

BMJ Open How pragmatic are randomised controlled trials evaluating minimally invasive surgery for oesophageal cancer? A methodological review of trial design using the Pragmatic-Explanatory Continuum Indicator Summary-2 (PRECIS-2) tool

Katy Chalmers , Sian Cousins , Natalie S Blencowe , Jane Blazeby 

To cite: Chalmers K, Cousins S, Blencowe NS, *et al.* How pragmatic are randomised controlled trials evaluating minimally invasive surgery for oesophageal cancer? A methodological review of trial design using the Pragmatic-Explanatory Continuum Indicator Summary-2 (PRECIS-2) tool. *BMJ Open* 2024;**14**:e078417. doi:10.1136/bmjopen-2023-078417

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<https://doi.org/10.1136/bmjopen-2023-078417>).

Received 01 August 2023
Accepted 22 November 2024



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY. Published by BMJ Group.

Centre for Surgical Research, University of Bristol, Bristol, UK

Correspondence to

Dr Katy Chalmers;
Katy.Chalmers@bristol.ac.uk

ABSTRACT

Background Surgical interventions are inherently complex and designing and conducting surgical randomised controlled trials (RCTs) can be challenging. Trial design impacts the applicability of trial results to clinical practice. Given the recent growth in numbers of surgical RCTs, there is a need to better understand the validity and applicability of trials in this field.

Objectives To examine the applicability and validity of RCTs comparing minimally invasive and open surgery for oesophageal cancer and to delineate areas for future research.

Eligibility criteria RCTs comparing open with minimal invasive oesophagectomy, published January 2012–June 2023. Abstracts, pilot and feasibility studies, and systematic reviews were excluded.

Sources of evidence Three sequential searches of Ovid MEDLINE, Embase and CENTRAL electronic databases and clinical trials registry databases.

Charting methods Two independent reviewers screened the articles and used appropriate, validated tools (Pragmatic-Explanatory Continuum Indicator Summary-2 (PRECIS-2) and Risk of Bias 2) to assess study quality. Trials were considered pragmatic if they were conducted in multiple centres and had a mean score of four or above on the PRECIS-2.

Results Nine RCTs were identified. One was judged to be pragmatic. The remaining eight were limited by narrow eligibility criteria, being single-centred or having strict intervention protocols. Two studies were low risk of bias, of which one was pragmatic, and three high, due to unblinded outcome assessment. The remaining four studies were of 'some concern' due to poor reporting.

Conclusions Only one trial identified in this review was considered pragmatic. More lenient criteria, as used in other reviews, may increase the proportion. There is a need for clearer guidance on the cut-off values that define a trial as pragmatic. It is recommended that the intended purpose of the trial, whether explanatory or pragmatic, receives more attention during surgical trial study design and conduct.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ This review provides a comprehensive assessment of applicability and validity by two reviewers independently, using validated tools.
- ⇒ Due to the original purpose of Pragmatic-Explanatory Continuum Indicator Summary-2 (PRECIS-2), there is no recognised standardised method for using PRECIS-2 scores to define trials as explanatory or pragmatic.
- ⇒ We provide detailed text to ensure transparency about how the PRECIS-2 and Risk of Bias 2 tools were applied.

INTRODUCTION

Surgical interventions are inherently complex. They involve multiple components and co-interventions,^{1–5} delivered by collaborating healthcare professionals with varying expertise, in complex environments. These factors require consideration during the design and conduct of randomised controlled trials (RCTs) as they determine how applicable trial results are to usual practice and how they may be used to inform decision-making.

Validity, which refers to the extent to which the observed treatment effect may be due to biases,⁶ is widely understood and accepted. Pragmaticism (ie, how applicable trial results are to usual care) is a less well recognised concept. A pragmatic trial aims to reflect 'real-world' settings to inform decisions to adopt interventions into clinical practice.⁷ For example, these trials would have multiple experienced surgeons delivering interventions, in numerous centres across a wide

range of eligible patients. In contrast, an explanatory trial is conducted in a tightly controlled, 'idealised' setting,⁷ for example, involving one or two expert surgeons, in a single centre, with a narrow subset of eligible patients.⁸ When assessing evidence from RCTs, it is essential to consider both the validity of the trial and how applicable (pragmatic) the trial is.

Given the recent growth in numbers of surgical RCTs, there is a need to better understand the validity and applicability of trials in this field. The aim of this study was to examine RCTs comparing minimally invasive and open surgery for oesophageal cancer, focusing on applicability to delineate areas for future research. This case study was chosen because it is complex with possible variation in eligibility criteria (eg, exclusion of patients with specific tumour gradings or co-morbidities), intervention delivery (eg, location of anastomoses and extent of lymph node dissection) and contextual variables that may impact intervention delivery and outcomes (eg, surgeon, team and centre experience and training, access to equipment such as robots). All these trial design characteristics can impact the applicability of results, and so it is crucial that it is considered. To our knowledge, none of the existing systematic reviews in this area have considered this issue.^{9–16}

METHODS

Annual literature searches were undertaken to inform the steering committee of an ongoing RCT (Randomised Oesophagectomy: Minimally Invasive or Open (ROMIO))¹⁷ of up-to-date evidence. Eligible studies were assessed for applicability and validity using validated tools. Data were synthesised narratively, and key findings presented.

Study eligibility

Included were RCTs (including ancillary studies, substudies and long-term follow-ups) comparing open with minimally invasive oesophagectomy in patients with oesophageal cancer. Any variant of minimally invasive oesophagectomy (total, hybrid or robotic) was eligible. Protocols of included RCTs and clinical trial registry database entries were retrieved where cited, and where protocols were not available, corresponding authors were contacted and asked to provide a copy. Reviews, conference abstracts, feasibility, pilot studies and those reporting only technical components of procedures were excluded.

Searches and screening

Three annual consecutive searches and a final update search of electronic databases (Ovid MEDLINE, Ovid Embase and CENTRAL) were conducted for RCTs published (i) January 2012–December 2017, (ii) January 2018–November 2018, (iii) December 2018–November 2020 and (iv) December 2020–June 2023. Comprehensive search strategies (online supplemental table 1) were tailored for each database using the concepts

oesophageal cancer, open surgery, minimally invasive surgery and RCTs. Retrieved articles were imported into Endnote (V.20) and deduplicated. One assessor (KAC) screened titles and abstracts. Full texts of remaining articles were retrieved and assessed for eligibility. If eligibility was unclear, a second reviewer (SEC) was approached, and if necessary, a third senior reviewer (JMB). Identified studies were presented and discussed with the ROMIO study management group, comprising experts in oesophageal cancer. This provided assurance that the searches and screening were comprehensive.

Data extraction

Data extraction focused on three main aspects: key trial characteristics and assessments of pragmaticism and validity. Where multiple publications and/or protocols related to the same RCT, data were extracted on a per trial, rather than per paper, basis. Data were extracted and assessments made by two researchers (KAC, SEC); a third reviewer (JMB) was consulted for technical advice or if disagreements occurred. Key trial characteristics included the type of surgical intervention (eg, totally robotic oesophagectomy), number of participating centre and patients, and study countries.

Assessment of pragmaticism

The Pragmatic-Explanatory Continuum Indicator Summary-2 (PRECIS-2) tool^{7 18} was used to assess characteristics of included RCTs that are important in determining applicability across nine domains (online supplemental table 2). Each domain was scored from 1 (very explanatory) to 5 (very pragmatic). Disagreements were discussed to facilitate determination of 'rules' to standardise judgments,¹⁹ in discussion with a third senior reviewer (JMB). Trials were then re-examined.

Aspects of PRECIS-2 that required specific attention because the interventions were surgical are explained here. 'Flexibility: delivery' addresses how much flexibility in the delivery of the intervention is allowed within the trial, for example, whether a strict prescribed protocol was used or whether delivery was at the surgeon's discretion. Judgments of procedural standardisation were informed by deconstructing interventions using a validated typology^{20 21} in consultation with a senior surgeon with procedural knowledge of both interventions (JMB). 'Flexibility: adherence' looks at whether the intervention was delivered as intended and typically relates to a patient's compliance with the intervention, such as taking medication. In surgical trials, patients are not involved in intervention delivery; PRECIS-2 authors therefore suggest this domain should be left blank.⁷ However, intervention adherence by surgeons may be examined by using methods to examine whether interventions were performed as planned (eg, by videorecording procedures). It is an important aspect of a surgical trial and merits individual acknowledgement.²² Therefore, for this review, 'flexibility: adherence' was assessed as more explanatory where trials had very detailed protocols

requiring strict surgeon adherence and more pragmatic in trials where few components were specified as required (or prohibited). In both types, videotaping can be used to examine adherence to intervention delivery protocols.

PRECIS wheels⁷ were created to illustrate assessments for each domain, which were examined in conjunction with overall mean scores to provide a descriptive summary of assessments. In line with PRECIS-2 guidelines, individual domains scoring four or five were considered pragmatic; three, equally explanatory and pragmatic; and two or one, explanatory. As there is no recognised standardised method for using the scores across domains to define trials as explanatory or pragmatic, in the current review, trials were considered to be pragmatic if they were conducted in multiple centres²³ and had a mean score of four or above. Descriptive accounts of all domains are provided for each trial.

Assessment of validity

The Risk of Bias (ROB) 2 tool²⁴ was used to assess validity. The judgments of five domains (online supplemental table 3) are considered collectively to produce an overall ROB (low risk, some concerns or high risk). Judgments were informed by the tool guidance¹² and inputted into an Excel-based macro with in-built algorithms used to compute overall ROB for each trial. In the first instance, RCTs were assessed independently by two researchers (KAC and SEC). Assessments were then discussed to facilitate determination of 'rules' to standardise judgments, and papers re-assessed. Disagreements were resolved by discussion or by JMB. Verbatim text from papers was extracted and used as a narrative to support judgments.

Assessment of inter-rater agreement

The extent of agreement between reviewers' assessments of external and internal validity was determined by calculating the number of assessments that matched, expressed as a percentage.

Patient and public involvement

Patients and public were not involved in the design or conduct of this systematic review.

Ethics approval

Ethics approval was not required for this review as there were no participants involved.

RESULTS

Full texts of 35 articles were reviewed and nine RCTs included^{25–38} (figure 1).

Key trial characteristics

Most RCTs (n=6) compared open and totally minimally invasive oesophagectomy (one robotically assisted) (table 1). Seven randomised <200 patients, including three with ≤100 patients. Seven were single centred. Five were conducted in China, one in each of the Netherlands, Austria and France and one was international (the

Netherlands, Spain and Italy). Four studies^{25 33 34 36} calculated and reported sample sizes—one study³⁴ was under powered as it did not recruit sufficient patients.

Pragmatism

Only one study provided information to enable PRECIS assessments in all nine domains.²⁶ No information was included for 'flexibility: adherence' in seven trials, 'primary analyses' in five trials and 'organisation' in four trials. Eligibility, setting and primary outcome were the only domains that could be populated for all (figure 2 and online supplemental table 4).

One study^{25 26} was considered to have adopted a more pragmatic approach as it was multicentred, recruited patients from outpatient clinics and performed an intention-to-treat analysis. However, additional data collection and standardisation and monitoring of interventions suggested that there were some aspects of trial design that were explanatory.

Most commonly, the recruitment domain was judged as pragmatic across studies.^{26 27 29 34 36–38} 'Primary analyses' was reported by four of the nine included studies, three^{26 33 36} used intention-to-treat analysis, ensuring that patients were analysed within the group they were randomised, and one³⁴ used per protocol analysis. The domains commonly judged to be more explanatory were 'trial setting' (most (n=7)^{27–29 34 36–38} were conducted in single centres), 'flexibility: delivery' (four^{26 33 34 38} reported methods to standardise the delivery of interventions) and 'flexibility: adherence' (two^{26 33} reported strategies to monitor adherence to intervention delivery).

Validity

Overall ROB was judged to be low in two studies.^{26 36} Of the 45 individual judgments made over the five domains, 23 were low ROB. 'Missing outcome data' was most frequently judged to be at low ROB.^{26 28 29 33 34 36–38} 'Measurement of outcome' was judged as low ROB in five studies; these used methods deemed to be appropriate to assess primary outcomes unlikely to be affected by lack of outcome assessor blinding.^{27 28 37} In six studies, there was insufficient information to make a judgment about 'deviations from intended interventions' as no information was reported about blinding of patients, carers or people delivering the intervention, crossovers and whether an intention-to-treat analysis was used. Only three trials^{26 34 36} comprehensively reported methods of randomisation and allocation (figure 3 and online supplemental table 5).

Assessment of inter-rater agreement

After establishing rules to standardise judgments, agreement for PRECIS-2 assessments was 91.7% (mean, SD 9.8). For validity, agreement after initial judgments was 65% (mean, SD 20.7), with the greatest agreement for 'missing outcome data' (100%) and 'randomisation process' (75%) domains and least for the 'deviations from intended interventions' (37.5%) domain. Following discussions with a ROB expert (JS), consensus

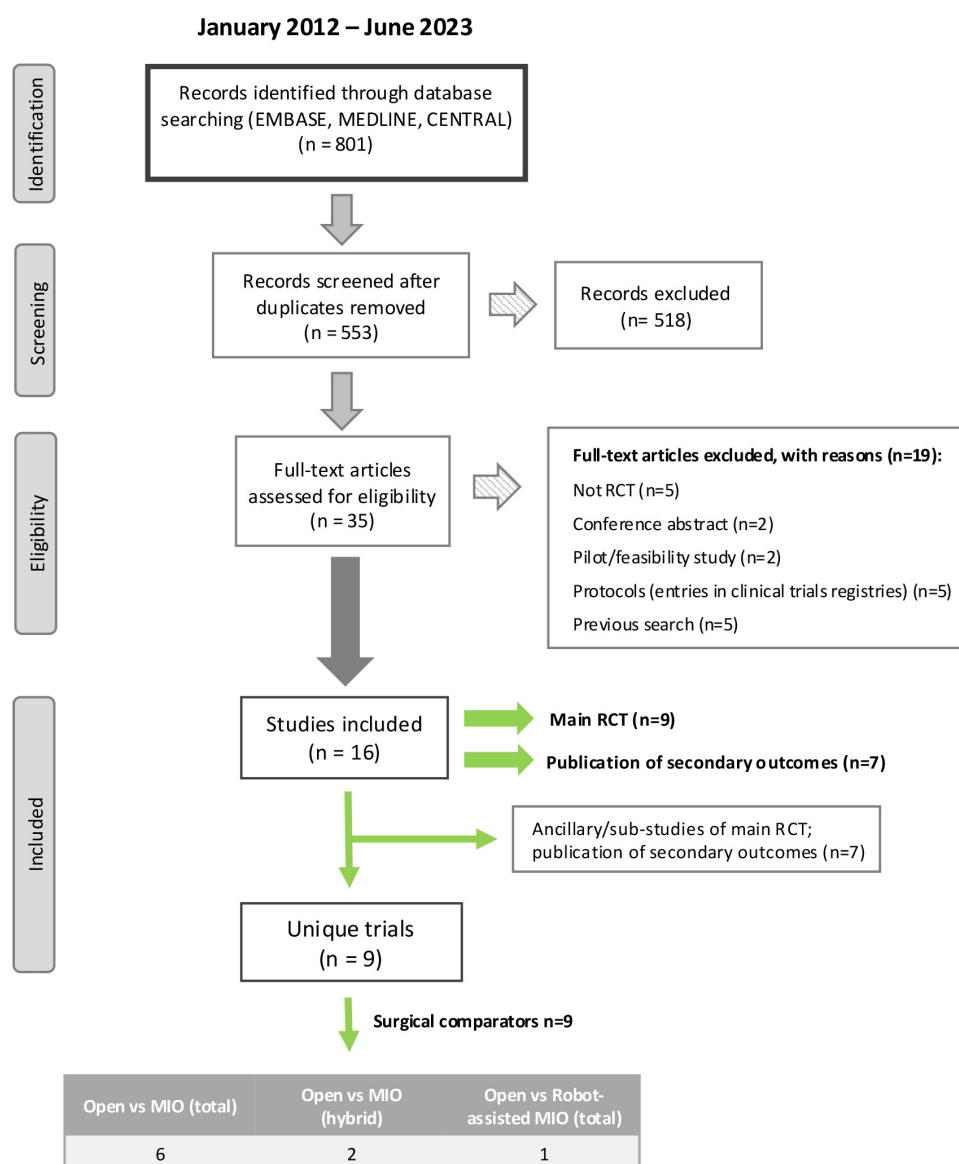


Figure 1 PRISMA flowchart depicting the search strategy and selection of articles for the review. MIO, minimally invasive oesophagectomy; PRISMA, Preferred Reporting Items for Systematic reviews and Meta-Analyses. RCT, randomised controlled trial.

was consequently reached for all judgments across all domains after re-assessment.

DISCUSSION

This review provides an in-depth examination of surgical RCTs comparing two widely used procedures for oesophageal cancer (open and minimally invasive oesophagectomy) and whether outcomes from these trials are applicable to clinical practice. Until recently, applicability (also referred to as generalisability) has received much less attention than validity. Applicability was assessed using the PRECIS-2 tool.⁷ In the nine RCTs identified, only one reported information required to assess all nine PRECIS-2 domains. One study was classified as more pragmatic and therefore applicable to the wider clinical setting. The others had characteristics deemed more

explanatory, including inclusion of only single centres, with narrow eligibility criteria, prescribed intervention delivery methods or an intense follow-up regime, characteristics not exhibited in ‘real-world’ clinical practice. There is a need for increased awareness of the importance of applicability within the surgical trials’ community if trial results are to be applied to a wider clinical setting. Higher quality pragmatic, multicentre surgical RCTs are required to generate evidence to inform clinical practice.

Defining trials as ‘explanatory’ or ‘pragmatic’ based on their intended purpose were first reported in 1967.³⁹ A trial to determine the effects of interventions delivered in conditions more aligned with usual care was described as being pragmatic,³⁹ and there may be more confidence that trial results are applicable to the wider target patient population. Certain trial characteristics are key

Table 1 Detailed characteristics of included studies

Study number	First author (publication year)	Countries	Number of centres	Number of patients randomised		Type of MIO	Primary outcome* (timing of outcome)
				Open	MIO		
1a	Biere (2012)	Netherlands /Spain/Italy	5	56	59	Total	Pulmonary infection (2 weeks)
1b†	Maas (2014)	Netherlands	1	13	14		Stress response measurements (1 day)
1c†	Maas (2015)	Netherlands	5	56	59		HRQL (1 year)
1d†	Biere (2017)	/Spain/Italy		56	59		Factors associated with respiratory infections (not stated)
1e†	Straatman (2017)			56	59		Overall survival (3 years)
2	Guo (2013)	China	1	110	111	Total	Length of operation (not applicable)
3	Hong (2013)	China	1	59	55	Total	Pulmonary infection (2 weeks)
4	Paireder (2018)	Austria	1	13	16	Hybrid laparoscopy and thoracotomy	Morbidity and mortality (4 weeks)
5	Ma (2018)	China	1	97	47	Total	Morbidity (not stated)
6	van der Sluis (2019)	Netherlands	1	56	56	Total robot-assisted	Morbidity (2 weeks)
7a	Mariette (2019)	France	13	104	103	Hybrid laparoscopy and thoracotomy	Morbidity (4 weeks)
7b‡	Mariette (2019)			104	103		HRQL (3 years)
8	Zhang (2020)	China	1	50	50	Total	Length of operation (not applicable)
9	Yu (2022)	China	1	45	45	Total	Length of operation (not applicable)

*If primary outcome was not specified, the first outcome listed in the abstract is given. For ancillary studies and substudies, this will differ from the original trial and will be the main outcome reported.
†Ancillary and sub-studies of Biere *et al.*²⁶
‡Publication of secondary outcomes.
HRQL, health-related quality of life; MIO, minimally invasive oesophagectomy.

in determining the degree of pragmatism. For example, study populations are determined by eligibility criteria; a pragmatic study would aim to include patients who would receive the intervention in usual care. Four of the RCTs examined excluded patients with tumours of the upper oesophagus,^{26 33 34 36} although these patients would have received the intervention in usual clinical practice.^{33 34} The number of study centres is also important; outcomes from one centre may not be comparable to another that differs in geography, culture or surgeon/centre experience. Only two studies were conducted in multiple centres.^{26 33} Dal-Ré *et al* suggested that single-centre studies should be considered ‘explanatory’ regardless of other domain scores,²³ although they may be useful in decision-making where that the single centre is representative of the wider context.¹⁸ Treatment effects have been found to be larger in single centre studies,^{40 41} and so results should be interpreted with caution. It is only in multicentre trials that the intervention may be evaluated in different contexts with multiple surgeons.^{37 38} The importance of selecting centres that reflect those where

procedures are delivered in usual care is highlighted by a comparison of data from the TIME trial²⁶ (a multicentred European trial) with data from the Dutch Upper Gastro-intestinal Cancer Audit, following the implementation of totally minimally invasive oesophagectomy in the Netherlands.⁴² Audit data showed increased total and pulmonary complications and reoperations in patients undergoing the minimally invasive intervention⁴²—adverse events not seen in the TIME trial.²⁶ It was concluded that this may reflect non-expert surgeons outside of high-volume centres performing this minimally invasive technique in a non-standardised fashion outside of a controlled trial environment.⁴² Assessment of applicability of the TIME trial using the PRECIS-2 tool in this review recognised that the experience required for both the centres and individual surgeons to complete the MIO was greater than for open oesophagectomy. However, Markar’s study⁴² has illustrated that perhaps a greater emphasis needs to be placed on the potential learning curve for surgeons and the experience of a non-specialist centre. This is a crucial point; for trial outcomes to be applicable,

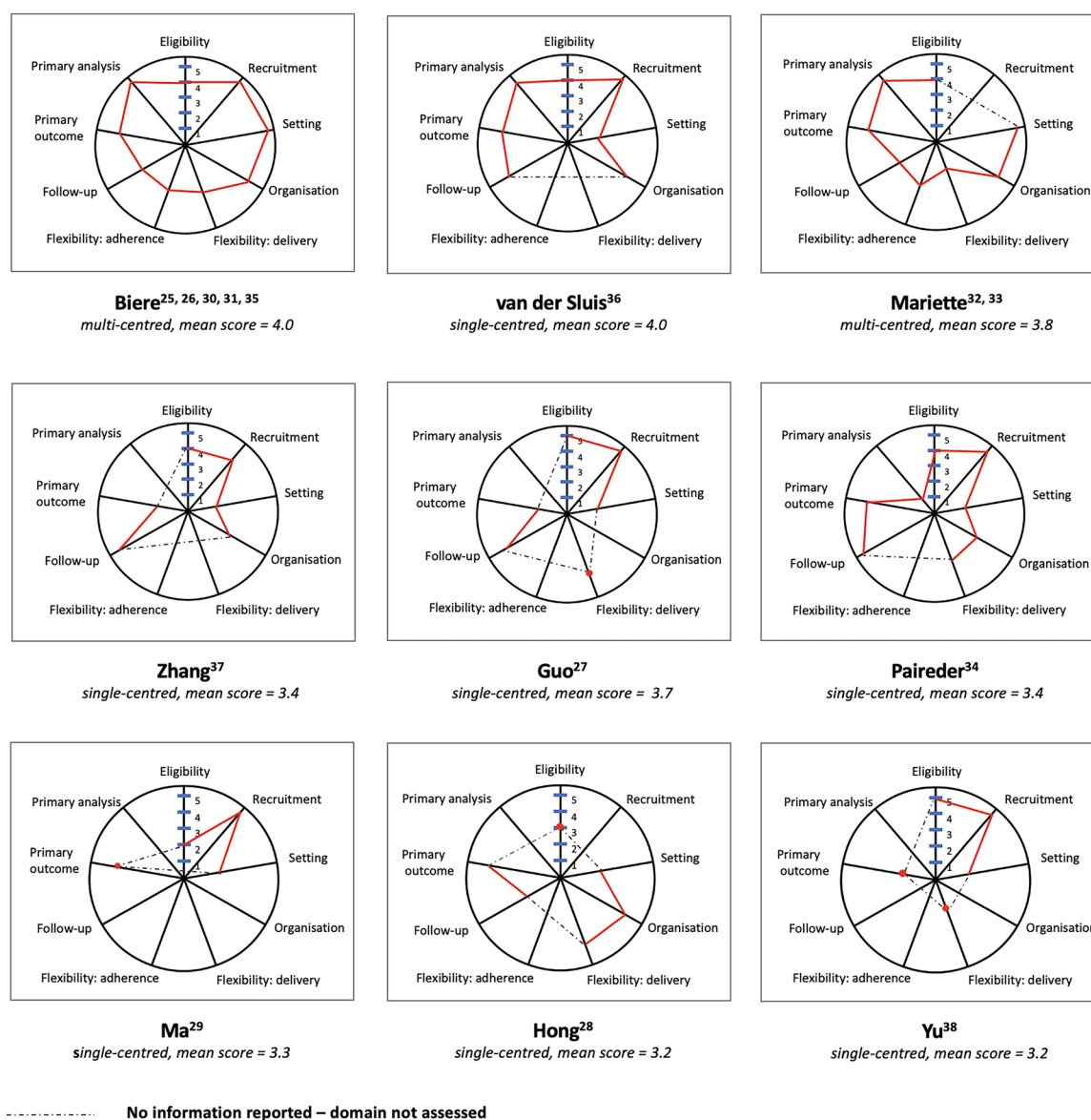


Figure 2 Pragmatic-Explanatory Continuum Indicator Summary-2 (PRECIS-2) scores for each domain plotted on PRECIS wheels, illustrating the pragmatism of included studies.

surgeons and centres conducting the interventions in RCTs need to reflect surgeons and centres nationally and perhaps internationally.

The majority of the RCTs examined were poorly reported. Tools such as the Standard Protocol Items: Recommendations for Interventional Trials checklist⁴³ have been developed to encourage trialists to address (and report) specific items in interventional trials, including rationale, objectives, study setting, PICO (population, intervention, comparator, outcome), sample size and recruitment strategies. The provision of this information would enable a more informed, comprehensive assessment of a trial's applicability. The Consolidated Standards of Reporting Trials checklist⁴⁴ and subsequent extensions recommend that the generalisability of a trial's findings should be discussed⁴⁵ and suggest greater depth of reporting throughout all sections of a paper

to facilitate judgements about applicability.⁴⁶ Cited over 1100 times, the checklist is evidently recognised, yet poor reporting of key methodological aspects was evident in the studies examined. This is reflected in other reviews in surgery⁴⁷ and medicine.⁴⁸

Some trial design features, such as intervention standardisation and monitoring adherence, are judged to be explanatory by PRECIS-2; however, in surgical trials, their absence may compromise internal validity. For example, some standardisation of surgical procedures may be required, so interventions can be compared across surgeons/centres and to monitor their delivery. In terms of analysis, data from complex surgical trials should also be carefully considered when applying PRECIS-2. Strict intention-to-treat (ITT) analyses state that all patients should be analysed in their assigned group regardless of whether they receive an intervention, crossover or are lost

Author (primary outcome)	Domain					
	Randomisation process	Deviations from intended intervention	Missing outcome data	Measurement of the outcome	Selection of the reported result	Overall bias
Biere ^{25, 26, 30, 31, 35} (Pulmonary infection)	Low risk of bias	Low risk of bias	Low risk of bias	Low risk of bias	Low risk of bias	Low risk of bias
Guo ²⁷ (Length of operation)	Some concerns	High risk of bias	Some concerns	High risk of bias	Some concerns	High risk of bias
Hong ²⁸ (Pulmonary infection)	Some concerns	Some concerns	Low risk of bias	Some concerns	Some concerns	Some concerns
Paireder ³⁴ (Measure of morbidity)	Low risk of bias	Some concerns	Low risk of bias	Low risk of bias	Some concerns	Some concerns
Ma ²⁹ (Measure of morbidity)	Some concerns	Some concerns	Low risk of bias	Low risk of bias	Some concerns	Some concerns
van der Sluis ³⁶ (Measure of morbidity)	Low risk of bias	Low risk of bias	Low risk of bias	Low risk of bias	Low risk of bias	Low risk of bias
Mariette ^{32, 33} (Measure of morbidity)	Some concerns	Low risk of bias	Low risk of bias	Low risk of bias	Low risk of bias	Some concerns
Zhang ³⁷ (Length of operation)	Some concerns	High risk of bias	Low risk of bias	High risk of bias	Some concerns	High risk of bias
Yu ³⁸ (Length of operation)	Some concerns	Some concerns	Low risk of bias	High risk of bias	Some concerns	High risk of bias

Low risk of bias Some concerns High risk of bias

Figure 3 ROB assessment for the trial's primary outcome, using the ROB 2 tool. ROB, risk of bias.

to follow-up. Often in surgical trials, patients are found to no longer be eligible when they undergo surgery for reasons such as advanced disease or the presence of previously unknown pathologies, both which result in patients not receiving the allocated intervention. The undertaking of preplanned sensitivity analyses, such as modified ITT, may be seen as less pragmatic by PRECIS-2; however, in real terms, the provision of these analyses provides an important perspective about the results and their interpretation. The use of per-protocol analyses in surgical trials is therefore important and often essential in non-inferiority designs. The definitions of a per-protocol analyses in a surgical trial, however, are often confusing and trialists need to carefully describe this in relevant documents to optimise transparency. Comprehensive and transparent reporting will facilitate more informed considerations about applicability to be made.

Strengths of the review include rigorous and comprehensive assessment of applicability and validity by two reviewers independently, including sourcing unpublished protocols. Detailed text provides transparency about how the ROB-2 and PRECIS-2 tools were applied. Furthermore, this review addresses a gap in the literature in this area, and it modifies the PRECIS-2 methods to make them more applicable to surgical interventions by using a published typology. A potential limitation to the

review is that due to the nature of the annual literature reviews, dual screening was not conducted. However, the narrow search criteria meant that the maximum number of entries to screen was <250, limiting this possibility, and findings were presented to experts involved in the recently completed ROMIO study¹⁷ to ensure comprehensiveness. A further point of note is that the PRECIS-2 tool was developed to aid trialists in the design stage rather than the assessment of completed studies. However, it is being used retrospectively,²³ with some adjustments.¹⁸ This, and its ability to facilitate a comprehensive and detailed assessment of trial design features important in applicability, meant that the current review used the PRECIS-2 tool. However, it should be noted that there is no recognised standardised method for using PRECIS-2 scores to define trials as explanatory or pragmatic. Readers are encouraged to look at individual domain scores depicted in the PRECIS wheels, to gain a more in-depth understanding of each trial's design.

Since completing this work, several publications have used,^{49 50} or propose to use,^{51 52} PRECIS-2 to determine the applicability of eligible studies in systematic reviews. Each has selected a different value to define a trial as pragmatic. Applying the criteria used by other authors to the eight trials identified in this review would have led to different conclusions. One method⁵⁰ would have

identified three studies^{26 33 36} as pragmatic, and the other three methods^{49 51 52} would have considered all studies as pragmatic.²⁶ This is because a more lenient approach to study parameter classification²⁶ was used in these reviews than in this review. It is recommended that standardised methods are established to interpret retrospective assessments to enable comparisons between studies. It is also recommended that new trials are clear about their intentions in design (ie, whether explanatory or more pragmatic) because of the lack of clarity seen in the included trials. This could be part of the trial protocol and accompanying publication. Until this is achieved, transparent reporting of all judgments and data, as reported in this review, are vital. In the area of oesophageal cancer surgery where robotic techniques are becoming more widespread, it is recommended that pragmatic trials are conducted maybe comparing robotic with minimal access techniques. A potential method to optimise investment would be to nest an IDEAL phase 2b more explanatory study within a larger pragmatic trial. The Ideal 2b study would only run in a few centres.⁵³

Using RCTs of oesophagectomy for oesophageal cancer as a case study, this review has demonstrated that although trials contained both pragmatic and explanatory characteristics, only one was classified as pragmatic. Stakeholders involved in surgical trials need to be aware of the importance of applicability. Further work in this area should raise awareness of existing guidelines^{43–46 54} and tools^{7 24} developed to improve the design, conduct and reporting of trials. It is also recommended that current methods are applied to reviews of RCTs of other surgical interventions to examine whether their findings are applicable. Trial-lists should be encouraged to incorporate considerations of applicability into the design stages of trials to ensure they are fit for the purpose. While well-designed, powered explanatory studies are possible and appropriate where the objective is to establish mechanisms and how it works in expert hands, A pragmatic approach is required in order to change clinical practice.

X Natalie S Blencowe @NatalieBlencowe

Acknowledgements We thank Dr. Paireder for his correspondence and providing additional information about the MIOMIE trial. We also thank Dr. Jelena Savović for her assistance with the Risk of Bias 2 tool.

Contributors KAC: methodology, validation, formal analysis, investigation, data curation, writing—original draft, visualisation, project administration. SEC: methodology, validation, investigation, writing—original draft, project administration. NSB: methodology, writing—review and editing. JMB: conceptualisation, methodology, writing—original draft, supervision, funding acquisition. All authors reviewed and approved the final manuscript. KC is the guarantor.

Funding This work was supported by the MRC ConDuCT-II (Collaboration and innovation for Difficult and Complex Randomized Controlled Trials in Invasive Procedures) Hub for Trials Methodology Research (MR/K025643/1) and the NIHR Biomedical Research Centre at the University Hospitals Bristol NHS Foundation Trust and the University of Bristol (BRC-1215-20011). The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health and Care Research, or the Department of Health. JMB is an NIHR Senior Investigator, and NSB is an MRC Clinician Scientist.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

ORCID iDs

Katy Chalmers <http://orcid.org/0000-0003-4923-3000>

Sian Cousins <http://orcid.org/0000-0003-0088-841X>

Natalie S Blencowe <http://orcid.org/0000-0002-6111-2175>

Jane Blazeby <http://orcid.org/0000-0002-3354-3330>

REFERENCES

- 1 Barkun JS, Aronson JK, Feldman LS, *et al*. Evaluation and stages of surgical innovations. *Lancet* 2009;374:1089–96.
- 2 Dimick JB, Sedrakyan A, McCulloch P. The IDEAL Framework for Evaluating Surgical Innovation: How It Can Be Used to Improve the Quality of Evidence. *JAMA Surg* 2019;154:685–6.
- 3 Ergina PL, Cook JA, Blazeby JM, *et al*. Challenges in evaluating surgical innovation. *Lancet* 2009;374:1097–104.
- 4 Hirst A, Philippou Y, Blazeby J, *et al*. No Surgical Innovation Without Evaluation: Evolution and Further Development of the IDEAL Framework and Recommendations. *Ann Surg* 2019;269:211–20.
- 5 McCulloch P, Altman DG, Campbell WB, *et al*. No surgical innovation without evaluation: the IDEAL recommendations. *Lancet* 2009;374:1105–12.
- 6 Higgins JPT, Altman DG, Gøtzsche PC, *et al*. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928.
- 7 Loudon K, Treweek S, Sullivan F, *et al*. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ* 2015;350:h2147.
- 8 Thorpe KE, Zwarenstein M, Oxman AD, *et al*. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *CMAJ* 2009;180:E47–57.
- 9 Akhtar NM, Chen D, Zhao Y, *et al*. Postoperative short-term outcomes of minimally invasive versus open esophagectomy for patients with esophageal cancer: An updated systematic review and meta-analysis. *Thorac Cancer* 2020;11:1465–75.
- 10 Booka E, Tsubosa Y, Haneda R, *et al*. Ability of Laparoscopic Gastric Mobilization to Prevent Pulmonary Complications After Open Thoracotomy or Thoracoscopic Esophagectomy for Esophageal Cancer: A Systematic Review and Meta-analysis. *World J Surg* 2020;44:980–9.
- 11 Guo W, Ma X, Yang S, *et al*. Combined thoracoscopic-laparoscopic esophagectomy versus open esophagectomy: a meta-analysis of outcomes. *Surg Endosc* 2016;30:3873–81.
- 12 Kauppila JH, Xie S, Johar A, *et al*. Meta-analysis of health-related quality of life after minimally invasive versus open oesophagectomy for oesophageal cancer. *Br J Surg* 2017;104:1131–40.
- 13 Oor JE, Wiezer MJ, Hazebroek EJ. Hiatal Hernia After Open versus Minimally Invasive Esophagectomy: A Systematic Review and Meta-analysis. *Ann Surg Oncol* 2016;23:2690–8.
- 14 Pu S, Chen H, Zhou C, *et al*. Major Postoperative Complications in Esophageal Cancer After Minimally Invasive Esophagectomy Compared With Open Esophagectomy: An Updated Meta-analysis. *J Surg Res* 2021;257:554–71.

- 15 Xiong WL, Li R, Lei HK, *et al.* Comparison of outcomes between minimally invasive oesophagectomy and open oesophagectomy for oesophageal cancer. *ANZ J Surg* 2017;87:165–70.
- 16 Zhou C, Zhang L, Wang H, *et al.* Superiority of Minimally Invasive Oesophagectomy in Reducing In-Hospital Mortality of Patients with Resectable Oesophageal Cancer: A Meta-Analysis. *PLoS ONE* 2015;10:e0132889.
- 17 Brierley RC, Gaunt D, Metcalfe C, *et al.* Laparoscopically assisted versus open oesophagectomy for patients with oesophageal cancer—the Randomised Oesophagectomy: Minimally Invasive or Open (ROMIO) study: protocol for a randomised controlled trial (RCT). *BMJ Open* 2019;9:e030907.
- 18 Zwarenstein M, Thorpe K, Treweek S, *et al.* PRECIS-2 for retrospective assessment of RCTs in systematic reviews. *J Clin Epidemiol* 2020;126:202–6.
- 19 Forbes G, Loudon K, Treweek S, *et al.* Understanding the applicability of results from primary care trials: lessons learned from applying PRECIS-2. *J Clin Epidemiol* 2017;90:119–26.
- 20 Blencowe NS, Mills N, Cook JA, *et al.* Standardizing and monitoring the delivery of surgical interventions in randomized clinical trials. *Br J Surg* 2016;103:1377–84.
- 21 Cousins S, Blencowe NS, Tsang C, *et al.* Optimizing the design of invasive placebo interventions in randomized controlled trials. *Br J Surg* 2020;107:1114–22.
- 22 Blencowe NS, Skilton A, Gaunt D, *et al.* Protocol for developing quality assurance measures to use in surgical trials: an example from the ROMIO study. *BMJ Open* 2019;9:e026209.
- 23 Dal-Ré R, de Boer A, James SK. The design can limit PRECIS-2 retrospective assessment of the clinical trial explanatory/pragmatic features. *J Clin Epidemiol* 2020;126:193–201.
- 24 Sterne JAC, Savović J, Page MJ, *et al.* RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:14898.
- 25 Biere SS, van Berge Henegouwen MI, Bonavina L, *et al.* Predictive factors for post-operative respiratory infections after esophagectomy for esophageal cancer: outcome of randomized trial. *J Thorac Dis* 2017;9:S861–7.
- 26 Biere SSAY, van Berge Henegouwen MI, Maas KW, *et al.* Minimally invasive versus open oesophagectomy for patients with oesophageal cancer: a multicentre, open-label, randomised controlled trial. *Lancet* 2012;379:1887–92.
- 27 Guo M, Xie B, Sun X, *et al.* A comparative study of the therapeutic effect in two protocols: video-assisted thoracic surgery combined with laparoscopy versus right open transthoracic esophagectomy for esophageal cancer management. *Chin -Ger J Clin Oncol* 2013;12:68–71.
- 28 Hong L, Zhang Y, Zhang H, *et al.* The short-term outcome of three-field minimally invasive esophagectomy for Siewert type I esophagogastric junctional adenocarcinoma. *Ann Thorac Surg* 2013;96:1826–31.
- 29 Ma G, Cao H, Wei R, *et al.* Comparison of the short-term clinical outcome between open and minimally invasive esophagectomy by comprehensive complication index. *J Cancer Res Ther* 2018;14:789–94.
- 30 Maas KW, Biere SSAY, van Hoogstraten IMW, *et al.* Immunological changes after minimally invasive or conventional esophageal resection for cancer: a randomized trial. *World J Surg* 2014;38:131–7.
- 31 Maas KW, Cuesta MA, van Berge Henegouwen MI, *et al.* Quality of Life and Late Complications After Minimally Invasive Compared to Open Esophagectomy: Results of a Randomized Trial. *World J Surg* 2015;39:1986–93.
- 32 Mariette C, Markar SR, Dabakuyo-Yonli TS, *et al.* Health-related Quality of Life Following Hybrid Minimally Invasive Versus Open Esophagectomy for Patients With Esophageal Cancer, Analysis of a Multicenter, Open-label, Randomized Phase III Controlled Trial. *Ann Surg* 2020;271:1023–9.
- 33 Mariette C, Markar SR, Dabakuyo-Yonli TS, *et al.* Hybrid Minimally Invasive Esophagectomy for Esophageal Cancer. *N Engl J Med* 2019;380:152–62.
- 34 Paireder M, Asari R, Kristo I, *et al.* Morbidity in open versus minimally invasive hybrid esophagectomy (MIOMIE): Long-term results of a randomized controlled clinical study. *Eur Surg* 2018;50:249–55.
- 35 Straatman J, van der Wielen N, Cuesta MA, *et al.* Minimally Invasive Versus Open Esophageal Resection: Three-year Follow-up of the Previously Reported Randomized Controlled Trial: the TIME Trial. *Ann Surg* 2017;266:232–6.
- 36 van der Sluis PC, van der Horst S, May AM, *et al.* Robot-assisted Minimally Invasive Thoracoscopic Esophagectomy Versus Open Transthoracic Esophagectomy for Resectable Esophageal Cancer: A Randomized Controlled Trial. *Ann Surg* 2019;269:621–30.
- 37 Zhang T, Jiang H, Ming CY, *et al.* Influence of different kinds of surgical resection on operation-related clinical indexes, inflammatory cytokines and complications in elderly patients with esophageal cancer. *Pak J Med Sci* 2020;36:532–7.
- 38 Yu Y, Han Y. Clinical Effect and Postoperative Pain of Laparo-Thoracoscopic Esophagectomy in Patients with Esophageal Cancer. *Evid Based Complement Alternat Med* 2022;2022:4507696.
- 39 Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutic trials. *J Chronic Dis* 1967;20:637–48.
- 40 Bafeta A, Dechartres A, Trinquart L, *et al.* Impact of single centre status on estimates of intervention effects in trials with continuous outcomes: meta-epidemiological study. *BMJ* 2012;344:e813.
- 41 Unverzagt S, Prondzinsky R, Peinemann F. Single-center trials tend to provide larger treatment effects than multicenter trials: a systematic review. *J Clin Epidemiol* 2013;66:1271–80.
- 42 Markar SR, Ni M, Gisbertz SS, *et al.* Implementation of Minimally Invasive Esophagectomy From a Randomized Controlled Trial Setting to National Practice. *JCO* 2020;38:2130–9.
- 43 Chan A-W, Tetzlaff JM, Gøtzsche PC, *et al.* SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ* 2013;346:e7586.
- 44 Begg C, Cho M, Eastwood S, *et al.* Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996;276:637–9.
- 45 Boutron I, Moher D, Altman DG, *et al.* Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med* 2008;148:295–309.
- 46 Zwarenstein M, Treweek S, Gagnier JJ, *et al.* For the CONSORT and Pragmatic Trials in Healthcare (Practihc) groups. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 2008;337:a2390.
- 47 Blencowe NS, Boddy AP, Harris A, *et al.* Systematic review of intervention design and delivery in pragmatic and explanatory randomised clinical trials. *Br J Surg* 2015;102:1037–47.
- 48 Dal-Ré R. Articles provided insufficient information to conduct an appropriate retrospective assessment of the pragmatic/explanatory features of medicine trials with the PRECIS-2 tool. *Eur J Clin Pharmacol* 2020;76:1093–102.
- 49 Fitzpatrick C, Gillies C, Seidu S, *et al.* Effect of pragmatic versus explanatory interventions on medication adherence in people with cardiometabolic conditions: a systematic review and meta-analysis. *BMJ Open* 2020;10:e036575.
- 50 Sharma T, Qamar I, Zwarenstein M. How pragmatic are randomized trials of remdesivir and favipiravir for in-hospital treatment of COVID-19: a descriptive methodological review of trial design using the PRECIS-2 framework. *J Clin Epidemiol* 2022;152:193–200.
- 51 Coumoundouros C, von Essen L, Sanderman R, *et al.* Implementation of e-mental health interventions for informal caregivers of adults with chronic diseases: a protocol for a mixed-methods systematic review with a qualitative comparative analysis. *BMJ Open* 2020;10:e035406.
- 52 Lim J, Lee H, Kim Y-S. Applying the PRECIS-2 tool for self-declared “pragmatic” acupuncture trials: protocol for a systematic review. *BMJ Open* 2022;12:e052861.
- 53 Metcalfe C, Avery K, Berrisford R, *et al.* Comparing open and minimally invasive surgical procedures for oesophagectomy in the treatment of cancer: the ROMIO (Randomised Oesophagectomy: Minimally Invasive or Open) feasibility study and pilot trial. *Health Technol Assess* 2016;20:1–68.
- 54 Hoffmann TC, Glasziou PP, Boutron I, *et al.* Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;348:bmj.g1687.