extension for PAF trials, each item was counted as either ''reported'' or ''not reported'', apart from items 6c, 7b, 11a, 11b, 18, and 19a, all of which could have also been counted as ''not applicable''.

2.6. Outcome measures

The primary outcome was completeness of reporting according to the CONSORT statement extension to randomized PAF trials checklist. This was defined as the number and proportion of trials reporting each of the 40 items from the CONSORT extension to randomized PAF trials checklist. The secondary outcomes were completeness of reporting according to the CONSORT statement extension to randomized PAF trials checklist for each individual trial, defined as the number of items reported for each included study and categorized based on the reporting item (eg, title and abstract, introduction, methods, results, discussion) and research factors associated with the completeness of reporting of randomized PAF trials.

2.7. Sample size calculation

A sample size calculation was performed a priori using a 95% confidence interval (CI) approach [16]. With a margin of error (E) of 0.15, and assuming that the proportion of studies with adequate reporting as per the CONSORT statement extension to randomized PAF trials checklist is 0.50 (P_0), a minimum of 44 PAF trials was required for adequate estimation of the primary outcome.

2.8. Statistical analyses

The completeness of reporting was summarized with descriptive statistics (ie, proportion of articles reporting each CONSORT statement item with corresponding 95% CIs estimated using Agresti-Coull intervals) for the primary outcome. For items 6c, 7b, 11a, 11b, 18, and 19a, the

proportions were calculated based on the total number of studies for which the item was applicable.

We evaluated the secondary outcomes using descriptive statistics, reported as the mean and standard deviation (SD) of the total number of CONSORT items reported. To calculate the mean number of CONSORT items reported, ''not applicable'' responses to reporting items 6c, 7b, 11a, 11b, 18, and 19a were excluded. We conducted a Poisson regression to explore the association between the following factors and completeness of reporting as measured by the number of reported CONSORT items per study: year of publication, industry funding, registration in a clinical trial database, multisite study, journal endorsement of CONSORT, published a pilot RCT protocol, and number of patients randomized. Based on previously published data, it was hypothesized that publication after creation of the CONSORT extension to randomized PAF trials, industry funding, multicenter studies, journal endorsement of CONSORT, and a larger sample size would be associated with an increased number of reported CONSORT items [17−19]. Additionally, other covariates that were anticipated to positively predict the completeness of reporting of PAF trials included publication of a study protocol and registration in a clinical trial database [20−22]. The results of the Poisson regression were reported with adjusted incidence rate ratios (IRRs) with associated 95% CIs. A sensitivity analysis with a generalized Poisson regression model was performed. Descriptive statistics were used to characterize the sample population.

3. Results

3.1. Results of the search

Database searches identified 2242 records. After excluding 331 duplicates, 1991 records were available for screening. After assessing 193 full texts, 72 studies were included. The reasons for exclusion of the studies excluded at the full-text review are recorded in Supplemental Table 1. A Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram is illustrated in Figure 1.

3.2. Characteristics of the included studies

Table 1 presents the pooled characteristics of the included studies stratified by study year. Supplemental Tables 2 and 3 present study characteristics for each of the individual included studies. Overall, 47.2% of included studies were published in 2021. The surgical subspecialties evaluated included general surgery (26.4%), orthopedic surgery (15.3%), urology (13.9%), cardiovascular surgery (9.7%), gynecology (8.3%), and ophthalmology (8.3%). The median number of patients approached for consideration of inclusion was 78 (interquartile range [IQR] 51−190), the median number of patients randomized was 43.5 (IQR 21.5−54.5), and the median number of patients completing trial follow-up was 40 (IQR 20−52.5). Twenty-one studies (29.2%) stated that their PAF trial was reported according to the CONSORT guidelines.

3.3. Reporting completeness based on CONSORT extension

The mean CONSORT reporting score across all included studies was 21.5 (SD 6.3), correlating with a proportion of 0.56 of the items. After excluding the items with ''not applicable'' options, the mean CONSORT reporting score was 20.1 (SD 6.1) of 34 items (0.59). In the studies published in 2011, the mean CONSORT reporting score was 19.4 (SD 5.8) compared to a mean of 23.9 (SD 6.2) for studies published in 2021. The items reported by the largest proportion of studies were ''4a. Eligibility criteria for participants'' (n = 70/72, 97.2%, 95% CI 90.3−99.7%) and ''5. The interventions for each group with sufficient details to allow replication, including how and when they were actually administered'' (n = 63/72, 87.5%, 95% CI 77.6−94.1%) (Table 2). The items reported by the smallest proportion of studies were ''7b. When applicable, explanation of any interim analyses and stopping guidelines'' (n = 3/46, 6.5%, 95% CI 1.4−17.9%) and ''6b. Any changes to pilot trial assessments or measurements after the pilot trial commenced, with reasons'' (n = 6/72, 8.3%, 95% CI 3.1−17.3%) (Table 2). The overall proportion of adequate reporting per item in the title and abstract section ranged from 69.4% to 81.9%, from 80.6% to 83.3% in the introduction section, from 6.5% to 97.2% in the methods section, from 19.4% to 83.3% in the results section, and from 15.3% to 80.6% in the discussion section.

3.4. Factors associated with CONSORT extension reporting

The adjusted IRRs for overall CONSORT adherence by study characteristic are reported in Table 3. The year of publication (ie, published in 2021 vs 2011) did not influence the number of reported CONSORT items (adjusted IRR 1.01, 95% CI 0.89−1.15). Registration in clinical trial databases (adjusted IRR 1.29, 95% CI 1.12−1.48) and sample sizes more than 50 at randomization (adjusted IRR 1.16 95% CI 1.04−1.30) were associated with increased reporting completeness. The remaining study characteristics were inconclusively associated with reporting completeness. Findings from the sensitivity analysis were similar (Supplemental Table 4).

4. Discussion

This was the first metaresearch study examining the reporting completeness of PAF trials in surgery. In keeping with our a priori hypothesis, reporting of pilot RCTs in surgery was poor. The mean proportion of reported CONSORT extension for randomized PAF trial items was 0.56.

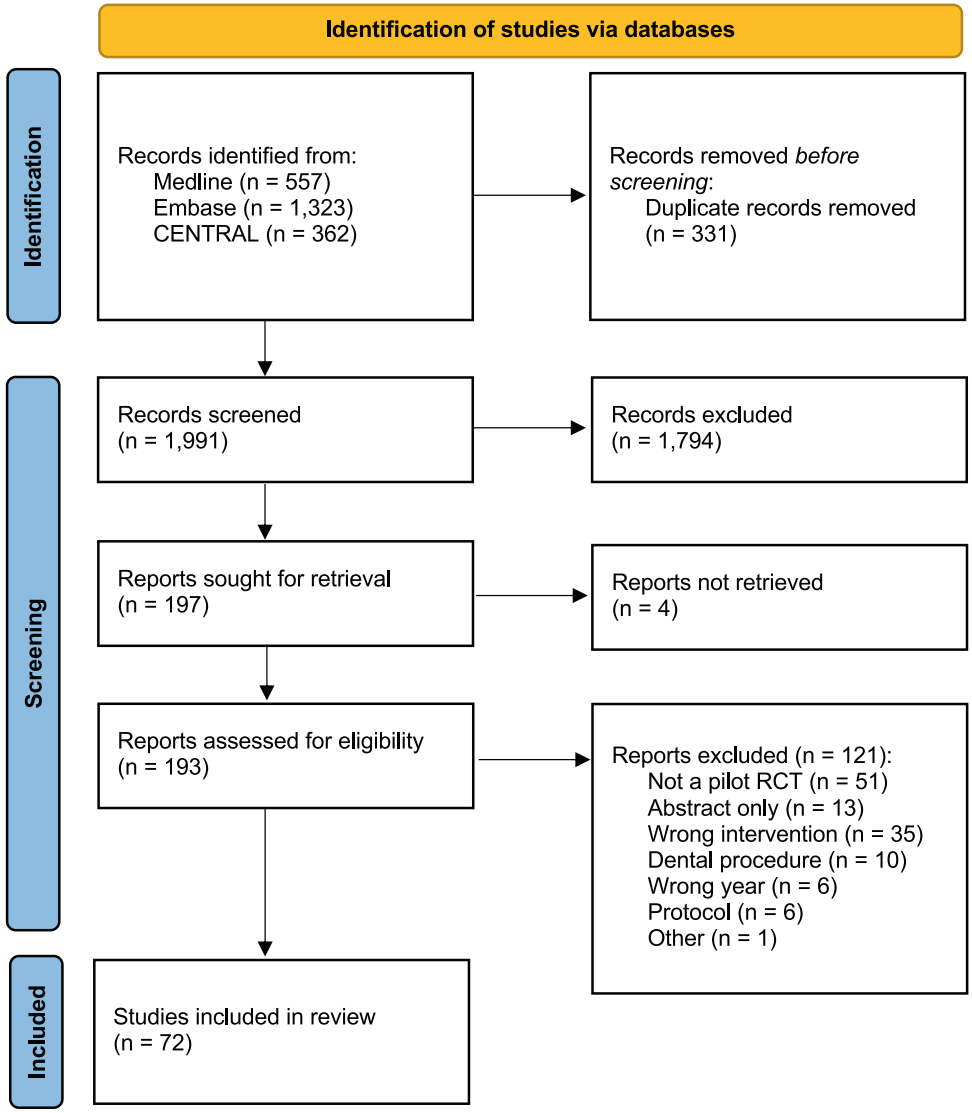

Figure 1. PRISMA diagram. Transparent reporting of systematic reviews and meta-analysis flow diagram outlining the search strategy results from initial search to included studies. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

This is similar to previously published proportions in chronic kidney disease literature (0.54) and psychotherapy interventions (0.52) [12,13]. Studies that were registered in a clinical trial database and studies with 50 or more patients randomized were associated with 29.0% and 15.9% increases in the number of CONSORT items reported, respectively.

Across our sample of surgical pilot RCTs, some CONSORT items were consistently reported across studies. These items included eligibility criteria, intervention details, trial objectives, and baseline patient demographics. These items are fairly similar to those found in the original CONSORT statement for randomized trials that was published in 1996 [10]. Reporting these research components has been common practice among surgical researchers for decades. However, CONSORT items that have not been ubiquitously recommended for such a prolonged period, such as the publication of a trial protocol and explanation of stopping guidelines, were reported in less than 20% of included studies [11]. Similarly, CONSORT items unique to pilot trials, such as changes to pilot trial methodology after trial commencement and prespecification of progression criteria, were reported in less than 20% of included studies. The lack of reporting prespecified progression criteria to a definitive trial is particularly alarming, as often times the purpose of PAF trials is to aid in the development of performing a rigorously designed and adequately powered RCT [23]. This may be difficult without well-defined progression criteria establishing the required changes to facilitate the achievement of critical feasibility aspects of the definitive trial. This could explain the low number of included PAF trials that stated they were preludes to definitive RCTs (22.2%). Such a large proportion of pilot trials failing to progress to definitive RCTs would indicate that

Table 1. Pooled study characteristics

	Overall ($N = 72$)		2021 ($N = 34$)		2011 ($N = 38$)	
Study characteristic	***N* studies**	**%**	***N* studies**	**%**	***N* studies**	**%**
Location of Corresponding Author						
North America	19	26.4	8	23.5	11	28.9
Europe	27	37.5	13	38.2	14	36.8
Asia	17	23.6	9	26.5	8	21.1
Russia	2	2.8	2	5.9	0	0.0
Other	7	9.7	2	5.9	5	13.2
Type of Journal						
Clinical	65	90.3	29	85.3	36	94.7
Methodological	7	9.7	5	14.7	2	5.3
Open Access Publication						
Yes	30	41.7	17	50	13	34.2
No	42	58.3	17	50	25	65.8
Type of RCT						
Parallel	69	95.8	31	91.2	38	100
Crossover	2	2.8	2	5.9	0	0.0
Factorial	1	1.4	1	2.9	0	0.0
Type of Surgery						
Cardiovascular	7	9.7	4	11.8	3	7.9
Orthopedics	11	15.3	6	17.6	5	13.2
General Surgery	19	26.4	8	23.5	11	28.9
Other	35	48.6	16	47.1	19	50.0
Number of Centers						
Single center	48	66.7	18	52.9	30	78.9
Multicenter	24	33.3	16	47.1	8	21.1
Sample Size						
Equal to or Less than 50	37	51.4	10	29.4	27	71.1
More than 50	35	48.6	24	70.6	11	28.9
Prelude to a Definitive RCT						
Yes	16	22.2	9	26.5	7	18.4
No	56	77.8	25	73.5	31	81.6
Study States Reporting as per CONSORT						
Yes	21	29.2	18	52.9	3	7.9
No	51	70.8	16	47.1	35	92.1
Study States Reporting as per CONSORT for Pilot RCTs						
Yes	-	-	5	14.7	-	-
No	-	-	29	85.3	-	-
Journal Endorses CONSORT Statement						
Yes	46	63.9	25	73.5	21	55.3
No	26	36.1	9	26.4	17	44.7
Industry Funded						
Yes	13	18.1	4	11.8	9	23.7
No	58	80.6	29	85.3	29	76.3

CONSORT, Consolidated Standards of Reporting Trials; *N*, number of studies; RCT, randomized controlled trial.

a large number of trials were deemed not feasible or that a significant portion of these trials were evaluating efficacy outcomes as opposed to feasibility outcomes. The latter is more likely given that 79.2% of the included trials formed conclusions based on clinical outcomes from the trials as opposed to conclusions on trial feasibility. Unfortunately, given the relatively small sample size of most included studies, the resultant clinical conclusions are likely based on underpowered analyses. Not only does this potentially bias our body of surgical RCT literature but also serves

Table 2. Pooled reporting of items on the CONSORT extension for pilot and feasibility trials

	Overall			2021			2011		
CONSORT item	***N* studies**	**%**	**95% CI**	***N* studies**	**%**	**95% CI**	***N* studies**	**%**	**95% CI**
Title and Abstract									
1a. Identification as a pilot or feasibility randomized trial in the title	50/72	69.4	57.5–79.8	26/34	76.5	58.8–89.3	24/38	63.2	46.0–78.2
1b. Structured summary of pilot trial design, methods, results, and conclusions	59/72	81.9	71.1–90.0	27/34	79.4	62.1–91.3	32/38	84.2	68.7–94.0
Introduction									
2a. Scientific background and explanation of rationale for future definitive trial, and reasons for randomized pilot trial	58/72	80.6	69.5–88.9	23/34	67.6	49.5–82.6	35/38	92.1	78.6–98.3
2b. Specific objectives or research questions for pilot trial	60/72	83.3	72.7–91.1	24/34	70.6	52.5–84.9	36/38	94.7	82.3–99.4
Methods									
Trial Design									
3a. Description of pilot trial design (such as parallel, factorial) including allocation ratio	34/72	47.2	35.3–59.3	13/34	38.2	22.2–56.4	21/38	55.3	38.3–71.4
3b. Important changes to methods after pilot trial commencement (such as eligibility criteria), with reasons	10/72	13.9	6.9–24.1	10/34	29.4	15.1–47.5	0/38	0.0	0–9.3
Participants									
4a. Eligibility criteria for participants	70/72	97.2	90.3–99.7	34/34	100	89.7–100	36/38	94.7	82.3–99.4
4b. Settings and locations where the data were collected	49/72	68.1	56.0–78.6	28/34	82.4	65.5–93.2	21/38	55.3	38.3–71.4
4c. How participants were identified and consented	36/72	50.0	38.0–62.0	19/34	55.9	37.9–72.8	17/38	44.7	28.6–61.7
Interventions									
5. The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	63/72	87.5	77.6–94.1	28/34	82.4	65.5–93.2	35/38	92.1	78.6–98.3
Outcome Measurement									
6a. Completely defined prespecified assessments or measurements to address each pilot trial objective specified in 2b, including how and when they were assessed	54/72	75.0	63.4–84.5	27/34	79.4	62.1–91.3	27/38	71.1	54.1–84.6
6b. Any changes to pilot trial assessments or measurements after the pilot trial commenced, with reasons	6/72	8.3	3.1–17.3	6/34	17.6	6.8–34.5	0/38	0.0	0–9.3

(*Continued*)

Table 2. Continued

CONSORT item	Overall			2021			2011		
	N studies	%	95% CI	*N* studies	%	95% CI	*N* studies	%	95% CI
6c. If applicable, prespecified criteria used to judge whether, or how, to proceed with future definitive trial[a]	10/58	17.2	8.6–29.4	10/34	29.4	15.1–47.5	0/24	0.0	0–14.2
Sample Size									
7a. Rationale for numbers in the pilot trial	24/72	33.3	22.7–45.4	15/34	44.1	27.2–62.1	9/38	23.7	11.4–40.2
7b. When applicable, explanation of any interim analyses and stopping guidelines[a]	3/46	6.5	1.4–17.9	2/13	15.4	1.9–45.4	1/33	3.0	0.1–15.8
Randomization									
8a. Method used to generate the random allocation sequence	46/72	63.9	51.7–74.9	26/34	76.5	58.8–89.3	20/38	52.6	35.8–69.0
8b. Type of randomization(s); details of any restriction (such as blocking and block size)	28/72	38.9	27.6–51.1	15/34	44.1	27.2–62.1	13/38	34.2	19.6–51.4
Allocation Concealment Mechanisms									
9. Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	36/72	50.0	38.0–62.0	19/34	55.9	37.9–72.8	17/38	44.7	28.6–61.7
Implementation									
10. Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	25/72	34.7	23.9–46.9	14/34	41.2	24.6–59.3	11/38	28.9	15.4–45.9
Blinding									
11a. If done, who was blinded after assignment to interventions (eg, participants, care providers, those assessing outcomes) and how[a]	38/60	63.3	49.9–75.4	16/23	69.6	47.1–86.8	22/37	59.5	42.1–75.2
11b. If relevant, description of the similarity of interventions[a]	44/54	81.5	68.6–90.7	15/23	65.2	42.7–83.6	29/31	93.5	78.6–99.2
Statistical Methods									
12. Methods used to address each pilot trial objective whether qualitative or quantitative	60/72	83.3	72.7–91.1	29/34	85.3	68.9–95.0	31/38	81.6	65.7–92.3
Results									
Participant Flow									
13a. For each group, the numbers of participants who were approached and/	35/72	48.6	36.7–60.7	25/34	73.5	55.6–87.1	10/38	26.3	13.4–43.1

(*Continued*)

Table 2. Continued

CONSORT item	Overall			2021			2011		
	N studies	%	95% CI	*N* studies	%	95% CI	*N* studies	%	95% CI
or assessed for eligibility, randomly assigned, received intended treatment, and were assessed for each objective									
13b. For each group, losses and exclusions after randomization, together with reasons	47/72	65.3	53.1–76.1	24/34	70.6	52.5–84.9	13/38	34.2	19.6–51.4
Recruitment									
14a. Dates defining the periods of recruitment and follow-up	57/72	79.2	68.0–87.8	29/34	85.3	68.9–95.0	28/38	73.6	56.9–86.6
14b. Why the pilot trial ended or was stopped	14/72	19.4	11.1–30.5	8/34	23.5	10.7–41.2	6/38	15.8	6.0–31.3
Baseline Data									
15. A table showing baseline demographic and clinical characteristics for each group	60/72	83.3	72.7–91.1	32/34	94.1	80.3–99.3	28/38	73.7	56.9–86.6
Numbers Analyzed									
16. For each objective, number of participants (denominator) included in each analysis. If relevant, these numbers should be by randomized group	51/72	70.8	58.9–81.0	27/34	79.4	62.1–91.3	24/38	63.2	46.0–78.2
Outcomes and Estimation									
17. For each objective, results including expressions of uncertainty (such as 95% confidence interval) for any estimates. If relevant, these results should be by randomized group	54/72	75.0	63.4–84.5	24/34	70.6	52.5–84.9	30/38	78.9	62.7–90.4
Ancillary Analyses									
18. Results of any other analyses performed that could be used to inform the future definitive trial[a]	24/72	33.3	22.7–45.4	10/34	29.4	15.1–47.5	14/38	36.8	21.8–54.0
Harms									
19. All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	50/72	69.4	57.5–79.8	21/34	61.8	43.6–77.8	29/38	76.3	59.8–88.6
19a. If relevant, other important unintended consequences[a]	9/27	33.3	16.5–54.0	4/9	44.4	13.7–78.8	5/18	27.7	9.7–53.5
Discussion									
Limitations									
20. Pilot trial limitations, addressing sources of potential bias and remaining uncertainty about feasibility	50/72	69.4	57.5–79.8	30/34	88.2	72.5–96.7	20/38	52.6	35.8–69.1

(*Continued*)

Table 2. Continued

CONSORT item	Overall			2021			2011		
	N studies	%	95% CI	*N* studies	%	95% CI	*N* studies	%	95% CI
Generalizability									
21. Generalizability (applicability) of pilot trial methods and findings to future definitive trial and other studies	50/72	69.4	57.5–79.8	16/34	47.1	29.8–64.9	34/38	89.5	75.2–97.1
Interpretation									
22. Interpretation consistent with pilot trial objectives and findings, balancing potential benefits and harms, and considering other relevant evidence	58/72	80.6	69.5–88.9	28/34	82.4	65.5–93.2	30/38	78.9	62.7–90.4
22a. Implications for progression from pilot to future definitive trial, including any proposed amendments	19/72	26.4	16.7–38.1	17/34	50.0	32.4–67.6	2/38	5.3	0.6–17.7
Other Information: Registration									
23. Registration number for pilot trial and name of trial registry	35/72	48.6	36.7–60.7	24/34	70.6	52.5–84.9	9/38	23.7	11.4–40.2
Protocol									
24. Where the pilot trial protocol can be accessed, if available	11/72	15.3	7.9–25.7	10/34	29.4	15.1–47.5	1/38	2.6	0.1–13.8
Funding									
25. Sources of funding and other support (such as supply of drugs), role of funders	46/72	63.9	51.7–74.9	31/34	91.2	76.3–98.1	15/38	39.5	24.0–56.6
26. Ethical approval or approval by research review committee, confirmed with reference number	30/72	41.7	30.2–53.9	27/34	79.4	62.1–91.3	3/38	7.9	1.7–21.4

CI, confidence intervals; CONSORT, Consolidated Standards of Reporting Trials; *N*, number.
[a] Total *N* only accounted for studies in which the item was applicable.

as an inefficient use of valuable and finite research resources [24,25].

Interestingly, our results suggest that the 2016 publication of the CONSORT extension for randomized PAF trials did not impact the reporting completeness of pilot RCTs [11]. Rather, other research factors such as registration in a clinical trial database, publication of protocols, and inclusion of a larger number of patients (ie, more than 50) were associated with improved reporting. In contemporary clinical research, the preregistration of pilot trials in recognized clinical trial databases and the subsequent publication of study protocols play pivotal roles in ensuring transparency, methodological rigor, and the overall advancement of evidence-based medicine [26,27]. Registering pilot trials in reputable databases fosters transparency by providing a comprehensive record of ongoing and completed studies, thereby reducing publication bias and selective reporting [27]. Furthermore, the publication of study protocols in peer-reviewed journals serves as a critical component of disseminating methodological details, objectives, and planned analyses, enabling stakeholders to critically evaluate the study design and execution [26,28]. Yet, adherence to these practices remains limited. A recent metaresearch study found that only 36.2% of published RCTs had publicly available protocols [29]. Journals should strongly encourage researchers to make their protocol available publicly prior to accepting pilot RCTs for publication. This, in combination with ongoing endorsement of CONSORT reporting checklists, may lead to substantial improvements in reporting completeness.

The strengths of this metaresearch study include the Poisson regression that identified modifiable factors that may improve reporting, adequate statistical power for the primary outcome, and the publication of a protocol a priori. The limitations include sampling bias, lack of statistical power for comparative analysis, and heterogeneity across study

Table 3. Incidence rate ratios for the total number of CONSORT pilot trial extension items reported

Variable	Adjusted IRR (95% CI)
Published in 2021 (vs 2011)	1.01 (0.89–1.15)
Industry funded (vs not industry funded)	0.91 (0.81–1.02)
Journal endorsement of CONSORT (vs no endorsement)	1.03 (0.91–1.17)
Registered in a clinical trial database (vs not registered)	1.29 (1.12–1.48)
Published study protocol (vs no published protocol)	1.23 (0.99–1.54)
Multisite study (vs single-center study)	1.02 (0.87–1.19)
More than or equal to 50 patients randomized (vs less than 50)	1.16 (1.04–1.30)

characteristics. We elected to only include surgical pilot RCTs published in 2011 and 2021 based on their relation to the publication of the CONSORT extension to randomized PAF trials (ie, 2016) [11]. Altman et al. suggest that it takes 4–6 years for a publication to become highly cited [30]. Therefore, it was assumed that the CONSORT extension to randomized PAF trials would have been integrated into clinical trials by 2021. The year 2011 was chosen as a comparator to assess the impact of the publication of the CONSORT extension to randomized PAF trials. However, it is possible that through only sampling articles from 1 calendar year following the publication of the CONSORT extension to randomized PAF trials, we captured data which were not necessarily representative of all surgical pilot RCTs published since the CONSORT extension. Moreover, given that the comparisons were not based on randomization, it is likely that the 95% CIs derived in the present study likely underestimate the uncertainty associated with these effect estimates. From a statistical perspective, this methodological review was designed to assess the certainty of the proportion of CONSORT items reported (ie, our primary outcome) using a 95% CI approach. It was not, however, powered for comparative analyses between years (ie, 2011 and 2021); therefore, the Poisson regression may have been underpowered to detect a difference in reporting completeness. The definition of surgical intervention was broad and nonspecific such that there was a wide range of surgical procedures included in this study. Moreover, studies were included across a variety of geographic locations and journals with wide ranging impact factors. This may have introduced between study heterogeneity and thus perhaps impacted measures of variance. Nonetheless, the broad definition was chosen intentionally to capture as many pilot RCTs as possible pertinent to surgical researchers. While the broad search criteria allowed for good capture of relevant articles, the screening process was not pilot tested; thus, it remains possible that relevant articles were missed. Finally, it must be acknowledged that the CONSORT checklist was not primarily designed to assess completeness of reporting but rather to guide reporting practices for authors. As such, whether this is the most appropriate means of assessing completeness of trial reporting is unknown.

5. Conclusion

The reporting of pilot RCTs in surgery as per the CONSORT extension for randomized PAF trials remains poor despite being widely disseminated in health research methodology literature for the past 7 years. As the number of PAF trials in surgery continues to grow, improved reporting completeness is imperative to improve the utility of these pilot trials. Endorsement of the CONSORT extensions by journals as well as promotion of the importance of registration in clinical trial databases and public availability of research protocols may improve reporting completeness. Future research should aim to identify other strategies to improve surgical pilot trial reporting completeness.

Ethical statement

Local ethics review board approval was not required.

CRediT authorship contribution statement

Tyler McKechnie: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Tania Kazi:** Writing – review & editing, Methodology, Investigation, Data curation. **Austine Wang:** Writing – review & editing, Methodology, Investigation, Data curation. **Sophia Zhang:** Writing – review & editing, Visualization, Methodology, Investigation, Data curation. **Alex Thabane:** Writing – review & editing, Visualization, Methodology, Investigation, Formal analysis, Data curation. **Keean Nanji:** Writing – review & editing, Visualization, Methodology, Investigation, Formal analysis, Data curation. **Phillip Staibano:** Writing – review & editing, Visualization, Validation, Resources, Methodology, Investigation, Formal analysis, Data curation. **Lily J. Park:** Writing – review & editing, Visualization, Resources, Methodology, Investigation, Data curation. **Aristithes Doumouras:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis. **Cagla Eskicioglu:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Data curation. **Lehana Thabane:** Writing – review & editing, Visualization, Validation, Supervision, Resources, Methodology, Investigation, Formal analysis, Conceptualization. **Sameer Parpia:** Writing – review & editing, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **Mohit**

Bhandari: Writing – review & editing, Visualization, Validation, Supervision, Software, Resources, Methodology, Investigation, Data curation, Conceptualization.

Data availability

Derived data supporting the findings of this study are available from the corresponding author on request.

Declaration of competing interest

The authors declare that there are no conflicts of interest.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2024.111335.

References

[1] Weiser TG, Haynes AB, Molina G, Lipsitz SR, Esquivel MM, Uribe-Leitz T, et al. Estimate of the global volume of surgery in 2012: an assessment supporting improved health outcomes. Lancet 2015;385:S11.
[2] Mattingly AS, Rose L, Eddington HS, Trickey AW, Cullen MR, Morris AM, et al. Trends in US surgical procedures and health care system response to policies curtailing elective surgical operations during the COVID-19 pandemic. JAMA Netw Open 2021;4(12):e2138038.
[3] Ahmed Ali U, van der Sluis PC, Issa Y, Habaga IA, Gooszen HG, Flum DR, et al. Trends in worldwide volume and methodological quality of surgical randomized controlled trials. Ann Surg 2013;258(2):199–207.
[4] Brooke BS, Nathan H, Pawlik TM. Trends in the quality of highly cited surgical research over the past 20 years. Ann Surg 2009;249(1):162–7.
[5] Hopewell S, Dutton S, Yu LM, Chan AW, Altman D. The quality of reports of randomised trials in 2000 and 2006. BMJ 2010;340:c723.
[6] Farrokhyar F, Karanicolas PJ, Thoma A, Simunovic M, Bhandari M, Devereaux PJ, et al. Randomized controlled trials of surgical interventions. Ann Surg 2010;251(3):409–16.
[7] McCulloch P, Altman DG, Campbell WB, Flum DR, Glasziou P, Marshall JC, et al. No surgical innovation without evaluation: the IDEAL recommendations. Lancet 2009;374(9695):1105–12.
[8] Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008;337:a1655.
[9] Fairhurst K, Blazeby JM, Potter S, Gamble C, Rowlands C, Avery KNL. Value of surgical pilot and feasibility study protocols. Br J Surg 2019;106(8):968–78.
[10] Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. JAMA 1996;276:637.
[11] Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ 2016;355:i5239.
[12] Kosa SD, Monize J, Leenus A, Leenus S, Samra S, Szwiega S, et al. Reporting quality of pilot clinical trials in chronic kidney disease patients on hemodialysis: a methodological survey. Pilot Feasibility Stud 2019;5(1):53.
[13] Bhatt M, Zielinski L, Sanger N, Shams I, Luo C, Bantoto B, et al. Evaluating completeness of reporting in behavioral interventions pilot trials. Res Soc Work Pract 2018;28(5):577–84.
[14] Murad MH, Wang Z. Guidelines for reporting meta-epidemiological methodology research. Evid Base Med 2017;22(4):139–42.
[15] McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. J Clin Epidemiol 2016;75:40–6.
[16] Isiguzo G, Zunza M, Chirehwa M, Mayosi BM, Thabane L. Quality of abstracts of pilot trials in heart failure: a protocol for a systematic survey. Contemp Clin Trials Commun 2017;8:258–63.
[17] Rios LP, Odueyungbo A, Moitri MO, Rahman MO, Thabane L. Quality of reporting of randomized controlled trials in general endocrinology literature. J Clin Endocrinol Metab 2008;93:3810–6.
[18] Kosa SD, Mbuagbaw L, Borg Debono V, Bhandari M, Dennis BB, Ene G, et al. Agreement in reporting between trial publications and current clinical trial registry in high impact journals: a methodological review. Contemp Clin Trials 2018;65:144–50.
[19] Mbuagbaw L, Thabane M, Vanniyasingam T, Borg Debono V, Kosa S, Zhang S, et al. Improvement in the quality of abstracts in major clinical journals since CONSORT extension for abstracts: a systematic review. Contemp Clin Trials 2014;38:245–50.
[20] Turner L, Shamseer L, Altman DG, Weeks L, Peters J, Kober T, et al. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. Cochrane Database Syst Rev 2012;2013(1):MR000030.
[21] Adam GP, Springs S, Trikalinos T, Williams JW, Eaton JL, Von Isenburg M, et al. Does information from ClinicalTrials.gov increase transparency and reduce bias? Results from a five-report case series. Syst Rev 2018;7(1):59.
[22] Wager E, Elia N. Why should clinical trials be registered? Eur J Anaesthesiol 2014;31(8):397–400.
[23] Thabane L, Ma J, Chu R, Cheng J, Ismaila A, Rios LP, et al. A tutorial on pilot studies: the what, why and how. BMC Med Res Methodol 2010;10:1.
[24] Ahmed Ali U, ten Hove JR, Reiber BM, van der Sluis PC, Besselink MG. Sample size of surgical randomized controlled trials: a lack of improvement over time. J Surg Res 2018;228:1–7.
[25] Raudys SJ, Jain AK. Small sample size effects in statistical pattern recognition; recommendations for practitioners. IEEE Trans Pattern Anal Mach Intell 1991;13(3):252–64.
[26] Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013;158:200.
[27] Viergever RF, Ghersi D. The quality of registration of clinical trials. PLoS One 2011;6:e14701.
[28] Lassere M, Johnson K. The power of the protocol. Lancet 2002;360(9346):1620–2.
[29] Schönenberger CM, Griessbach A, Taji Heravi A, Gryaznov D, Gloy VL, Lohner S, et al. A meta-research study of randomized controlled trials found infrequent and delayed availability of protocols. J Clin Epidemiol 2022;149:45–52.
[30] Altman DG, Goodman SN. Transfer of technology from statistical journals to the biomedical literature. JAMA 1994;272:129–32.