



HAL
open science

Reporting of cluster randomised crossover trials 2023: extension of the CONSORT 2010 statement with explanation and elaboration Name / Job title Affiliation

Joanne Mckenzie, Monica Taljaard, Karla Hemming, Sarah Arnup, Bruno Giraudeau, Sandra Eldridge, Richard Hooper, Brennan Kahan, Tianjing Li, David Moher, et al.

► To cite this version:

Joanne Mckenzie, Monica Taljaard, Karla Hemming, Sarah Arnup, Bruno Giraudeau, et al.. Reporting of cluster randomised crossover trials 2023: extension of the CONSORT 2010 statement with explanation and elaboration Name / Job title Affiliation. *BMJ - British Medical Journal*, 2025, pp.e080472. <10.1136/bmj-2024-080472>. <hal-05075241>

HAL Id: hal-05075241

<https://hal.science/hal-05075241v1>

Submitted on 20 May 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Reporting of cluster randomised crossover trials 2023: extension of the CONSORT 2010 statement with explanation and elaboration

Name / Job title	Affiliation
Joanne E McKenzie Professor	School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia ORCID 0000-0003-3534-1641 Twitter/X @jomck15
Monica Taljaard Professor	Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada ORCID 0000-0002-3978-8961
Karla Hemming Professor	Institute of Applied Health Research, University of Birmingham, Birmingham B15 2TT, UK ORCID Twitter/X @karlahemming
Sarah J Arnup Research Fellow	School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia ORCID 0000-0002-9497-1902
Bruno Giraudeau Professor	Université de Tours, Université de Nantes, INSERM SPHERE U1246, Tours, France. INSERM CIC, CHRU de Tours, Tours, France ORCID 0000-0003-3031-8258
Sandra Eldridge Professor	Wolfson Institute of Population Health, Queen Mary University of London, London, UK
Richard Hooper Professor	Wolfson Institute of Population Health, Queen Mary University of London, London, UK ORCID 0000-0002-1063-0917
Brennan C Kahan Principal Research Fellow	MRC Clinical Trials Unit at UCL, London, UK ORCID 0000-0001-9957-0844 Twitter/X @Brennan_Kahan
Tianjing Li Associate Professor	Department of Ophthalmology, School of Medicine, University of Colorado Anschutz Medical Campus, Aurora, Colorado, United States of America ORCID 0000-0001-5371-4558
David Moher Professor	Centre for Journalology, Clinical Epidemiology Program, Ottawa Hospital Research Institute, and School of Epidemiology and Public Health, University of Ottawa, Ottawa, Canada ORCID 0000-0003-2434-4206 Twitter/X @dmoher
Elizabeth L Turner Associate Professor	Department of Biostatistics and Bioinformatics, and Duke Global Health Institute, Duke University, Durham, NC, USA ORCID 0000-0002-7638-5942
Jeremy M Grimshaw Professor	Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada

	Department of Medicine, University of Ottawa, Ottawa, ON, Canada ORCID 0000-0001-8015-8243 Twitter/X @GrimshawJeremy
Andrew B Forbes Professor	School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia ORCID 0000-0003-4269-914X Twitter/X @AndrewF02554942

Corresponding Author:

Professor Joanne E McKenzie, Methods in Evidence Synthesis Unit, School of Public Health and Preventive Medicine, Monash University, 553 St Kilda Road, Melbourne Victoria 3004, Australia. Telephone: +61 3 9903 0380. Email address: joanne.mckenzie@monash.edu

Word count: excluding examples 10746, including examples 14833

Standfirst

In this report we present the Consolidated Standards of Reporting Trials (CONSORT) extension for the cluster randomised crossover (CRXO) trial. A CRXO trial involves randomisation of groups of individuals ('clusters') to different sequences of interventions over time. The design has gained popularity in settings where cluster randomisation is required as it can largely overcome the loss in power due to clustering in parallel cluster trials. However, the design has many methodological complexities, requiring tailored reporting guidance. The guideline was developed using a survey and in-person consensus meeting, informed by a systematic review examining the quality of reporting in CRXO trials and relevant CONSORT statements for individual, crossover, cluster and stepped-wedge designs. We provide recommended reporting items, along with explanations and examples.

Summary points

- The cluster randomised crossover (CRXO) trial has been gaining popularity over time due to the design's methodological strengths as compared with other cluster randomised designs (i.e. increased statistical efficiency, reduced risk of chance imbalances in cluster characteristics).
- The design has allowed randomised evaluation of interventions directed at clusters which otherwise could not have been undertaken due to the availability of too few clusters to achieve the desired power (e.g. when there are a limited number of intensive care units within a country).
- There are many methodological complexities in the design, conduct and analysis of the CRXO trials that requires tailored reporting guidance.
- We report a CONSORT extension for the CRXO trial design. The reporting guideline brings together advice from relevant CONSORT statements and addresses issues unique to the CRXO design. Coupled with this, new items, and amendments to others, reflect evolving requirements of the scientific community.
- The CONSORT CRXO extension includes a checklist of 28 items (with a total of 43 sub-items), along with a separate checklist for journal and conference abstracts, and sample size items.

Introduction

Complete and accurate reporting of randomised trials is necessary to enable users to understand the benefits and harms of an intervention, details of the intervention (such that it could be implemented in practice), the population to which the findings apply, and any methodological limitations that may threaten the validity of the findings. Additionally, complete reporting allows the findings of a trial to be incorporated in evidence synthesis products such as systematic reviews and clinical practice guidelines. Complete and accurate reporting also facilitates reproducibility (1). The Consolidated Standards of Reporting Trials (CONSORT) statement, first published in 1996 (2) and last updated in 2010 (3), aims to improve the quality of reporting of parallel group individually randomised designs through the provision of a minimum set of recommendations for reporting (hereafter referred to as 'CONSORT 2010'). Extensions have been developed, providing modified and additional recommendations for variants of randomised trial designs (e.g. crossover randomised trials (4), cluster randomised trials (5)), specific interventions (e.g. social and psychological interventions (6)) and outcome types (e.g. harms (7)).

The Cluster Randomised Crossover (CRXO) trial is a particular type of randomised trial design that has become more prominent in recent decades (8). In this design, groups of individuals (known as 'clusters') are randomly assigned to different sequences of treatments over time (see Box 1 for an example). The design has several potential benefits. The cluster randomisation feature allows evaluation of interventions directed at clusters, can provide logistical benefits and can reduce the potential for within-cluster contamination (9). The crossover feature can yield increased statistical efficiency and reduce the impact of chance imbalances in cluster characteristics (10).

While the CRXO design offers many advantages, there are methodological complexities in the design, conduct and analysis that require tailored reporting guidance. Current guidelines for the individual parallel group (3) and crossover designs (4), cluster parallel group (5) and stepped wedge designs (11) provide reporting recommendations of relevance to the CRXO design, but these are scattered across the statements and do not address all issues unique to the CRXO design. Furthermore, according to a systematic review examining the quality of reporting of CRXO trials, incomplete reporting was common for aspects of the methods and results (8, 12). Thus, our aim was to develop a CONSORT reporting extension for the CRXO design that consolidated reporting recommendations across relevant statements, modified items as necessary and added relevant items to address identified reporting gaps. Here we report the consensus process used to develop the extension, and the resulting items with explanation and elaboration, along with examples of complete reporting.

Design features of a CRXO trial

In a CRXO design, clusters are randomly allocated to sequences of treatments which are delivered in different periods of time (see Table 1 for a glossary of terms). Typically, the sequences alternate between intervention and control conditions, such that each cluster will cross between all treatment conditions at least once. Some CRXO trials include a washout period in which no treatment is delivered, as a means to mitigate any potential individual-level or cluster-level carryover effects.

Figure 1(A) depicts the commonly employed two-treatment (A or B), two-period and two-sequence (AB or BA) cross-sectional design, where observations in different periods in a cluster are of different people (8). Figure 1(B) depicts a cohort design, where observations in different periods in a cluster are of the same people. 'Mixture' designs are also possible in which observations are from a mix of

the same and different people in different periods. Figure 1(C) extends 1(A) in the number of periods (from two to four), but still has two treatments (A or B) and two sequences (ABAB, BABA). Designs such as this, in which clusters cross between treatments more than once, are known as ‘multiple-period’ or ‘longitudinal’ cluster designs (10).

The CRXO design shares features in common with parallel group cluster trials (where *clusters* are randomised to treatment conditions, which they remain in for the duration of the study), individually randomised crossover trials (where individuals are *randomised to sequences of treatments*) and stepped-wedge trials (where *clusters* are *randomised to sequences of treatments*, but where the clusters can only cross in one direction (e.g. from A to B)) (13, 14). For this reason, the CRXO design shares some of the potential benefits and limitations of these designs.

Development of the CONSORT CRXO extension

Development of this extension was based on guidance for developing health research reporting guidelines (15). We registered our intent to develop the CONSORT CRXO extension on the Enhancing the QUALity and Transparency Of health Research (EQUATOR) network (23 June 2017). A summary of the development process is depicted Figure 2, and details are provided in Supplementary materials 1.

The CONSORT CRXO extension

Scope of the guideline

This CONSORT extension has been developed for CRXO designs in which clusters of participants are randomly allocated to sequences of treatments over time, where the sequences alternate between intervention and control conditions (e.g. AB, BA). However, much of the extension applies to a broader range of designs where at least some of the clusters cross forwards and others backwards at least once between treatments (e.g. AA, AB, BA, BB). Note that trials where crossing between interventions arises solely from protocol deviations are excluded. One-way crossover (or cross-forward) designs where clusters can only cross in one direction (e.g. from A to B) are also excluded (14). A prominent subset of these cross-forward designs is the stepped-wedge design, for which a CONSORT extension is available (11).

How to use the CONSORT CRXO extension

The checklist includes 28 items (with a total of 43 sub-items) (Table 2), along with a separate checklist for journal and conference abstracts (Table 3), and sample size items (Table 4). Some items are identical to those in the CONSORT 2010 statement or relevant extensions, while others have been modified or are new (see Supplementary materials 2, Tables S1 and S2 for details, along with evidence of reporting quality for items in Table S3). New items on research ethics review (item 26), data sharing (item 27) and patient and public involvement (item 28) have been included. In line with recommendations to provide greater clarity of the treatment effects being estimated, we incorporate estimands into item 12a (16, 17). For CRXO trials, items in this extension should be addressed, replacing those in the CONSORT 2010 statement.

For each item, the CRXO item, an explanation, and example(s) are presented. For explanations and examples for items 2a, 4b, 6b, 11b, 14b, 21, 22, 23, and 25, readers are referred to the CONSORT 2010 statement. Other CONSORT extensions may also need to be consulted depending on the attributes of the trial (e.g. the CONSORT extension for health equity in trials where this is a focus (18)). The checklist structure guides where to report information, but this structure should not be seen as prescriptive. Given most journals offer online supplements (and the availability of open

access repositories), journal space restrictions should not preclude complete and accurate reporting of this minimum set of items.

Title and Abstract

Item 1a: Title

Identification as a cluster randomised crossover trial in the title.

Explanation: Identification of the study design in the title facilitates appropriate indexing in bibliographic databases (3) for searching, and ensures ease of identification of randomised trials for potential inclusion in evidence synthesis products (e.g. systematic reviews). Further, alerting readers to the study design early in the article allows them to consider implications for the analysis methods and any potential biases.

Example of item 1a

“Cluster-Randomized, Crossover Trial of Head Positioning in Acute Stroke” (19)

Item 1b: Abstract

Structured summary of trial design, methods, results, and conclusions (see Table 3).

Explanation: Clear, accurate and sufficiently detailed abstracts are necessary for a number of reasons: readers commonly use abstracts to decide whether or not to read the full article; some readers may not have access to the full article and may base healthcare decisions on the abstract only; in systematic reviews using multi-stage screening, the eligibility of studies at the first stage of screening is based upon abstracts; and finally, to ensure appropriate indexing in electronic databases. A minimum set of items to be reported in an abstract of a CRXO trial is presented in Table 3. An adapted version of a published abstract demonstrating reporting against the items is provided (see Supplementary materials 3, Table S1).

Introduction

Item 2b: Objective

Specific objectives or hypotheses.

Explanation: Specifying objectives allows readers to understand the questions the trial is designed to address (20). In cluster trials, the treatments may be intended to change outcomes at different levels, for example, the cluster level (e.g. clinicians’ adherence to guideline recommendations) or the patient level (e.g. mortality), or both. Furthermore, trials may be undertaken for different purposes; that is, to inform clinical decision making (pragmatic intention) or to understand the effects of the treatments in ideal experimental conditions (explanatory intention). Clearly reported objectives will provide clarity about the purpose of the trial and its intended effects. The articulation of research questions may be aided by the use of a question formulation framework (20); for example, the PICOTS framework (**P**atient/**P**opulation, **I**ntervention, **C**omparison, **O**utcomes, **T**iming and **S**etting) or the estimands framework (see item 12a).

Example of item 2b

Example (Reporting of PICOTS elements; pragmatic intention) – “...we conducted a pragmatic, multicenter, cluster randomized, crossover trial. We hypothesized that umbilical cord milking would reduce admission to the neonatal intensive care unit (NICU) compared with early cord clamping in nonvigorous newborns born between 35 and 42 weeks’ gestation...” (21)

Methods: Trial design

Item 3a: Trial design

3a.1 Rationale for a cluster crossover design.

3a.2 Description of the realised trial design including:

- number of treatment conditions;
- definition of cluster (i.e. the unit of randomisation);
- number and duration of periods;
- number and composition of sequences (e.g. ABAB, BABA);
- number of clusters randomised to each sequence;
- duration of any washout periods;
- whether the participants assessed in different periods are the same people, different people, or a mixture of the two; and
- consideration of potential for carryover effects.

A diagram of the trial is recommended when there are more than two periods and/or treatment conditions.

Explanation: A rationale for the choice of trial design should be provided. For a CRXO trial, this should include justification of the cluster and crossover aspects of the design (12). Justifying cluster randomisation is important because this will likely increase the required number of participants, as compared to an individually randomised trial. Thus, more participants will likely be exposed to an intervention of unknown effectiveness than if individual randomisation were used. Furthermore, cluster randomisation may increase the potential for bias arising from selective identification and recruitment of participants. The crossover element may exacerbate this risk in cross-sectional designs, since participants in the second period are, by definition, recruited after randomisation and it may be harder to blind those responsible for recruiting in the second period. The crossover element may lead to an increased risk of cluster withdrawal due to additional burden of participating in multiple periods. Hence, justifying the need for the crossover element is also important.

Possible justifications for adopting a *cluster* randomised design include the ability to evaluate interventions directed at clusters, operational and logistical benefits, and the need to avoid within-cluster contamination (when interventions are delivered at the individual-level) (10); possible justifications for using the *crossover* include infeasibility of a parallel design due to the availability of too few clusters to achieve the desired power (22); to reduce the impact of chance imbalances in cluster characteristics; and, to increase the statistical efficiency (with associated benefits such as minimising the number of participants exposed to the interventions, and reducing the cost).

A key requirement of the crossover design is that the effect of an intervention given in one period does not carry over into the next period (23). In a CRXO trial, carryover can occur at both the cluster and individual levels (24). At the cluster level, the treatment delivered in one period may continue to be delivered in subsequent period(s). At the individual level in cross-sectional designs, participants that are exposed to a treatment for a long period (e.g. patients receiving care in an intensive care unit) may unintentionally receive the other treatment if they are still in the cluster when it crosses over. In open or closed cohort designs where participants are followed over time, carryover may arise in the same ways as in individually randomised crossover designs (i.e. the treatment itself may persist (e.g. physical persistence of a drug) or the effects of the treatment may persist (e.g. a drug has a curative effect) (8, 23, 25). Therefore, methods for managing potential carryover (e.g. the use of washout periods, along with justification for the duration, different participants in each period,

blinding of trialists involved in the delivery of the intervention), or justification for why carryover at both the cluster and individual level is expected to be negligible, should be provided. In a 2x2 crossover design, carryover is particularly problematic since it cannot be distinguished from a treatment by period interaction (23).

In CRXO trials where period effects are expected, any design approaches used to limit the possible impact of such period effects should be noted. An example of a period effect is seasonal variation in outcomes which might be expected when the trial spans a year and the outcome (e.g. incidence of influenza) varies by season. A learning effect is another form of period effect which might arise when fidelity to each intervention improves similarly in the second period. The use of balanced designs and “dual pair sequences” (e.g. ABB and its ‘dual’ BAA) may limit the potential for period effects to bias the estimation of treatment effects (10).

Providing specific details of the CRXO design is important to allow assessment of the potential for bias, and whether the sample size and analysis methods were appropriate. Where possible, justification for the chosen design elements should be provided (e.g. providing a rationale for the chosen composition of sequences). A diagram can efficiently and clearly communicate the design details, particularly for complex designs (more than two periods or treatment conditions or both) or simple designs with staggered commencement of clusters. Key details to depict for each sequence include the commencement date and duration of each treatment condition and washout period, and the number of clusters allocated (see Example 3 of item 3a.2). Such a diagram will clearly signal any imbalance in the treatments evaluated in each period, and therefore the need to control for period effects.

Example of item 3a.1

Example 1 (Rationale for cluster randomisation and crossover) – “Cluster-randomisation with hospitals as clusters was used to overcome confounding by indication and to prevent contamination between isolation strategies. Crossover of strategies at the hospital level was aimed at reducing between-hospital variability and, thus, increasing statistical efficiency.” (26)

Examples of item 3a.2

Example 2 (Description of cross-sectional design) – “Hospitals were randomized 1:1 ... to UCM [Umbilical Cord Milking] or ECC [Early Cord Clamping] in Period 1 (5 hospitals per treatment group from January 2019 to January 2020) until half of the required enrollment was reached (N=600 consented subjects); they were then crossed over to the other intervention during Period 2 (February 2020 to May 2021) for the remaining half of consented subjects (N=600). A 1- to 2-month washout period occurred after Period 1.”(21)

Example 3 (Description of closed cohort design; number of schools allocated to each sequence shown in figure 2 of Makris et al(27)) – “The trial was conducted in six primary schools with two periods (40-days organic diet vs. 40-days of conventional diet) ... A washout period was not required as it was intrinsically included in the two periods, since the first urine sample [for measurement of the primary outcome] of the second period was collected about 12 days after the beginning of the second period. Moreover, the pesticide half-lives are short (half-lives ranging 6.4–16.5 hours for pyrethroids and 5–33 hours for neonicotinoids), so no carryover effect was expected.”(27)

Example 4 (Diagram of realised design from MedBridge trial (28)) –

		2017												2018								
Hospital	Ward	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct
Uppsala	AIM	I2			I1	C			C		C		I1		I2							
	IM	C			I2	I1			I1		I1		I2		C							
Gävle	Geriatric	C		I2					I1		I1	I2	I2		C							
	Stroke	I1		C		I2			I2		I2	C	I1		I1							
Enköping	IM 1			I1		I1	I2			I2		C		I1		C		I2				
	IM 2			C		C	I1			I1		I2		C		I2					I1	
Västerås	ASN									I2		I1		C		C		I1			I2	
	DN									I1		C		I2		I2		C				I1

AIM indicates acute internal medicine; ASN, acute stroke and neurology; DN, diabetes and nephrology; and IM, internal medicine.

Item 3b: Trial design

Important changes to planned methods after trial commencement (such as eligibility criteria), with reasons.

Explanation: Changes from the planned methods may need to be made after the trial has commenced. Important changes should be fully reported to enable readers to appropriately interpret the results. For example, changes in a CRXO design might arise from delays in recruitment of either clusters or participants within clusters, which could lead to unintended staggered commencement of the trial in different clusters, or differences in period lengths. Such changes have the potential to bias estimates of treatment effect, and therefore are important to describe.

Example of item 3b

Example (Deviation from planned dates) – “The study had 1 major deviation. Originally, the initial phase was to run from November 2017 to February 2018, and the crossover phase from March to June 2018. However, because of logistics issues faced by all clinics with obtaining seasonal influenza vaccine supplies ahead of the midyear season, commencement of the crossover phase was delayed by 1 month. We retained the study design of two 4-month phases and ran the initial and crossover phases from November 2017 to February 2018 and from April to July 2018, respectively.”(29)

Methods: Participants

Item 4a: Participants

Eligibility criteria for clusters and participants.

Explanation: A comprehensive description of the eligibility criteria is required to help readers determine the population to whom the results of the trial were intended to be generalised. The criteria should be framed in such a way that a reader can recognise whether the trial includes typical or atypical clusters and participants (30). Inclusion and exclusion criteria need to be reported for both clusters and participants. In addition, in some cluster trials, there may be multiple levels of participants (e.g. emergency department staff and patients attending the emergency departments), and eligibility criteria should be reported for each.

Example of item 4a

FLUID trial – “Inclusion criteria at *hospital level*: participating hospitals were required to have a level II or III ICU [Intensive Care Unit] as these hospitals have the capability of admitting patients that are more severely ill and in turn may receive more fluid administration than hospitals with a level I ICU.

Exclusion criteria at the *hospital level*: we excluded hospitals that had fewer than 6000 acute care admissions per year (<1500 admissions per study period).

Inclusion criteria at *patient level*: adult and paediatric patients admitted to the participating hospitals for the first time in the previous 90 days (index admission) over the duration of each study period were included in FLUID (to avoid exposure and thus potential contamination with either crystalloid fluid in the prior 90 days).

Exclusion criteria at *patient level*: neonates were excluded from FLUID since RL [Ringer's Lactate] is neither used nor recommended for use in this population. Patients who were readmitted to hospital during study period 1 or 2 were excluded to avoid contamination with previous FLUID exposure. Patients admitted during the run-in or run-out study periods were also excluded." (31)

Methods: Intervention

Item 5: Intervention

The treatment conditions with sufficient details to allow replication, and whether they were delivered at the level of the cluster, the individual, or both.

Explanation: A complete description of the intervention conditions is necessary for their future implementation (32). The TIDieR (Template for Intervention Description and Replication) checklist, provides an extension of the CONSORT statement (item 5) (3), for describing interventions (32). Complete description also allows assessment of the potential for carryover effects (Table 1). Unlike other cluster designs (parallel, stepped-wedge), the CRXO design is not suitable for evaluating treatments where their effects persist, such as educational interventions targeted at changing clinicians' practice (23).

For cluster trials, reporting whether the treatments are delivered at the cluster level, the individual participant level, or both, is important. This information allows assessment of whether individuals can avoid treatments (which is important for determining the coverage of the treatment), and has implications for the identification of research participants and requirements for informed consent (items 10c and 26).

In CRXO trials, treatments are commonly delivered at the individual participant level (e.g. daily bathing of critically ill children in paediatric intensive-care units with chlorhexidine gluconate compared with standard bathing practices), since these treatments can be withdrawn (8). Treatments delivered at the level of the cluster can also be evaluated using the CRXO design, as long as the intervention can be withdrawn (e.g. enhanced cleaning of hand contact surfaces within intensive care wards compared with standard cleaning, where the active treatment component of extra hand gel on wards can be removed).

Examples of item 5

Example 1 (Description of the treatment conditions; level of delivery to the individual) – “In each group, the intervention was applied at the patient level.

For patients included during experimental periods, hospital pharmacists performed the medication reconciliation at discharge. ... Hospital pharmacists completed a short form documenting the reason for hospitalisation, home medication modifications, new medication and laboratory results necessary for community pharmacists to understand and/or accept the prescription... They also checked the discharge prescriptions (drug added and/or omitted, different dosage, route or duration of treatment) and, if needed, made an intervention on the physician's prescription according to SFPC [French Society of Clinical Pharmacy] standards

[standards provided in a figure] to change the prescription. Then they explained to the patient the drug initiated and the modifications to the home medication. They phoned the patient's community pharmacist to explain the patient's inclusion in the study, the discharge time, and the modifications in treatment. They also sent the prescription sheet to the community pharmacist before patient discharge. The patient or caregiver then visited the community pharmacist as usual.

For the control group, patients received the usual care already implemented both at the hospital (classical drug dispensation by staff pharmacists) and by their community pharmacist (drug dispensation according to the prescription sheet written by the hospital physician in addition to the general practitioner's sheet [if present])." (33)

Example 2 (Level of delivery to the cluster) – “The study fluid interventions were implemented at the hospital level using a hospital policy or strategy, with the aim to answer our study question at the hospital level.” (31)

Methods: Outcomes

Item 6a: Outcomes

Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed (for specific guidance see CONSORT for outcomes).

Explanation: Outcomes should be completely defined, such that the resulting treatment effect estimates can be interpreted and used in other research (e.g. systematic reviews). Here, we focus on issues particular to CRXO trials, and refer readers to the CONSORT 2010 statement and the CONSORT-Outcomes extension for further discussion of aspects common to all designs and examples (3, 34).

In cluster randomised trials, outcomes in each period can be measured at the level of the individual, or the level of the cluster. The latter could reflect a true cluster level outcome or arise because only aggregated individual level data are available (e.g. the number of hospital-associated infections per period may be available only in summary form) (11). The level of the outcome measure has implications for the method of analysis (see item 12a).

The timing of outcome assessments should be reported, including whether the outcomes were measured at discrete points in time common to all participants (e.g. via a survey at the end of each period), or at a time point specific to each participant (e.g. number of adverse events per patient until the time of their hospital discharge). Reporting of the timing of follow-up facilitates assessment of whether participants had similar (or variable) exposure to the treatment condition, and if there was the possibility of contamination of observations collected in the second (or later) periods (the latter may arise in cross-sectional designs in which the treatment is delivered at the level of the cluster).

In CRXO trials, data may be sought from routinely collected sources (e.g. registries, hospital medical records) or purposively collected (e.g. questionnaires, patient diaries). Reporting the methods for data collection is important for allowing readers to assess whether the methods are appropriate for measuring the intended outcomes, and facilitates an assessment of the potential for bias arising from measurement of the outcomes (35).

Examples of item 6a

Example 1 (Description of outcomes; primacy; outcome measured at level of individual; time point specific to each participant) – “Primary outcome and timing: All the outcomes pertain to the

individual participant. The primary outcome was admission to the NICU [Neonatal Intensive Care Unit] in the first 24 hours of life for predefined criteria as follows: respiratory distress (tachypnea, grunting, or retractions), bradycardia or tachycardia, hypotonia, lethargy or difficulty arousing, hypertonia or irritability, poor feeding or emesis, hypoglycemia, oxygen desaturations or cyanosis, need for oxygen, apnea, seizures or seizure-like activity, hyperbilirubinemia, and/or temperature instability.

Secondary outcomes and timing: The predefined secondary safety and efficacy outcomes included therapeutic hypothermia, volume expanders, phototherapy, hemoglobin at 24 hours of life, and peak serum bilirubin. ..." (21)

Example 2 (Outcome measured at level of cluster; use of administrative data) – “Primary feasibility outcome: Adherence to the fluid protocol: Adherence to the study fluid was measured not at the individual patient-level, but according to the aggregate use of the study fluid throughout the hospitals (all hospital wards, monitored units, and departments) using the hospital inventory system; monitoring fluid exposure or adherence according to individual patients was not feasible due to the sheer number of hospital admissions.

...

Data collection: All follow-up and collection of data for enrolled patients at the participating hospitals were captured through health administrative data that are housed at the Institute for Clinical Evaluative Sciences. There were no individual patient level data collected by research coordinators in the participating hospitals." (31)

Example 3 (Description of outcomes; primacy; outcome measured at level of individual (implicit); time point specific to each participant; data collection methods) – “The primary outcome was symptomatic VTE [venous Thromboembolism] within 90 days of surgery. ...

Data source and timing: Patients received a web link via email or text message to complete online data collection at 90 days postoperatively. ... Data collection forms were reviewed by registry staff and all patients who responded “yes” to having had a venous thromboembolism, a secondary operation within 90 days or 6 months, a major bleeding event within 90 days, or a joint-related readmission within 90 days had this outcome verified through written documentation from treating physicians. These results were verified by the trial outcome verification committee. For the primary outcome, the date and type of VTE [venous thromboembolism] (below- or above-knee deep venous thrombosis [DVT] or pulmonary embolism) and side of DVT (left or right) were recorded. Mortality data were collected through linkage between the registry and the National Death Index." (36)

Methods: Sample size

Item 7a: Sample size

How sample size was determined. Method of calculation and relevant parameters with sufficient detail so the calculation can be reproduced (Table 4). Assumptions made about correlations between outcomes of participants from the same cluster.

Explanation: Despite improvements over time, most randomised trials remain insufficiently powered to detect the small effects that are generally observed (37). Reporting the sample size calculation in sufficient detail is important to ensure reproducibility and for allowing assessment of the plausibility of the assumptions, and identification of any scientific or ethical concerns arising (38).

The method used to determine the sample size should be described, specifically noting how the calculation accounts for both the cluster randomisation and the multiple period aspects of the design (e.g. (39-41)). In the CRXO, as in any cluster randomised trial, it is important to account for the intra-cluster correlation coefficient (ICC); however, in a CRXO trial, a single ICC may not be adequate since the strength of the intra-cluster correlation may depend on the separation in time of the observations. The recommended minimum items to report are given in Table 4. The details to provide are to some extent dependent on the particulars of the design (e.g. whether the design is cross-sectional or cohort).

Justification for the choice of parameter values such as ICCs should be provided (e.g. based on estimates reported in a repository of correlations or extrapolation from known determinants of correlations such as whether the outcome is a process or clinical outcome) (42, 43). Often sample size parameters (e.g. ICCs, cluster autocorrelation coefficient) are based on estimates with large uncertainty. Therefore, it is important in the protocol to provide an indication of the sensitivity of the sample size, or power, to the assumed parameter values, and this should be available with the trial report (38).

In some CRXO trials, the sample size may be fixed by constraints on the number of clusters, participants, or periods. In this circumstance, the effect size that the trial was powered to detect should be reported. If no power calculation was performed, this should be reported. Retrospective power calculations based on the results of the trial are of little merit and are discouraged (44).

Example of item 7a

PEPTIC trial (some core items are described elsewhere in the publication: 2 periods, 2 sequences, 1:1 allocation ratio) – “There is no established consensus regarding what constitutes a minimal clinically important difference in in-hospital mortality for critically ill adults. This (...) makes it challenging to establish what difference in mortality one should power an ICU study to detect. ... Accordingly, we sought to conduct the largest trial that we considered feasible in order to provide the most precise estimates of the likely mortality treatment effect possible with PPIs [Proton pump inhibitors] vs. H2RBs [histamine-2 receptor blockers]. In the study protocol the sample size was calculated [reference to sample size method] to have 80% power at a 5% significance level to detect a 2.4% absolute reduction in in-hospital mortality from a baseline mortality of 15%, which corresponds to a relative risk reduction of 16%. This calculation assumed a mean cluster-period size of 310 patients with coefficient of variation of 0.50, a within-cluster-within-period correlation of 0.035, and a within-cluster-between-period correlation of 0.025, yielding a cluster autocorrelation of 0.71.” (Supplement 2 of Peptic Investigators (45))

Item 7b: Interim analyses

When applicable, explanation of any interim analyses and stopping guidelines.

Explanation: Interim analyses of outcomes are undertaken to inform the decision of whether a trial should be stopped early if there is substantial evidence of harm, effectiveness, or futility. Authors should report the dates or times at which they were carried out, what triggered them, whether these were planned or ad hoc, and the statistical methods used (including any formal stopping rules). When stopping rules/guidelines are used, it is important to report whether these were planned before commencement of the trial and prior to assessment of any interim analysis results, or some time thereafter.

For CRXO trials, there is little methodological guidance on how to do an interim analysis of outcomes, and interim analyses are challenging due to the complexity of this design. For example,

whether commencement of clusters is concurrent or staggered will impact the type of interim analysis. If commencement of clusters is concurrent, in a two-period design when the analysis is undertaken at the end of the first period, the design would reduce to a parallel cluster randomised trial, necessitating a different analysis from the planned final analysis. If commencement of clusters is staggered, some clusters may have completed both treatments, others only one, and some a mix of partially completed treatments. Any interim analyses should address these issues and the properties of any stopping rules should be determined, with these details reported. As with any trial, the planned interim analyses (where a decision is to be made about continuation of the trial) should be allowed for in power calculations to control the overall Type I error rate (11).

There may be other important reasons for considering stopping a trial arising from the trial conduct itself; for example, inability to recruit clusters or clusters unable to recruit participants as planned, poor quality of the primary outcome. Reporting the date the trial was stopped, along with the reasons for stopping, is recommended.

Example of item 7b

“An interim analysis was not initially planned, as both treatments are considered standard practice for venous thromboembolism prophylaxis in Australia and the trial is investigating an adverse event as the primary outcome. However, due to concerns of an increased adverse event rate (symptomatic thromboembolism and death) in one of the prophylaxis groups, a Data Safety Monitoring Board (DSMB) was convened 1 year into patient recruitment. ... The DSMB were advised by the Trial Management Committee (TMC) to conduct an interim analysis. In conjunction with the DSMB (prior to the interim analysis), the TMC applied the Haybittle-Peto stopping rule of a two-sided significance of 0.001 for the primary outcome in population 1*.

This stopping rule was chosen as it does not require adjustment of the significance threshold for the final analysis and allows further interim analyses using the same threshold (if required). After the first interim analysis (in September 2020), the DSMB recommended continuing the trial and performing a second interim analysis in November 2020. ...

Interim analyses were conducted for thromboembolism and mortality within 90 days for population 1*. To account for unequal cluster sizes, incomplete crossover or clusters which had not yet crossed over, a composite analysis was designed. For clusters which had crossed over, including with partial completion of the second period, the cluster weighted estimator intended for the primary outcome was used. Harmonic mean weighting when there are unequal cluster sizes has been shown to improve precision and 95% confidence interval coverage compared with unweighted or inverse variance estimates [Reference provided for methods]. Clusters which had not crossed over were analysed using the cluster period summaries, weighted by cluster size, in a parallel design approach, i.e., as if it were a cluster randomised trial without crossover. Estimates for the two approaches were combined using inverse variance weights to provide a final estimate. Confidence intervals were constructed using the Haybittle-Peto boundary of 0.001.” (Supplement 2 of Sidhu *et al* (36))

* Population 1 = all registered patients undergoing primary total hip arthroplasty or total knee arthroplasty for a diagnosis of osteoarthritis.

Methods: Randomisation

Item 8a: Allocation schedule generation

Method used to generate the random allocation schedule.

Explanation: In a CRXO trial, clusters are randomly allocated to a sequence of treatments, where the sequence defines the order in which each cluster will receive the treatments. For example, in a two-treatment two-period design, there may be two sequences: treatment A followed by treatment B (sequence 1) and treatment B followed by treatment A (sequence 2). The allocation schedule dictates the random allocation of clusters to either sequence 1 or 2. Note that our use of the term ‘sequence’ is different to its usage in the CONSORT statement (3) and the cluster CONSORT extension (5), where for these parallel group designs, ‘sequence’ refers to the randomly generated list of allocations to *a single treatment* (e.g. treatment A or treatment B). In this CRXO statement, the term ‘schedule’ is instead used to refer to this randomly generated list of allocations to *sequences of treatments* (e.g. AB, or BA; see Box 2 for an example).

The set of candidate sequences and the method used to generate the random allocation of clusters to these sequences needs to be sufficiently described so that an assessment can be made as to its adequacy. Use of general terms such as ‘random’, ‘randomisation’, ‘random allocation’ are insufficient. Details of the method are required (e.g. computer-generated random numbers).

Example of item 8a

“Each hospital unit is randomised to perform medical reconciliation or usual care for the first 14-day period and is crossed over to the other group for a second 14-day period.” (Figure 1 of Pourrat *et al* (33))

“The randomisation sequence was generated ... using a computerised process.” (33)

Item 8b: Randomisation method

Type of randomisation; details of any restricted randomisation, if used.

Explanation: In a CRXO trial, randomisation of clusters to the sequences of treatments is often done at a single point in time before the trial commences. In this circumstance, it may be possible to balance (restrict) the randomisation such that an equal number of clusters are allocated to each of the sequences of treatments. This can be achieved using a single block balanced across sequences for the entire study. For example, in a two-period two-treatment and two-sequence (AB or BA) design with twenty clusters, a single block of size twenty balanced across the two sequences would allocate ten clusters to AB and ten to BA.

Stratification in CRXO trials may also be used when there are clusters with distinct characteristics (e.g. different case-mix, different hospital protocols for the delivery of care) which are related to the trial outcomes, or where commencement of clusters is staggered in time and there is potential for secular / seasonal trends. In stratified designs, balanced allocation of clusters to sequences are generated independently within each stratum. Other methods of restricted randomisation that may be employed include covariate-constrained allocation (46), minimisation (47) and matching (48).

When reporting the type of randomisation, it is important to indicate whether clusters were allocated at one time, in batches or sequentially, and whether simple or restricted randomisation (e.g. permuted blocks) was used. If permuted block randomisation was used, details should be provided on how the blocks were generated, the block size(s), and whether the block size was fixed or randomly varied. If stratification was used, details should be provided about the stratification

variables and the cutpoints for categorisation (3). If another method of restricted randomisation was used, the method should be described, along with the variables used to restrict the randomisation.

Example of item 8b

Example 1 (Restricted randomisation using stratification; sequential allocation) – “The ICUs [Intensive Care Units] were consecutively randomized in a 1:1 ratio using computer-generated randomization with random block sizes of 2, 4, and 6 and stratified by number of ICU beds (≤ 10 or >10).” (49)

Consecutive randomization was “as per the date of approval by the institutional review board” (Supplement 1 of Rosa *et al* (49))

Example 2 (Restricted randomisation using stratification; batch allocation) – “Participating ICUs [Intensive Care Units] were randomized to the order of treatments in a 1:1 ratio when ethics and regulatory approvals were in place at each study site. Randomization was performed using computer-generated random numbers and was stratified by region and time period with a minimum of 4 ICUs randomized at a time.” (45)

Example 3 (Simple randomisation) – “Schools were randomized *a priori* to two groups that differed in the sequence of the treatments; organic diet followed by conventional diet (Group 1) or conventional diet followed by organic diet (Group 2) (27).”

“Schools were randomly assigned following simple randomization procedures of no restriction or matching.” (Supplement 5 of Makris *et al* (27))

Item 9: Allocation concealment

Specification that allocation was based on clusters; description of any methods used to conceal the allocation from the clusters until after their recruitment.

Explanation: The benefits of randomisation are only realised when randomisation is properly implemented. In individually randomised trials, this requires generation of a random allocation schedule, and concealment of upcoming assignments from the participants and from those responsible for recruiting the participants, until after recruitment (known as allocation concealment) (50). Analogously, in cluster trials, where randomisation occurs at the level of the cluster, allocation concealment involves concealment of upcoming assignments from the *clusters* (i.e. the ‘gatekeeper’ or ‘guardian’) and from those responsible for recruiting the *clusters*. The randomisation process in CRXO trials may involve randomising clusters sequentially, in batches, or all at once (51). The strategies implemented to conceal the allocation will differ depending on the randomisation process, and should be completely described. As with other cluster randomised trials, allocation concealment is likely to be preserved when randomisation occurs at a single point in time (51).

In cluster randomised trials, participants may be recruited (or identified) after the clusters have been allocated. This can lead to recruitment/identification bias arising from selective recruitment (or identification) of participants across treatment conditions when those responsible for the recruitment are aware of the cluster’s allocation in a particular period. In this statement, the process for recruitment of individual level participants is considered separately (item 10b) to the randomisation process.

Example of item 9

Example 1 – “Crossover and randomization took place at cluster (ward) level within each hospital. ... The computer-generated codes were held by the statistician to assure allocation concealment until the moment of randomization.” (Supplement 3 of Kempen *et al* (28))

Example 2 – “Use of sequentially numbered sealed opaque envelopes enabled concealment of the allocation sequence at the hospital level until the last hospital was randomised.” (26)

Methods: Implementation of randomisation

Item 10: Implementation

Replaced by 10a, 10b and 10c (listed below).

Explanation: Consistent with the CONSORT extension for other clustered designs (5, 52), it is important that all steps in the implementation of the randomisation and recruitment processes for both clusters and individual level participants are described. Reporting of these processes allows readers to assess the potential for bias. Details of the allocation and enrolment process of the clusters are described in item 10a, with the corresponding detail provided for participants in item 10b. The consent process for participants is also addressed under implementation of the randomisation methods (item 10c) since this is an integral component of enrolment, with potential implications for bias (e.g. when consent procedures differ across treatment conditions which may lead to different types of participants being recruited). Use of the ‘timeline cluster tool’, which depicts the time sequence of trial processes (including the randomisation process) and the blinding status of participants and trial staff at each stage, allows readers to easily assess threats to the internal validity of the trial (53).

Item 10a: Inclusion of clusters

Who generated the random allocation schedule, who enrolled clusters, and who assigned clusters to sequences of treatments in the schedule.

Explanation: Reporting who generated the random allocation schedule, who enrolled the clusters, and who assigned the clusters to the sequences of treatments is important for allowing assessment of the potential for bias (51). A guiding principle to achieve adequate allocation concealment involves separation between the person who generates the allocation schedule and those who enrol and assign the clusters to the sequences of treatments. This can be achieved by having someone (preferably a statistician) independent of the trial generate the allocation schedule. When this person is unaware of the identity of the cluster (e.g. where they receive a non-identifying unique code for each cluster), they may also assign the clusters to the allocation schedule (i.e. equivalently maintaining separation). Failure to separate the above processes risks potential subversion of the randomisation schedule and potential biased estimates of treatment effect.

Example of item 10a

Example 1 (who generated the random allocation schedule, who enrolled clusters) – “Potentially eligible health services were identified in consultation with Departments of Health and Aboriginal Community Controlled Health Organisations and enrolled by study coordinators (AT, BH, and SGB). ... The trial statistician (HW) did the randomisation Participating health services were told their group allocation once participation agreements were signed.” (54)

Example 2 (who generated the random allocation schedule, who assigned clusters to the sequences of treatments) – “To remove the potential for allocation bias, one statistician generated the allocation codes and another randomly permuted clinic names within strata. A third person matched the allocation codes with clinic names to reveal the allocations.” (55) [Note that the role of the third person should be specified.]

Item 10b: Inclusion of participants

Mechanism by which individual participants were included in clusters for the purposes of the trial (such as complete enumeration or random sampling; continuous recruitment or ascertainment; or recruitment at a fixed point in time), including who recruited or identified participants.

Explanation: As with other cluster randomised designs (5, 52), there are different sources from which the individual participant data can be obtained (e.g. directly from the patient or from review of their medical records), different sampling strategies (e.g. complete enumeration or random sampling), and different timings of the data collection (e.g. continuously or at fixed points in time). These mechanisms have differing risks of bias.

In some CRXO trials, data are not directly sought from participants, but instead obtained from routinely collected sources (such as hospital medical records). In these trials, participants are 'identified' (ascertained), but may not be approached directly for recruitment into the trial. Alternatively, participants may be 'recruited' into the trial and may need to provide data prospectively, for example through questionnaires. Recruitment may occur continuously or at fixed points in time. Recruitment or identification may involve approaching or identifying all participants or a sample.

Reporting the mechanisms for including individual participants, covering who was involved in recruiting or identifying participants, and whether they were blinded to the cluster's treatment allocation in a particular period, allows assessment of the likelihood of bias arising from selective recruitment or identification of participants (51). In trials where participants are recruited or identified prior to randomisation, there is no potential for recruitment/identification bias (since selection of participants cannot be influenced by the cluster's treatment allocation, which is unknown prior to randomisation). However, in CRXO trials, the possibility of using this approach is limited to only a closed cohort design (where the same participants are included in all periods), which is a less common design (8). More commonly, participants are identified or recruited post randomisation. In this circumstance, recruitment or identification of participants by a person blind to the allocation of the cluster can help mitigate recruitment/identification bias. In trials where there is complete enumeration, there may be less scope for selective recruitment or identification of participants, even when the recruiter is aware of the cluster's assigned treatment; nevertheless, blinding the recruiter to the treatment assignment is always advisable (56).

Example of item 10b

Example 1 (Continuous recruitment; recruiters unblinded to cluster's treatment allocation) – "All adult patients were eligible, except those who stayed in the hospital longer than 21 days, who did not return home, who were in a moribund state, or who were not able to understand the topic of the study or complete a questionnaire." (33)

*"Participants are recruited by unblinded hospital pharmacists." (Figure 1 of Pourrat *et al* (33))*

Example 2 (Random sampling) – "Because of the variability in cluster size, the female participants were only randomly selected for enrolment by probability proportional to size sampling. For the male participants' enrolment, all the eligible participants were enrolled except for in three clusters whose size allowed random selection." (57)

Item 10c: Consent

Whether consent was sought, from whom, when and for what; whether this differed between treatment conditions. Justification for any waiver or modification of informed consent requirements.

Explanation: Recruitment and consent procedures in cluster randomised trials are more complex than in individually randomised trials (58). In individually randomised trials, informed consent is usually obtained from individual participants, and ‘consent’ generally encompasses being randomised, receiving study interventions and providing data for outcome assessment. In cluster randomised trials however, clusters are randomised — often before individual participants can even be identified for the trial, while study interventions and data collection procedures can be administered to different types of participants. Thus, the randomisation units, the units exposed to the intervention, and the units on which outcomes are assessed may be different. For example, hospitals may be randomised, the intervention may be administered to providers, while outcomes are assessed on patients. Reporting of explicit details about the recruitment and consent procedures in CRXO trials is essential to avoid confusion. A simple statement that “participants provided informed consent” is not adequate; instead, referring to specific participants (e.g. providers, patients) and indicating what consent pertains to (e.g. exposure to the intervention or data collection) aids clarity (59). While “gatekeepers” or “cluster guardians” can play an important role in facilitating the recruitment of clusters into the trial (60), gatekeeper permission for trial participation should not be confused with consent from research participants (61).

Clarity around the timing of any consent relative to the timing of randomization and intervention delivery is also particularly important in cluster randomised trials, because post-randomization consent increases the risks of bias (62). In a minority of CRXO trials (those with closed cohort designs), it may be feasible to identify and recruit individual participants before cluster randomisation, but the majority of CRXO trials have cross-sectional designs in which participants are recruited prospectively (8). In CRXO and multiple-period CRXO trials, it is important to describe consent procedures in adequate detail so that it is clear whether consent procedures differed over time and between treatment conditions in different periods as this can lead to bias (63). Steps taken to mitigate the risks of bias should also be clearly described (51, 56). Making recruitment materials and consent forms available in supplementary materials can aid clarity around what information was conveyed to participants about the trial, while providing Timeline Cluster diagrams can aid clarity around the timing of consent (53).

In some CRXO trials in which outcomes are available in routinely collected databases and study interventions pose only minimal risk to research participants, a research ethics committee may give permission for a CRXO to proceed without the explicit informed consent of research participants. When an ethics committee has approved the study with a waiver of informed consent, this should be clearly stated together with the justification for the waiver according to the regulatory requirements of the country or countries in which the CRXO took place.

Example of item 10c

Example 1 (Consent procedures differ between treatment conditions for patients; timing of consent) – “6a. Participant recruitment in the medication reconciliation group. ... [Participants] receive complete information and provide oral consent for intervention and for data collection.” “6b. Participant recruitment in the usual care group. [Participants] receive partial information because they are not aware of the existence of the medication reconciliation group and provide oral consent for data collection.” (Figure 1 of Pourrat *et al* (33))

Timeline Cluster diagram – Supplementary materials 3, Figure S1 (33)

Example 2 (Waiver of consent to exposure to intervention) – “We conducted a cluster randomized multiple crossover trial to evaluate a real-time best practice alert embedded into electronic health records... Institutional review board approval was obtained before patient enrollment... Patient

and provider consent was waived because the alert supplemented the current standard of care, there was no requirement to respond, and the alerts were unlikely to be harmful and might have proven beneficial.” (64)

Methods: Blinding

Item 11a: Blinding

Who was blinded after assignment to sequences (e.g. cluster level participants, individual level participants, those assessing outcomes) and how.

Explanation: Clear reporting of whether those involved in the delivery of the treatments, receipt of the treatments, outcome collection and adjudication, and data analysis, were blinded to the treatment allocation is important for assessing the likelihood of bias. In some CRXO trials participants are unaware of their treatment allocation because they do not know they are in the trial (e.g. when they are not directly recruited), or they may know they are in a study, but not a trial (51). The level of awareness that participants have about their participation in a trial can affect the likelihood of bias. Therefore, explicit reporting of participants’ knowledge of their participation in the trial is important.

Performance bias may arise when healthcare providers or individual level participants are aware of the treatment allocation, and this knowledge leads to deviations from the intended treatment (e.g. use of co-interventions, increased or decreased compliance with the treatment) (35). Similarly, detection bias may arise when the outcome assessor or adjudicator is aware of the treatment allocation, and this leads to differential assessment of the outcome. The outcome assessors can be cluster participants (e.g. patients or healthcare providers that self-report outcomes); healthcare providers responsible for outcome assessment (e.g. a clinical examination); trial personnel; or persons who are independent of delivery of the treatment. In many trials, there will be different types of outcome assessors; therefore, clearly reporting which outcomes were assessed by whom, and whether they were blind to the treatment allocation is necessary.

Parallel and stepped-wedge cluster randomised trials commonly evaluate treatments delivered to cluster-level participants (e.g. educational interventions delivered to healthcare providers to improve practice) (65, 66), and in this circumstance, it will generally be difficult to blind cluster-level participants (51). However, in CRXO trials, treatments delivered at the individual participant level are more likely (since these treatments can be withdrawn – see item 5), and this may provide more opportunity to employ blinding of cluster-level and individual-level participants (8). The consequence of a lack of blinding of those delivering the treatments (e.g. clinicians) on performance bias may be of particular concern in CRXO trials, since if changes in individual-level participant outcomes are observed in the first period by those delivering the treatment, this may lead to the continued use of the treatment condition in subsequent periods (a type of contamination).

Example of item 11a

Example 1 – “Participants and ward staff were not blinded to treatment allocation.” (28)
“All primary and secondary outcome data collection and assessments were blinded to treatment allocation ... Statisticians were blinded to treatment allocation until database closure.”
(Supplement 3 of Kempen *et al* (28))

Example 2 – “Blinding of clinicians and research staff will remain throughout the entirety of the trial. ... Trial participants, care providers, outcome assessors, and data analysts will be blinded to the assignment of intervention. The only research personnel who will be unblinded are the

statistician performing the randomization, the PM [Project Manager], and the research pharmacists in charge of supplying locking kits used during the study. ... Data analysis will occur outside of the unit with the type of intervention received by each patient concealed until the completion of analysis.” (67)

Methods: Statistical methods

Item 12a: Statistical methods

Target estimand for each primary and secondary outcome including whether it pertains to the cluster-level or individual-level; statistical methods for their estimation including how period effects, clustering and repeated measures were taken into account. Any assessment of carryover effects should be reported.

Explanation: Primary reasons for reporting the statistical methods are to allow for reproduction of the results and for readers to assess whether the methods are appropriate for the design, and the research question. The guiding principle should be to provide enough detail such that a knowledgeable individual with access to the original data could verify the reported results (68). Reference(s) to the chosen statistical methodology should be provided, along with details of the statistical software packages (and versions) used to implement the methods, and ideally the statistical code (see items 24 and 27). The link between which statistical method(s) apply to which outcomes, should be made clear. Pre-specified analyses should be distinguished from those that were not planned. This distinction is important because pre-specification limits bias arising from selective reporting of favourable results from fitting multiple analyses (commonly referred to as ‘p-hacking’), and so pre-specified analyses are considered more trustworthy (69, 70).

Reporting the precise description of the treatment effect to be estimated for each outcome (i.e. the estimand) allows readers to understand the specific questions addressed (71), which should align with the study objectives (item 2b). The attributes for completely specifying an estimand are outlined in Box 3. In addition to reporting the standard attributes for individually randomised trials (i.e. attributes (i) to (v)), in cluster trials, whether a *participant-average* or *cluster-average* treatment effect was targeted (attribute (vi)) should be reported (72). The latter will depend on the target of interest; that is, whether interest lies in answering the question “How effective is the intervention on average for a population of participants?” (*participant-average treatment effect*) or “How effective is the intervention on average for a population of clusters?” (*cluster-average treatment effect*). It is additionally important to specify whether interest lies in the *marginal* or *cluster-specific* treatment effect (73).

For cluster randomised trials, analyses can be undertaken at the individual- or cluster-level, with the latter by aggregating observations within each cluster (or cluster-period) into a single summary measure. The approach taken in the trial’s analysis should be reported. Note that the level of analysis is distinct from the level of targeted inference (72). Recent literature concerning parallel cluster trials has indicated that participant-average and cluster-average treatment effects may differ if there is informative cluster size, namely when cluster size is predictive of participant outcomes or of the magnitude of treatment effect (i.e. treatment effect heterogeneity). In such cases the different levels of analysis may be targeting different estimands (72). Although further work is needed to assess the impact of informative cluster size in CRXO trials, it should be reported whether informative cluster size has been considered or addressed in the analysis.

It is well recognised that the statistical analysis of cluster trials must account for the correlation of participant observations within clusters (intra-cluster correlation). In CRXO trials, the strength of this

correlation may depend on the separation in time between the observations (14, 74), along with (where applicable) correlation in repeated measurements from the same participants. Analyses that account for these correlations must be used: analyses that do not allow for these aspects typically yield incorrect estimates of standard errors of treatment effects (and therefore confidence intervals with coverage that is less or greater than the nominal level (e.g. 95%)) (9), and in some circumstances, may bias treatment effects (e.g. certain generalised linear mixed models where the random effects are misspecified) (75). Therefore, reporting the estimation methods and assumptions made about the intra-cluster correlation of observations at different periods or times (e.g. constant correlation, continuous-time correlation decay, exponential decay) is recommended (12).

In CRXO trials where commencement of the trial is staggered across clusters or period lengths differ across clusters, adjustment for secular / seasonal trends in the analysis should be considered, and description of the modelling approach provided (e.g. including calendar time periods as a fixed effect).

Any tests for carryover effects, or adjustment for carryover, should be reported. The two-stage analysis procedure, which involves first testing for carryover effects and determining the subsequent analysis approach based on the result of the test, has been shown to adversely affect estimation of the treatment effect and its statistical properties (76). For these reasons, testing for carryover is not recommended, but if undertaken, or an analysis of only the first period data is presented as a result of the two-stage strategy, this needs to be reported. Similarly, when an attempt is made to adjust for carryover (e.g. in designs with more than two periods), the model used and its assumptions should be described explicitly (77).

Given that CRXO trials typically include a small number of clusters (8), small sample standard error or degrees of freedom adjustments may need to be applied to maintain the nominal type I error rate (e.g. 5%) and confidence interval coverage (78). Any adjustments applied should be reported. Similarly, adjustments for prognostic factors at the participant level to mitigate any imbalance in participant characteristics between periods within a cluster should be reported, noting if they were pre-specified or not.

Any methods for handling of missing data at the individual and cluster level should be described (e.g. (79, 80)). For trials with deviations from the allocated sequences or in the planned crossover timings, details should be provided about how this was dealt with in the analysis. The handling of these two components – missing data and treatment of deviations from allocated sequences – defines whether an analysis follows the principle of ‘intention-to-treat’ (ITT) (3, 81). However, challenges in applying the ITT principle in the analysis of parallel cluster randomised trials have been recognised (62). These issues are further exacerbated in multiple period cluster designs where different definitions of adhering to the allocated sequence might be adopted (e.g. according to the planned dates of crossover, or to the order of the allocation sequence (11)). For these reasons, explicitly reporting how deviations are handled and methods for missing data are preferable to using labels such as ‘intention-to-treat’.

For binary outcomes, reporting both absolute and relative effect sizes is recommended (see item 17b). When a model-based approach has been adopted, the method used to calculate the risk difference should be reported. A number of methods exist, such as regression-based approaches that directly yield estimates of risk differences (e.g. fitting a binomial model with an identity link), or indirect approaches, such as marginal standardisation (82).

Example of item 12a

Example 1 (individual level analysis; allowance for clustering, period and secular trends; method for calculating absolute effect) – “The analysis of the primary outcome used individual patient-level data and generalized estimating equations with a logarithmic link function, an exchangeable working correlation matrix, and robust standard errors using the ICU [Intensive Care Unit] as the clustering unit. Because randomization was performed in batches of ICUs, covariate adjustment for randomization batch, the order of administration of the treatments, and batch × order interaction was performed to allow for separate order and secular time effects occurring in each of the randomization batches. ... Treatment comparisons are presented as risk ratios (RRs) and 95% CIs from the generalized estimating equation analysis, and as absolute risk differences and 95% CIs obtained by marginalizing (ie, standardizing) of the RR model [Reference provided]. The approach to the analyses of the secondary end points appears in the [Reference to supplement]. ... All statistical analyses were completed using Stata version 15 (StataCorp)” (45)

Example 2 (handling of missing data) – “The primary analysis will use all available data with no imputation. In addition, if the mRS [Modified Rankin Scale] at 90 days (the primary endpoint) or NIHSS [National Institute of Health Stroke Scale] at 7 days are missing for more than 10% of patients, missing data will be imputed using a fully conditional specification (FCS). The imputation model will include mRS at 90 days, the NIHSS and mRS scores at 7 days (or hospital discharge, if sooner), a variable indicating the cluster, a variable indicating the period, a variable indicating the intervention, and all baseline variables that are listed in Section 2.6 [sic].... The mRS at 90 days and NIHSS at 7 days will be imputed using an ordinal logistic model. Other variables will be imputed using either linear regression (for continuous/ordinal variables) or a discriminant function method (for nominal variables). Ten sets of imputed data will be created and analysed using the model described in Section Estimates of the treatment effect (β in Model 1) and its standard-errors will be combined to obtain a pooled common OR [Odds Ratio] and 95% CI.” (Statistical Analysis Plan of (19))

Example 3 (cluster level analysis; allowance for period and prognostic factor) – “As primary analysis, the difference in risk of PJI [Periprosthetic Joint Infection] after the intervention treatment compared with the control treatment will be estimated using linear regression for aggregated cluster-period data, with the proportion of events within a cluster-period as the dependent variable and treatment, cluster, period, and proportion of females within the cluster-period as independent variables. An estimated difference in risk with CIs and a 2-sided p-value will be presented.” (83)

Item 12b: Additional statistical methods**Methods for additional analyses, such as subgroup analyses, sensitivity analyses, and adjusted analyses.**

Explanation: CRXO trials, like other designs, will commonly investigate subgroup differences and may perform adjusted analyses (3). Sensitivity analyses may also be undertaken to examine whether the results are robust to the assumptions of the analysis. These analyses address the same question (i.e. estimand), but make different (plausible) assumptions that have the potential to lead to different results (84), for example, sensitivity of the treatment effect and its standard error to the choice of intra-cluster correlation structure. Details of such sensitivity analyses should be reported.

Example of item 12b

“A post-hoc series of analyses were undertaken to gauge the sensitivity of the point estimate and confidence interval from the primary analysis of in-hospital mortality to 90 days to alternative analysis models and small-sample corrections to the robust/sandwich variance estimator, and to

the use of degrees of freedom from the t-distribution.” (Supplement 2 of Peptic Investigators (45))

Results: Participant flow (a diagram is strongly recommended)

Item 13a: Participant flow

The numbers of clusters that were assessed for eligibility and were randomly assigned to each sequence. For each sequence-period (i.e. each cell) or treatment condition:

- the numbers of clusters that received intended treatments and were analysed for the primary outcome; and,
- the numbers of participants who were assessed for eligibility, received intended treatments and were analysed for the primary outcome.

Item 13b: Participant attrition

For each sequence-period (i.e. each cell) or treatment condition, losses and exclusions for both clusters and participants with reasons.

Explanation 13a and 13b: Providing information on the flow of clusters and participants at each stage of a randomised trial allows assessment of generalisability and the likelihood of selection and attrition bias. A flow diagram provides a succinct visual depiction of this information, which is particularly valuable for more complex trials (3).

For CRXO trials, the flow diagram can be structured in a number of ways (e.g. by each sequence-period or treatment condition). Different structures will be more or less suitable depending on the CRXO trial’s specific design features, such as the number of sequences and periods, and whether the same people, different people, or a mixture, are assessed in different periods. In some circumstances, presenting multiple flow diagrams using different structures may aid interpretation.

For a standard CRXO trial, the flow diagram is ideally structured by sequence-period, where *for each* sequence the cluster and participant level numbers are reported for each period. Such a structure allows assessment of differential recruitment/identification and attrition across treatments *and* across periods. Structuring the flow diagram by only treatment condition could be considered for more complex cross-sectional designs (e.g. > 2 periods, treatments or sequences) where a flow diagram structured by sequence-period may be unwieldy. However, structuring by only treatment condition has the disadvantage of not allowing assessment of differential inclusion and attrition across treatments and periods. Furthermore, this structure is problematic for cohort designs, since it is likely to over-represent the numbers of individual level participants as they contribute measurements under multiple treatment conditions.

Example of item 13 a and b

Example 1 (Flow diagram structured by sequence-period cross-sectional design) – Supplementary materials 3, Figure S2.

Example 2 (Flow diagram structured by treatment condition) – Supplementary materials 3, Figure S3.

Results: Recruitment

Item 14a: Recruitment

Dates of treatment periods and washout periods.

Explanation: Providing the dates of the treatment periods and washout periods is important for providing historical context for the trial (in particular, the treatments evaluated) (3) and for assessing

trial that was open to all EDs). This is necessary information for generalising the findings (item 21). In addition, reporting of baseline characteristics permits assessment of the balance of these characteristics across the randomised sequences, and in trials with post-randomisation recruitment of participants, can allow assessment of potential recruitment/identification bias and potential changes to the case-mix of participants over time.

In multiple-period trials, the concept of 'baseline' is more complex because of the longitudinal nature of the design (52). Here, the CONSORT stepped-wedge terminology is adopted where a 'baseline' characteristic is considered one that is measured before exposure to the treatment condition or one that is not expected to be influenced by the treatment (e.g. age). Therefore, 'baseline' characteristics may apply differently to cluster- and individual-level characteristics. 'Baseline' cluster characteristics can usually be measured prior to randomisation (and are often time-invariant), while 'baseline' individual characteristics may be assessed when the participant is recruited/identified (which may be post randomisation), but before they have been exposed to a treatment condition.

The baseline characteristics table can be structured in a number of ways, with different structures being more or less suitable depending on the CRXO trial's specific design features (e.g. number of sequences and periods; whether the same people, different people, or a mixture, are included in different periods), and the type of characteristics to be summarised (i.e. cluster or individual level). Different structures have advantages and disadvantages, so presenting multiple may aid in interpretation. Regardless of the structure of the table, the statistics reported should be descriptive (e.g. means and standard deviations), and not inferential (e.g. standard errors) (3).

For a standard CRXO cross-sectional trial, tabulating baseline individual level characteristics for each sequence-period allows assessment of the balance of these characteristics across the randomised sequences, the potential for recruitment/identification bias, and any systematic change in characteristics across the periods (12). For time-invariant cluster level characteristics, tabulating by sequence will be sufficient for assessing the balance of these characteristics across the randomised sequences. Tabulation by treatment condition could be considered for more complex cross-sectional designs (e.g. > 2 periods, treatments or sequences), where presentation by sequence-period may be unwieldy. Although this structure does not reflect the design, it still may be useful for identifying selective recruitment/identification of participants.

Example of item 15

Example 1 (Baseline table by sequence-period) – Supplementary materials 3, Table S2.

Example 2 (Baseline table by treatment condition) – Supplementary materials 3, Table S3.

Results: Numbers analysed

Item 16: Numbers analysed

The number of observations and clusters included in each analysis for each treatment condition and whether the analysis was according to the allocated schedule.

Explanation: The number of clusters and observations by treatment condition should be reported for analyses of all outcomes. For the primary outcome(s), this information will be attainable from the flow diagram of clusters and participants through the trial (item 13a). For analyses of secondary outcomes, reference may be made to the flow diagram when the same number of clusters and participants are included. For secondary outcomes where not all clusters or participants contribute

to each analysis, in addition to providing the number of clusters and observations by treatment condition, it can be useful to also provide these numbers by sequence (e.g. Tables 5(A) and 5(B)), or ideally, as a flow chart as per the primary outcome (i.e. structured by sequence-period, where *for each* sequence the cluster and participant numbers are reported for each period).

The type of summary information may differ depending on whether the same participants, different participants, or a mixture, are assessed in different periods. For example, where the same participants contribute repeated measurements across the periods, they will have been exposed to both treatment conditions; therefore, the total number of observations can be presented by treatment condition, or as the number of participants in the study with the average number of observations per participant under each treatment condition (52).

Providing information on whether the analysis was according to the allocated schedule, or if not, how deviations were dealt with in the analysis, is recommended. A detailed consideration of this issue is given under item 12a (Statistical methods).

Examples of item 16

Example 1 (Numbers by treatment condition) – “Of 38 assessed ICUs, eight fulfilled all eligibility criteria 10 980 patients were admitted during the study period: 2204 during baseline, 4069 during cycling [treatment condition 1], and 4707 during mixing [treatment condition 2]. 1598 (18.2%) patients were present during the monthly point-prevalence surveys (745 during cycling and 853 during mixing), and were therefore included in the main analysis”(85)

Example 2 (Table with the number of clusters and participants reported for each period within each sequence; noted analysis based on intention-to-treat) – Supplementary materials 3, Table S4.

Results: Outcomes and estimation

Item 17a: Outcomes and estimation

For each primary and secondary outcome, summary statistics by sequence-period (i.e. each cell) or treatment condition; the estimated effect size and its precision (e.g. 95% confidence interval); and any within-cluster correlations or variance components estimated in the analysis.

Explanation: For each primary and secondary outcome, reporting summary statistics provides necessary information for assessing the magnitude or prevalence of the outcome in the trial participants, and the potential for any ceiling or floor effects. For a standard CRXO trial, presenting summary statistics by sequence-period can allow assessment of the likelihood of any period effects, and allows readers to judge the appropriateness of the analysis. Presentation of summary statistics by treatment condition may be necessary for more complex designs (e.g. > 2 periods, treatments or sequences), where presentation by sequence-period may be unwieldy. Presentation of summary statistics by treatment condition is a natural approach; however, a simple comparison of summary statistics by treatment could lead to a biased estimate of treatment effect when period effects exist.

Estimated treatment effects with confidence intervals should be reported. P values may be provided, but these should be in addition to effect estimates and confidence intervals, and should be reported as exact values (i.e. not based on a significance threshold, for example, $P = 0.032$, not $P < 0.05$). The completeness of reporting of outcomes has been shown to be associated with statistical significance, where statistically significant outcomes are more often completely reported than non-significant outcomes, thus distorting the evidence base about the effects of treatments (86).

Reporting of estimates of correlations arising from the CRXO design is critical for planning future trials (5, 52). For a CRXO trial with two periods, this will include estimates of the within-period intra-cluster correlation coefficient (ICC) and the between-period intra-cluster correlation coefficient (or cluster autocorrelation coefficient). With more than two periods, the parameter estimates of the assumed within-cluster correlation structure over time should be reported. Given the number of possible correlations, explicit reporting of the type of correlation alongside the estimate is necessary. Variance components may be reported as an alternative to ICCs, particularly for non-continuous outcomes (87). When reporting ICCs for binary outcomes, indicating the scale (e.g. proportions or logistic) is necessary to know whether the ICC can be used in the variance inflation formula (88).

Example of item 17a

Table with summary statistics (percentages) reported for each period within each sequence; estimated effect sizes (with 95% CIs) – Supplementary materials 3, Table S4.

Table with within-period and between-period intra-cluster correlation coefficients – Supplementary materials 3, Table S5.

(Summary statistics by treatment condition) – “The number of patients with at least one drug-related problem in the intervention and control groups was 236 (44.0%) and 280 (50.6%) respectively (OR 0.77, 95% CI 0.61–0.98). The intervention reduced the frequency of prescription and/or dispensing errors, patient errors and treatment gaps (OR 0.52, 95% CI 0.29–0.93; 0.84, 0.66–1.07; and 0.65, 0.43–0.99, respectively; [Supplementary materials 3, Table S4]). Within-period and between-period intra-cluster correlation coefficients are reported in [Supplementary materials 3, Table S5].” (33)

Item 17b: Binary outcomes

For binary outcomes, presentation of both absolute and relative effect sizes is recommended.

Explanation: Reporting of both absolute (e.g. risk difference) and relative effect sizes (e.g. risk ratio, odds ratio) for each binary outcome is recommended so as to allow readers to fully interpret the impact of the intervention. Relative effect measures are generally more stable across different levels of baseline risk (89), and so have the advantage of being more generalisable to populations or settings where the baseline risk may differ. However, presentation of only relative effects can lead to over-statement and over-interpretation of the intervention effect when the outcome is rare (90). For this reason, reporting of absolute effect measures is also recommended (3).

Example of item 17b

Example 1 (Risk ratio and risk difference) – “A total of 2459 of 13 415 patients (18.3%) in the proton pump inhibitor group died at the hospital by day 90 and 2333 of 13 356 patients (17.5%) in the histamine-2 receptor blocker group died at the hospital by day 90 (RR, 1.05 [95% CI, 1.00 to 1.10]; absolute risk difference, 0.93 percentage points [95% CI, –0.01 to 1.88 percentage points]; P = .054).” (45)

Example 2 (Table with odds ratios and risk differences presented for each binary outcome) – Supplementary materials 3, Table S4.(33)

Results: Ancillary analyses

Item 18: Ancillary analyses

Results of any other analyses performed, including subgroup analyses, sensitivity analyses, and adjusted analyses, distinguishing pre-specified from exploratory.

Explanation: Readers are referred to the CONSORT statement for discussion of subgroup analyses and adjusted analyses (3). Reporting results from any sensitivity analyses undertaken is important for allowing readers to assess how departures from assumptions of the primary analyses may impact the interpretation of the results. Pre-specified sensitivity analyses should be distinguished from those that were not planned (71). At a minimum, reporting the estimated effect size and its precision from each sensitivity analysis is recommended; however, other relevant parameter estimates may be reported depending on the type of sensitivity analysis (e.g. reporting estimates of correlations if a different intra-cluster correlation structure is assumed). Sensitivity analyses may be best presented in a table that includes the results from the primary analyses, sensitivity analyses, and details of the assumptions that underlie the primary and sensitivity analyses.

Example of item 18

“Post-hoc sensitivity analyses for the primary outcome variable

[Supplementary materials 3, Table S6] provides the point estimate (RR) and 95% CI for a range of different analysis models for the primary endpoint, with small sample adjustments to the robust/sandwich variance estimator and the use of t-distribution with 32 degrees of freedom. Overall, there was no variation of the point estimate (RR) across approaches, and minimal variation in the 95% CI endpoints (maximum difference of 0.01 from the primary analysis result).” (Supplement 2 of Peptic Investigators (45))

Results: Harms

Item 19: Harms

Important harms or unintended effects in each treatment condition (for specific guidance see CONSORT Harms 2022 statement).

Explanation: Readers are referred to the CONSORT 2010 statement and the CONSORT Harms 2022 statement for examples and explanation (3, 7). However, for open or closed cohort CRXO designs where participants are followed over time, particular care is required in interpreting adverse events data because of the risk of carryover effects. The effects of any adverse events experienced in the first period may persist into the second (or later) periods, or, the effects of the treatment condition received in the first period may persist into later periods. Therefore, adverse events experienced in later periods cannot necessarily be attributed to the treatment condition the participant is receiving at the time they experience the event.

Discussion

Item 20: Limitations

Trial limitations, addressing sources of potential bias, imprecision, and if relevant, multiplicity of analyses. Consider potential carry-over effects.

Explanation: Reporting of any limitations is important since this allows readers to assess the potential for bias in the results. A key limitation of a CRXO design is the potential for carryover whereby the effects of the treatment in one period persist into a subsequent period, thus potentially

biasing estimates of the treatment effects. Other possible CRXO design limitations are outlined in item 3a, and should be reflected upon when they apply.

Possible limitations common to all randomised trial designs include imprecise estimates of treatment effect (e.g. for secondary outcomes) and the risk of spurious statistically significant findings arising from undertaking multiple statistical tests (e.g. for different outcomes, subgroup analyses). Interpretation and reporting of the results cognisant of these limitations is recommended. Readers are referred to the CONSORT 2010 statement for further discussion of imprecision and multiplicity of results (3).

Example of item 20

“Carryover effects are often plausible in individual-patient crossover trials where patients receive both study treatments. However, such carryover effects are less plausible in our cluster randomized crossover study because (i) different patients were present in the different observation periods and were not allocated to both treatments, and (ii) retention of ICU [Intensive Care Unit]-level practices from the previous period on the basis of better or worse outcomes than expected is unlikely both because details of overall patient outcomes within the prior treatment period are not readily available to staff and because random variability renders it implausible that true effects would be [*sic*] distinguishable from random variation at a single ICU level.” (Supplement 2 of Peptic Investigators (45))

Other information

Item 24: Trial protocol

Where the full trial protocol and statistical analysis plan can be accessed, if available.

Explanation: Readers are referred to the CONSORT 2010 statement for examples and explanation (3). Note however, that the CONSORT 2010 item was limited to only trial protocols. Statistical analysis plans (SAPs), which provide technical details of the analysis (to allow for reproduction of the results), are now standard practice (91). As with trial protocols, authors should ensure their SAP (and any amendments) are accessible to users. There are a number of mechanisms to achieve this; for example, publishing SAPs in journals (e.g. *Trials*) and referencing the SAP in the trial publication; including the SAP as a supplementary file to the trial publication; uploading the SAP to the clinical trial registry entry; or, depositing the SAP to a general purpose or institutional open access repository (e.g. Open Science Framework).

Item 26: Research ethics review

Whether the study was approved by a research ethics committee, with identification of the review committee(s).

Explanation: The International Committee of Medical Journal Editors (ICMJE) requires that authors include a statement in the report about research ethics approval by an independent local, regional, or national review body (68). The CONSORT 2010 statement did not include an item on research ethics approval. However, extensions to the CONSORT statement for pilot studies (92) and stepped-wedge trials (52) have introduced such an item. The relevant ethics committee(s) should be clearly identified, preferably together with the application or reference number. If a relevant research ethics committee granted an exemption from review, this should be reported together with the justification for the exemption.

Example of item 26

“The National Healthcare Group Domain Specific Review Board approved the study, with waiver of informed consent from patients (DSRB number: 2017/00441). (29)

Item 27: Data sharing

Where the individual de-identified participant data (including data dictionary), statistical code, and any other relevant documents/materials can be accessed.

Explanation: The call for data sharing for researchers, including clinical trialists, is now loud, as part of a larger ecosystem movement towards open science (e.g., UNESCO (93)). Very recently the US National Institutes of Health have indicated that data sharing will become a normative practice for their grantees (94). In Canada there is now a new mandate by the three federal funders requiring data sharing by their grantees as of 2022 (95). In Canada, data sharing will start by requiring grant applicants to describe their data management plans (DMPs) in their grant application, followed by actual data sharing at a later date. Similar data sharing requirements are elsewhere (96). While this is positive news, clinical trial data sharing is not yet a normative practice.

In the context of clinical trials, including CRXO trials, data sharing means that once the trial is finished and a report is made available (e.g., preprint server; publication, preferably open access) the reporting includes sharing relevant trial information to facilitate bringing best evidence to best practice, to enhance and facilitate evidence syntheses (97) and to help foster meta-research. This typically includes the trial protocol, statistical analysis plan or both (if not already publicly available) (see item 24), data, the analytical code that was used to arrive at the reported trial results, data dictionaries and other material(s). Examples of the latter could include intervention materials; for example, in a trial examining the effects of freshly prepared school lunches (intervention condition) versus usual packed lunches from home (control condition) the authors would share details of the menu schedule, recipes, the food groups and nutrient content.

When it comes to sharing data, authors need to be aware of the axiom, as open as possible, as closed as necessary. How the data are shared might depend on funder policies and/or journal requirements. Data sharing recommendations and mandates are time consuming and require resources, which some authors may not be able to meet (98). Data sharing does not typically include sharing of signed patient consent forms. Authors should follow the FAIR (Findability, Accessibility, Interoperability, and Reuse of digital assets) principles when sharing clinical trial data (99).

There have been concerns about the reproducibility of research, including clinical trials, for at least the last decade (100). Unfortunately, there has been little investigation about the reproducibility of trial results in biomedicine (101). Clinical trial data reuse can only happen if trialists share their data. Even during COVID-19, trial data reuse did not happen (102) and none of the COVID-19 vaccine trialists shared the data underpinning their results.

Example of item 27

The publication of results for the *ORGANIKO LIFE + CRXO* trial included the following supporting information: study questionnaires, statistical analysis plan, data files (comma-separated values format), analysis script files (for the statistical package R) and the methods for measuring the biomarkers (27).

Item 28: Patient and public involvement

Details of any patient and/or public involvement in the design, conduct and reporting of the trial; and, when applicable, other stakeholders' involvement.

Explanation: Patient and public involvement (PPI) in research has been defined as research carried out ‘with’ or ‘by’ members of the public, rather than involvement as a participant in a randomised trial (103). In this definition, ‘public’ includes patients, potential patients, carers and people who use health and social care services, as well as organisations that represent people who use these services (103). In the context of randomised trials, PPI might include: input in trial oversight committees (e.g. trial steering committee); defining the research questions; outcomes to be measured; visit schedules; methods of data collection; recruitment and consent procedures; lay write-up of the trials; and advocacy of the trial findings (104, 105).

Many funding bodies encourage PPI in the clinical trials they fund (106, 107), and some journals are asking clinical trialists to explicitly report on the PPI in the trial (e.g. BMJ, BMJ Open (108)). Despite such requirements, reporting of PPI has been very rare (109). Non, incomplete and inconsistent reporting can limit understanding of PPI in a particular trial, and efforts to build a PPI evidence-base (e.g. cataloguing the types of PPI used in trials) (105).

The Guidance for Reporting Involvement of Patients and the Public 2 (GRIPP2) provides a checklist for reporting PPI in health and social care (105). The short-form version (GRIPP2-SF) should be used to guide reporting in randomised trials (where PPI involvement is not the primary focus of the trial). In some cluster trials (including CRXO trials), there may be involvement from other stakeholders (e.g. hospital administrative staff, health professionals) in addition to PPI (106). These stakeholders may be involved in various aspects of the trial (e.g. advising on the content of the intervention, recruitment procedures), and their involvement should also be reported.

Example of item 28

Example 1 (Patient participant involvement) – “The LUSTRUM [Limiting Undetected Sexually Transmitted infections to RedUce Morbidity] PPI (Patient and Public Involvement) Group consists of 27 lay people, with a broad mix of demographic characteristics and a range of sexual health experiences. The group were first involved in the design of the research at the grant application stage. Their opinions guided the trial protocol and topic guides for the process evaluation interviews. The trial is conducted by qualified healthcare professionals and, as such, it is not feasible for the group to undertake any aspects of the data collection. The group strongly supported the service-level consent design which reduces the burden of participating in the research. Lay summaries of our findings will be circulated to the group to gain their perspective and feedback on our results; this may inform further analytical decisions and inputs.” (110)

Example 2 (Other stakeholder involvement) – “PREP-IT [A Program of Randomized trials to Evaluate Pre-operative antiseptic skin solutions In orthopedic Trauma] stakeholders include patient advisors who experienced orthopedic trauma and/or surgery, and subsequent infection complications; professional-association representatives; a multidisciplinary healthcare team; and dissemination professionals. Stakeholders codeveloped the PREP-IT study concept and protocol following the 10-Step Framework.” (111)

Conclusion

The CONSORT CRXO extension consolidates reporting items across relevant guidelines, modifying items as required, and includes additional items. The modified and new items address unique features of the CRXO design, as well as reflecting evolving requirements of the scientific community (108, 112). For CRXO trials to provide the best opportunity for informing policy and clinical practice, either directly, or through evidence synthesis products, complete and accurate reporting is required. The CONSORT CRXO extension is a first step to facilitating this. Journals, editors and peer-reviewers can also play a crucial role through endorsement and uptake of the guideline in editorial processes.

Tables

Table 1. Glossary of terms

Term	Explanation
Carry-over effect	The effect of the treatment delivered in one period persisting into a subsequent period.
Cell	The combination of the sequence and period. See Sequence-period.
Cluster-period	The combination of a cluster and period. For example, a standard CRXO design has two cluster-periods for each cluster.
Cohort*	Participants are repeatedly assessed over a series of measurement points. An <i>open</i> cohort design is one in which participants can join and leave the trial throughout its duration. A <i>closed</i> cohort design is one in which participants cannot join the trial once it has started.
Cross-sectional*	Different participants are measured at each measurement occasion.
Estimand	A clear description of what the treatment effect being estimated represents. Attributes for completely specifying an estimand for cluster trials are outlined in Box 3.
Gatekeepers, cluster guardians	Gatekeepers are individuals or bodies (such as a school principal or municipal council) who may be called on to protect the interests of organisations or communities that are the setting for a cluster randomised trial. (61)
Dual pair sequences (duals)	A pair of sequences that are the inverse of one another (e.g. ABB and its dual BAA) (10). Designs consisting of sequences and their duals help to ensure a time-balanced design.
Multiple-period, or longitudinal, CRXO design	One in which clusters cross between treatments more than once (10)
Period*	A grouping of observations by time of measurement.
Period effect	Systematic differences in responses across periods that are not due to the treatments (25). Examples of period effects include seasonal or secular trends, and order and learning effects (see item 3a for specific examples).
Participant*	A participant is someone on whom investigators seek to measure the outcome of interest.
Research participant*	A research participant denotes a human research subject from the standpoint of ethical considerations.
Sequence (or sequence of treatments)**	Order in which a cluster will receive the treatment conditions over time. More than one cluster can be allocated to each sequence.
Sequence-period	The combination of the sequence and period. For example, a standard CRXO design has the four sequence-periods: <i>Sequence AB</i> 1) Period 1 – treatment condition A 2) Period 2 – treatment condition B <i>Sequence BA</i> 3) Period 1 – treatment condition B 4) Period 2 – treatment condition A
Standard CRXO cross-sectional design	A CRXO design with two-periods, two-treatments (A or B) and two-sequences (AB, BA), with different participants in each period.
Washout period	A length of time between treatment periods when no treatment is delivered to reduce the risk of carry-over effects.

* Definition used in, or with minor modification from, the CONSORT extension for stepped-wedge trials (11).

** Following the CONSORT extension for stepped wedge trials we avoid the term 'group' or 'arm' to refer to the allocated treatment as is used in the standard CONSORT statement. We instead use the term treatment in a generic way to refer to either the active treatment or comparator. We use the phrase 'intervention condition' to refer to the active treatment and the 'control condition' to the comparator.

Table 2. Checklist of information to include when reporting a CRXO trial

Section / Topic	Item no.	Checklist item
Title and Abstract		
	1a	Identification as a cluster randomised crossover trial in the title.
	1b	Structured summary of the trial design, methods, results, and conclusions (see separate CRXO checklist for abstracts)
Introduction		
Background and objectives	2a	Scientific background and explanation of rationale.
	2b	Specific objectives or hypotheses.
Methods		
Trial design	3a.1	Rationale for a cluster crossover design.
	3a.2	Description of the realised trial design including: <ul style="list-style-type: none"> • number of treatment conditions; • definition of cluster (i.e. the unit of randomisation); • number and duration of periods; • number and composition of sequences (e.g. ABAB, BABA); • number of clusters randomised to each sequence; • duration of any washout periods; • whether the participants assessed in different periods are the same people, different people, or a mixture of the two; and • consideration of potential for carryover effects. <p>A diagram of the trial is recommended when there are more than two periods and/or treatment conditions.</p>
	3b	Important changes to planned methods after trial commencement (such as eligibility criteria), with reasons.
Participants	4a	Eligibility criteria for clusters and participants.
	4b	Settings and locations where the data were collected.
Intervention	5	The treatment conditions with sufficient details to allow replication, and whether they were delivered at the level of the cluster, the individual, or both.
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed (for specific guidance see CONSORT for outcomes).
	6b	Any changes to trial outcomes after the trial commenced, with reasons.

Sample size	7a	How sample size was determined. Method of calculation and relevant parameters with sufficient detail so the calculation can be reproduced. Assumptions made about correlations between outcomes of participants from the same cluster (see separate CRXO checklist for sample size items).
	7b	When applicable, explanation of any interim analyses and stopping guidelines.
Randomisation:		
Schedule generation	8a	Method used to generate the random allocation schedule.
	8b	Type of randomisation; details of any restricted randomisation, if used.
Allocation concealment mechanism	9	Specification that allocation was based on clusters; description of any methods used to conceal the allocation from the clusters until after their recruitment.
Implementation	10a	Who generated the random allocation schedule, who enrolled clusters, and who assigned clusters to sequences of treatments in the schedule.
	10b	Mechanism by which individual participants were included in clusters for the purposes of the trial (such as complete enumeration or random sampling; continuous recruitment or ascertainment, or recruitment at a fixed point in time), including who recruited or identified participants.
	10c	Whether consent was sought, from whom, when and for what; whether this differed between treatment conditions. Justification for any waiver or modification of informed consent requirements.
Blinding	11a	Who was blinded after assignment to sequences (e.g. cluster level participants, individual level participants, those assessing outcomes) and how.
	11b	If relevant, description of the similarity of interventions.
Statistical methods	12a	Target estimand for each primary and secondary outcome including whether it pertains to the cluster-level or individual-level; statistical methods for their estimation including how period effects, clustering and repeated measures were taken into account. Any assessment of carryover effects should be reported.
	12b	Methods for additional analyses, such as subgroup analyses, sensitivity analyses, and adjusted analyses.
Results		
Participant flow (a diagram is strongly recommended)	13a	The numbers of clusters that were assessed for eligibility and were randomly assigned to each sequence. For each sequence-period (i.e. each cell) or treatment condition: <ul style="list-style-type: none"> the numbers of clusters that received intended treatments and were analysed for the primary outcome; and, the numbers of participants who were assessed for eligibility, received intended treatments and were analysed for the primary outcome.
	13b	For each sequence-period (i.e. each cell) or treatment condition, losses and exclusions for both clusters and participants with reasons.

Recruitment	14a	Dates of treatment periods and washout periods.
	14b	Why the trial ended or was stopped.
Baseline data	15	A table showing baseline cluster level characteristics by sequence, and individual level characteristics for each sequence-period (i.e. each cell) or treatment condition.
Numbers analysed	16	The number of observations and clusters included in each analysis for each treatment condition and whether the analysis was according to the allocated schedule.
Outcomes and estimation	17a	For each primary and secondary outcome, summary statistics by sequence-period (i.e. each cell) or treatment condition; the estimated effect size and its precision (e.g. 95% confidence interval); and any within-cluster correlations or variance components estimated in the analysis.
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended.
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses, sensitivity analyses, and adjusted analyses, distinguishing pre-specified from exploratory.
Harms	19	Important harms or unintended effects in each treatment condition (for specific guidance see CONSORT Harms 2022 statement).
Discussion		
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and if relevant, multiplicity of analyses. Consider potential carry-over effects.
Generalisability	21	Generalisability (external validity, applicability) of the trial findings. Generalisability to clusters or individual participants, or both (as relevant).
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence.
Other information		
Registration	23	Registration number and name of trial registry, or state the trial was not registered.
Protocol	24	Where the full trial protocol and statistical analysis plan can be accessed, if available.
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders.
Research ethics review	26	Whether the study was approved by a research ethics committee, with identification of the review committee(s).
Data sharing	27	Where the individual de-identified participant data (including data dictionary), statistical code, and any other relevant documents/materials can be accessed.
Patient and public involvement	28	Details of any patient and/or public involvement in the design, conduct and reporting of the trial; and, when applicable, other stakeholders' involvement.

This table can be downloaded as a separate document in Supplementary materials 4, where section and page numbers can be inserted.

Table 3. Items to report in conference and journal abstracts of a cluster randomised crossover (CRXO) trial

Abstract item	Extension for CRXO trials
Title	Identification of study as a CRXO trial
Trial design	Description of the trial design (including number of sequences, periods and clusters; and whether participants assessed in different periods are the same people, different people, or a mixture)
Methods:	
Participants	Eligibility criteria for clusters and participants and the settings where the data were collected
Interventions	The treatment conditions
Objective	Specific objectives or hypotheses
Outcome	Clearly defined primary outcome
Randomisation	How clusters were allocated to sequences of treatments
Blinding (masking)	Whether or not participants, healthcare professionals, those recruiting and those assessing outcomes were blinded
Results:	
Numbers randomised	Number of clusters randomised to each sequence of treatments
Recruitment	Trial status
Numbers analysed	Number of clusters included in the analysis. Number of observations in the analysis for each treatment condition.
Outcome	For the primary outcome, summary statistics by treatment condition and the estimated effect size (confidence interval)
Harms	Important adverse events or side effects (even if not observed), or that harms were not assessed
Conclusions	General interpretation of the results
Trial registration	Registration number and name of trial register
Funding	Source of funding
Research ethics review	Ethical approvals

Table 4. Core items to report under the sample size calculation (item 7a)*

Recommended items to report	Further explanation
Outcome(s)	Outcome(s) for which the sample size calculation was undertaken.
Target difference	For a superiority trial, the magnitude of difference in the outcome of interest (“effect size”) that the trial is designed to detect (see Power), along with the basis for this choice. Specify the target difference as an absolute or relative effect, or both (e.g. risk difference, risk ratio) (38). For equivalence and non-inferiority trials, as above, but where the target difference is replaced by the equivalence or non-inferiority margin.
Level of significance (probability of a type I error)	State clearly whether a one or two-sided test was used.
Power	The probability of detecting the target difference (or larger) as

	statistically significant, given such a difference exists.
A measure of variation for the outcome	For continuous outcomes, a standard deviation is required; for binary outcomes, the control condition proportion should be specified. For other outcomes see Cook <i>et al</i> (38).
Number of clusters and allocation ratio	The total number of clusters and the allocation ratio of the sequences (i.e. the proportion of clusters allocated to each sequence).
Number of periods	State the number of periods.
Average cluster size	The total cluster size or the cluster size per period, clearly denoting which is reported.
The assumed correlation structure	The assumed intra-cluster correlation coefficient (ICC) and whether the ICC is time dependent or time independent. If time dependent, state the parameters that were assumed to accommodate the time dependency, for example, the within-period ICC and the between-period ICC or the cluster autocorrelation coefficient, or any variance components. For binary outcomes it is important to report the scale of the correlations or variance components (e.g. proportions scale or logistic scale). For rate outcomes it might be more appropriate to report coefficients of variation of outcomes. (11)
Within person correlations	Where the design includes repeated measurements on the same individual, describe the assumed correlation structure at the individual level (e.g. an individual autocorrelation coefficient), including if any decay in correlation in repeated measures on the same individual has been accounted for.
Method used	Description of the sample size methodology used (e.g. via provision of a formula, or results and code for simulation-based sample size calculations), including the statistical packages (and functions) used for implementation.
Any adjustments made	
Allowance for multiple primary outcomes	Any allowance made for multiple primary outcomes (e.g. by applying a correction to the level of significance in the sample size calculation (113)).
Allowance for variation in cluster size	Any allowance made for unequal cluster sizes. This can include variation in total cluster sizes or variation in cluster-period sizes.
Allowance for attrition	Any inflation due to anticipated attrition at the cluster or participant levels, or both.
Small sample adjustment	Any corrections applied when the number of clusters are small. Sample size calculations for CRXO trials based on the normal distribution are only accurate when the number of clusters is large. Use of these formulas when the derived number of clusters is small may underestimate the required number of clusters, resulting in a

	trial with less power than planned. A range of small sample size adjustments might be applied, such as adding clusters or using values for the <i>t</i> -distribution in place of the normal distribution in the sample size calculation (9) (pg 149) (39, 114).
Sensitivity analysis	Any analyses undertaken to examine the sensitivity of the sample size, or power, to the chosen parameters (e.g. ICCs, control condition proportion). Justification for the chosen parameters should be provided.

* Adapted (where applicable) from Table 5 of the CONSORT extension for stepped-wedge cluster randomised trials (11) and informed by the DELTA² guidance on choosing the target difference (38).

Tables 5(A) and 5(B) offer two possible table structures for tabulating the number of clusters and observations included in each analysis per outcome by sequence for a standard CRXO trial (two-periods, two-treatments and two-sequences (AB, BA)). Table 5(B) provides more complete information because it separates the number of clusters that contribute to only one period by sequence. Note that when there is no withdrawal of clusters, the columns marked A only or B only are not required.

Table 5(A)

	Number of clusters (number of observations)			
Outcome	AB clusters	BA clusters	A only clusters	B only clusters
X	8 (...)	8 (...)	2 (...)	2 (...)
Y				
...				

Table 5(B)

	Number of clusters (number of observations)					
	Sequence AB			Sequence BA		
Outcome	AB clusters	A only clusters	B only clusters	BA clusters	B only clusters	A only clusters
X	8 (...)	2 (...)	0 (...)	8 (...)	2 (...)	0 (...)
Y						
...						

Boxes

Box 1. Example of a CRXO trial: The PEPTIC CRXO trial (45)

Primary aim: To compare Proton Pump Inhibitors (PPIs) versus Histamine-2 Receptor Blockers (H₂RBs) for stress ulcer prophylaxis on in-hospital mortality among adults requiring mechanical ventilation.

Primary outcome: In-hospital all-cause mortality up to 90 days from the date of admission.

Treatment conditions: PPIs versus H₂RBs.

Clusters: Intensive Care Units (ICUs).

Participants: Adult patients requiring invasive ventilation within 24 hours of ICU admission.

Design: Two-period two-sequence cross-sectional design. 50 ICUs (across 5 countries) with 25 ICUs randomised to use PPIs in the first period, then crossover to use H₂RBs in the second period; and, 25 ICUs randomised to use H₂RBs in the first period, then crossover to use PPIs in the second period. The trial was pragmatic in its intent to inform clinical decision-making, and as such, while the default prescription for stress ulcer prophylaxis was determined by the ICU's treatment allocation in a particular period, clinicians were able to use the alternative treatment if they considered this preferable. Periods were of length 6 months. There was no washout between periods.

Box 2. Distinguishing the terminology of 'sequence' and 'schedule' and their use in relation to random allocation of units in parallel group and CRXO trials

Randomisation unit (i.e. participant or cluster)	Parallel-group randomised trial (2 treatments, 1 period)	CRXO trial (2 treatments, 2 periods)
	<i>Allocation sequence</i>	<i>Allocation schedule</i>
1	A	AB
2	B	BA
3	A	BA
4	A	AB
5	B	AB
6	B	BA

Box 3. Estimands in cluster randomised trials

Concept	Explanation
Estimand	Randomised trials are used to answer questions about how well an intervention works. However, there is not one single type of question of interest, and different trials might try to answer different types of questions; for instance about the effect of the treatment if all participants adhered; the effect in those patients who would adhere; or the effect of assignment to treatment irrespective of whether they adhered not. The exact question being addressed is described by the <i>estimand</i> , which provides a clear description of what the treatment effect being estimated represents. It precisely describes what the numerical result (the

	<p>estimate) will mean (e.g. whether a relative risk of 0.80 implies a 20% relative reduction regardless of whether participants adhered, or only if all patients were to adhere) (115).</p> <p>For cluster randomised trials, estimands are comprised of the following attributes: (i) population; (ii) treatment conditions; (iii) endpoint; (iv) summary measure (e.g. marginal risk ratio); (v) strategies used to handle intercurrent events (such as patient- or cluster-level non-adherence); and (vi) whether a participant-average treatment effect or a cluster-average treatment effect is to be used.</p> <p>Certain estimand attributes link to other items in the checklist; for example, the estimand ‘population’ links to item 4a (the trial population) and 12a (which includes the analysis population), and the estimand’s ‘treatment conditions’ link to item 5 (intervention). However, the estimand attribute differs to these other items in that it pertains to the trial’s treatment effect (e.g. which participants the treatment effect applies to; which treatments are being compared for the treatment effect) rather than the specific study methods (e.g. which participants were eligible to be enrolled, which participants were included in the analysis; which treatments clusters are randomised to).</p>
Estimator	The method of analysis used to calculate the estimated treatment effect.
Estimate	The numerical value computed by the estimator using the trial outcome data (e.g. in a cluster crossover trial reporting an odds ratio of 1.51, the value 1.51 represents the ‘estimate’).
Example of a participant-average estimand (adapted from the PEPTIC CRXO trial (45))	The marginal, participant-average risk ratio [summary measure] for in-hospital all-cause mortality up to 90 days from the index ICU admission [endpoint] in patients aged 18 years or older requiring invasive mechanical ventilation within 24 hours of admission to ICU, without upper gastrointestinal bleeding [population], between a strategy of proton pump inhibitor vs. histamine-2 receptor Blocker given until death/discharge/development of a clinically important upper gastrointestinal bleeding event, or until no longer indicated [treatment conditions], regardless of non-receipt of treatment/discontinuation/or receipt of alternate treatments (i.e. treatment policy strategy for handling intercurrent events) [intercurrent events].
Example of a cluster-average estimand (adapted from the BETR CRXO trial (116))	The marginal, cluster-average difference in infection rates [summary measure] for incident cases of <i>C difficile</i> infection during hospital admission (exposure period) [endpoint] among patients placed in a previously contaminated hospital room [population], between standard terminal disinfection (bleach) and enhanced terminal disinfection (bleach and ultraviolet light) [treatment conditions], regardless of whether there was documented use of the ultraviolet device [intercurrent event].

Figures

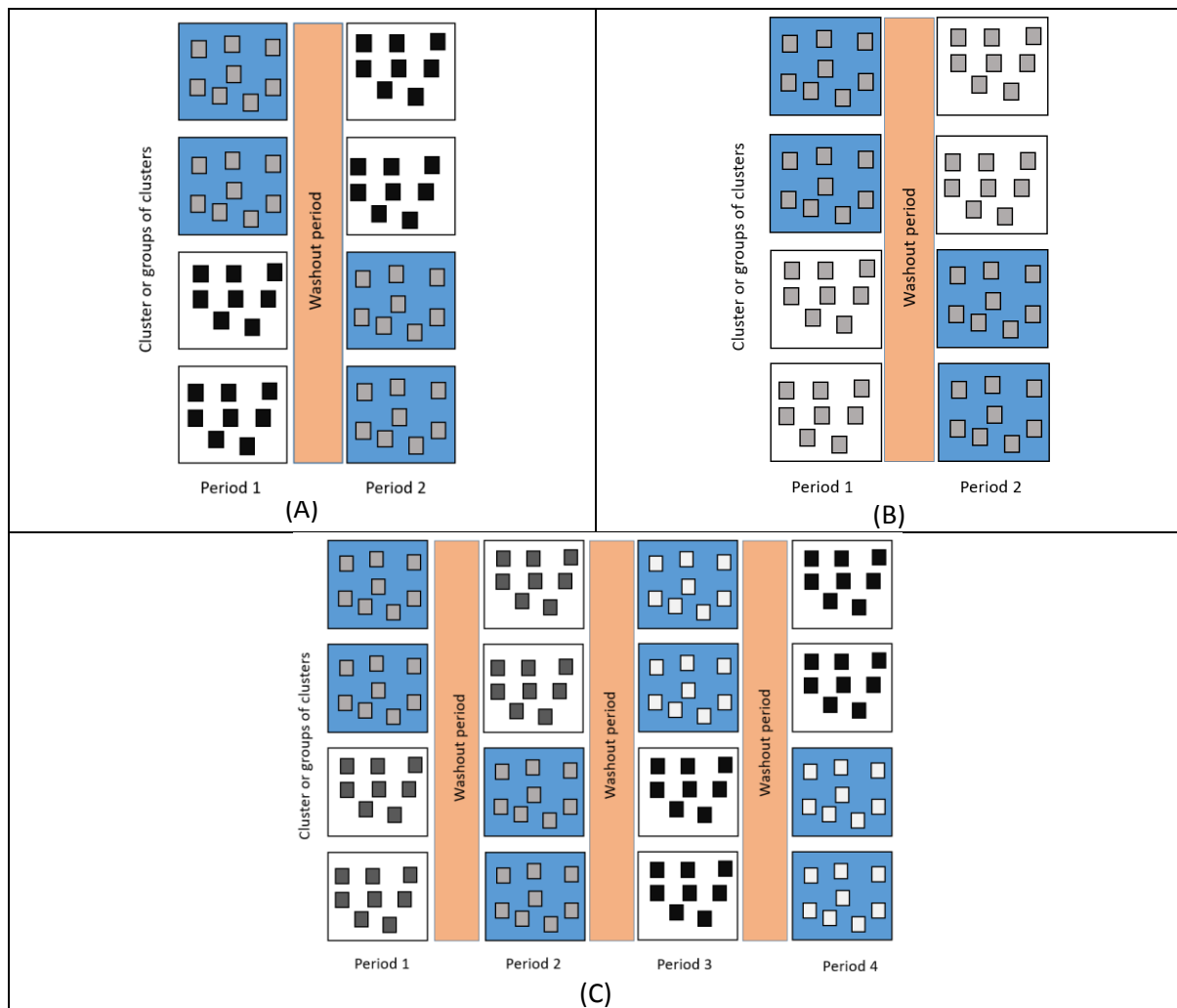


Figure 1. Depicts three different CRXO designs. The large boxes represent clusters (or groups of clusters), with different colours representing different treatments. The internal boxes represent participants, with changes in grey shading from period to period representing different participants across periods (i.e. cross-sectional design).

(A) standard CRXO cross-sectional design (indicated by different grey shading of internal boxes across periods) with two-treatments, two-periods, two-sequences (AB, BA).

(B) standard CRXO cohort design (indicated by the same grey shading of internal boxes across periods) with two-treatments, two-periods, two-sequences (AB, BA).

(C) CRXO cross-sectional design with two-treatments, four-periods, two-sequences (ABAB, BABA).

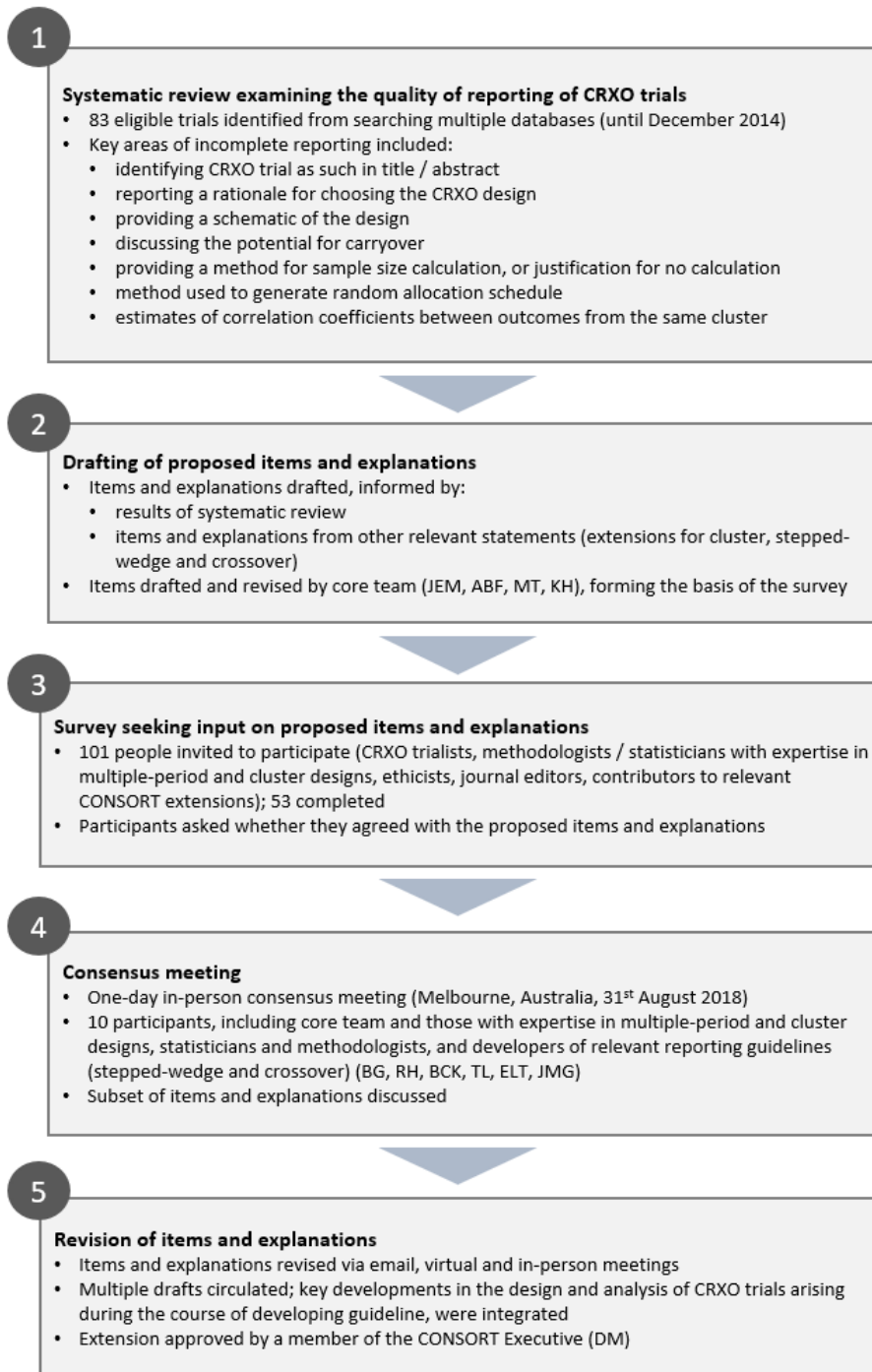


Figure 2. Summary of the five-step process used to develop the CONSORT CRXO extension. Further details are available in Supplementary materials 1.

Declarations

Acknowledgements

We thank the following participants for completing the survey (listed in alphabetical order):

Adrian Aldcroft, Craig Anderson, Deborah Ashby, Giraudeau Bruno, Marion Campbell, Mike Campbell, Leonid Churilov, Andrew Copas, Catherine Crespi, Katie Dainty, Stephanie Dixon, Kerry Dwan, Diana Elbourne, Sandra Eldridge, Gene Feder, Alan Girling, Cory Goldstein, Michael Grayling, Jeremy Grimshaw, Martin Gulliford, James Hargreaves, Karla Hemming, Stephane Heritier, Richard Hooper, Jim Hughes, Brennan Kahan, Mona Kanaan, Jessica Kasza, Neil Klar, Caroline Kristunas, Gillian Lancaster, Shun Fu Lee, Tianjing Li, Graeme MacLennan, James Martin, John Matthews, David Moher, Katy Morgan, David Osrin, Rupert Pearse, John Preisser, Stephen Senn, Monica Taljaard, Steven Teerenstra, Lehana Thabane, Jennifer Thompson, David Torgerson, Rebecca Turner, Elizabeth Turner, Obi Ukoumunne, Rieke van de Graaff, Rebecca Walwyn, and Charles Weijer. We thank Matthew J Page (Monash University) for taking notes at the consensus meeting.

Consensus meeting participants (listed in alphabetical order):

Andrew Forbes, Bruno Giraudeau, Jeremy Grimshaw, Karla Hemming, Richard Hooper, Brennan Kahan, Tianjing Li, Joanne McKenzie, Matthew Page (scribe), Monica Taljaard, Elizabeth Turner

The figure in Item 14a is reprinted from *The Lancet Infectious Diseases*, Vol. 19, Marjolein F Q Kluytmans-van den Bergh, Patricia C J Bruijning-Verhagen, Christina M J E Vandenbroucke-Grauls, Els I G B de Brauwier, Anton G M Buiting, Bram M Diederer, Erika P M van Elzakker, Alex W Friedrich, Joost Hopman, Nashwan al Naiemi, John W A Rossen et al., Contact precautions in single-bed or multiple-bed rooms for patients with extended-spectrum β -lactamase-producing Enterobacteriaceae in Dutch hospitals: a cluster-randomised, crossover, non-inferiority study, Pages 1069-1079, Copyright (2019), with permission from Elsevier.

Contributors

JEM and ABF are joint lead authors. *Survey*: JEM, MT, KH, JMG, ABF sought ethics approval. JEM drafted proposed items and explanations, with input first from ABF, then MT and KH. JEM administered the survey and analysed the data. *Consensus meeting*: JEM prepared the materials with input from ABF, MT and KH. JEM, ABF and BG gave background presentations. JEM, ABF, KH, MT presented proposed modifications to items and explanations. JMG chaired the meeting. All authors except SJA, DM, and SE attended. JEM and Matthew J Page took notes. JEM drafted the meeting summary. *Drafting the extension*: JEM revised the items and explanations with input first from ABF, then MT and KH. JEM led the drafting and editing of the article. All authors were involved in revising the article critically for important intellectual content. All authors approved the final version of the article. JEM is the guarantor of this work. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Data sharing statement

The survey data produced for the present study are available upon reasonable request to the authors (with ethics approval).

Transparency statement

JEM and ABF affirm that the manuscript is an honest, accurate, and transparent account of the development of the CONSORT extension for CRXO trials; and that no important aspects have been omitted.

Details of funding

The consensus meeting was funded through an Australian National Health and Medical Research Council (NHMRC) project grant (GNT1108283). JEM was supported by an NHMRC Career Development Fellowship Level 1 (GNT1143429) and an NHMRC Investigator Grant Leadership 1 (GNT2009612) during the conduct of this research. SJA was supported by the Research Support Package of JEM's NHMRC Investigator Grant from 2022. BCK is funded by the UK Medical Research Council (grants MC_UU_00004/07 and MC_UU_00004/09). TL was supported by grant UG1 EY020522 from the National Eye Institute, National Institutes of Health. JMG held a Tier 1 Canada Research Chair in Health Knowledge Transfer and Uptake (until December 2022). The funders had no role in considering the study design or in the collection, analysis, interpretation of data, writing of the report, or decision to submit the article for publication.

Details of ethics approval

Ethics approval was obtained from the Monash University Human Research Ethics Committee (approval number 13429). We sought consent from survey participants to take part in the online survey and for inclusion of their name in the list of participants acknowledged for contributing expert knowledge in papers, presentations and meetings about the study.

Patient and public involvement

Patients and public were not involved in the development of this reporting guideline. We have included an item that recommends reporting details of any patient and/or public involvement in the design, conduct and reporting of the CRXO trial. We refer users of this reporting guideline to the short-form version of the Guidance for Reporting Involvement of Patients and the Public 2 (GRIPP2); which was developed in collaboration with patient partners.

Competing interests

All authors have completed the Unified Competing Interest form (available on request from the corresponding author) and declare: support for the consensus meeting and for some authors from research funding organisations (see Details of funding); no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Provenance and peer review

Not commissioned; externally peer reviewed.

Licence for Publication

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd to permit this article (if accepted) to be published in BMJ and any other BMJPG products and sublicences such use and exploit all subsidiary rights, as set out in our licence (<http://group.bmj.com/products/journals/instructions-for-authors/licence-forms>).

References

1. Goodman SN, Fanelli D, Ioannidis JP. What does research reproducibility mean? *Sci Transl Med*. 2016;8(341):341ps12.
2. Begg C, Cho M, Eastwood S, Horton R, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA*. 1996;276(8):637-9.
3. Moher D, Hopewell S, Schulz KF, Montori V, et al. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *Bmj*. 2010;340(mar23_1):c869.
4. Dwan K, Li T, Altman DG, Elbourne D. CONSORT 2010 statement: extension to randomised crossover trials. *Bmj*. 2019;366:l4378.
5. Campbell MK, Piaggio G, Elbourne DR, Altman DG. Consort 2010 statement: extension to cluster randomised trials. *Bmj*. 2012;345:e5661.
6. Montgomery P, Grant S, Mayo-Wilson E, Macdonald G, et al. Reporting randomised trials of social and psychological interventions: the CONSORT-SPI 2018 Extension. *Trials*. 2018;19(1):407.
7. Junqueira DR, Zorzela L, Golder S, Loke Y, et al. CONSORT Harms 2022 statement, explanation, and elaboration: updated guideline for the reporting of harms in randomised trials. *Bmj*. 2023;381:e073725.
8. Arnup SJ, Forbes AB, Kahan BC, Morgan KE, et al. Appropriate statistical methods were infrequently used in cluster-randomized crossover trials. *Journal of clinical epidemiology*. 2016;74:40-50.
9. Eldridge S, Kerry SM. A practical guide to cluster randomised trials in health services research. Chichester, West Sussex, U.K.: John Wiley & Sons; 2012.
10. Hemming K, Taljaard M, Weijer C, Forbes AB. Use of multiple period, cluster randomised, crossover trial designs for comparative effectiveness research. *Bmj*. 2020;371:m3800.
11. Hemming K, Taljaard M, McKenzie JE, Hooper R, et al. Reporting of stepped wedge cluster randomised trials: extension of the CONSORT 2010 statement with explanation and elaboration. *Bmj*. 2018;363:k1614.
12. Arnup SJ, Forbes AB, Kahan BC, Morgan KE, et al. The quality of reporting in cluster randomised crossover trials: proposal for reporting items and an assessment of reporting quality. *Trials*. 2016;17(1):575.
13. Hemming K, Haines TP, Chilton PJ, Girling AJ, et al. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *Bmj*. 2015;350:h391.
14. Hooper R, Bourke L. Cluster randomised trials with repeated cross sections: alternatives to parallel group designs. *Bmj*. 2015;350:h2925.
15. Moher D, Schulz KF, Simera I, Altman DG. Guidance for Developers of Health Research Reporting Guidelines. *PLoS Med*. 2010;7(2):e1000217.
16. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). ICH Harmonised Guideline: Addendum on estimands and sensitivity analysis in clinical trials. Accessed 19 June 2023. https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf; 2019.
17. Kahan BC, Hindley J, Edwards M, Cro S, et al. The estimands framework: a primer on the ICH E9(R1) addendum. *Bmj*. 2024;384:e076316.
18. Welch VA, Norheim OF, Jull J, Cookson R, et al. CONSORT-Equity 2017 extension and elaboration for better reporting of health equity in randomised trials. *Bmj*. 2017;359:j5085.
19. Anderson CS, Arima H, Lavados P, Billot L, et al. Cluster-Randomized, Crossover Trial of Head Positioning in Acute Stroke. *N Engl J Med*. 2017;376(25):2437-47.
20. Booth A, Noyes J, Flemming K, Moore G, et al. Formulating questions to explore complex interventions within qualitative evidence synthesis. *BMJ Global Health*. 2019;4(Suppl 1):e001107.
21. Katheria AC, Clark E, Yoder B, Schmolzer GM, et al. Umbilical cord milking in nonvigorous infants: a cluster-randomized crossover trial. *Am J Obstet Gynecol*. 2023;228(2):217 e1- e14.

22. Bellomo R, Forbes A, Akram M, Bailey M, et al. Why we must cluster and cross over. *Critical care and resuscitation : journal of the Australasian Academy of Critical Care Medicine*. 2013;15(3):155-7.
23. Senn S. The AB/BA Design with Normal Data. *Cross-over Trials In Clinical Research*: John Wiley & Sons, Ltd; 2003. p. 35-88.
24. Copas AJ, Lewis JJ, Thompson JA, Davey C, et al. Designing a stepped wedge trial: three main designs, carry-over effects and randomisation approaches. *Trials*. 2015;16:352.
25. Edited by Higgins JPT, Li T, Sterne JA, on behalf of the RoB 2 working group on crossover trials. Revised Cochrane risk of bias tool for randomized trials (RoB 2) Additional considerations for crossover trials. <https://www.riskofbias.info/welcome/rob-2-0-tool/rob-2-for-crossover-trials>; 2021.
26. Kluytmans-van den Bergh MFQ, Bruijning-Verhagen PCJ, Vandenbroucke-Grauls C, de Brauwier E, et al. Contact precautions in single-bed or multiple-bed rooms for patients with extended-spectrum beta-lactamase-producing Enterobacteriaceae in Dutch hospitals: a cluster-randomised, crossover, non-inferiority study. *Lancet Infect Dis*. 2019;19(10):1069-79.
27. Makris KC, Konstantinou C, Andrianou XD, Charisiadis P, et al. A cluster-randomized crossover trial of organic diet impact on biomarkers of exposure to pesticides and biomarkers of oxidative stress/inflammation in primary school children. *PLoS One*. 2019;14(9):e0219420.
28. Kempen TGH, Bertilsson M, Hadziosmanovic N, Lindner KJ, et al. Effects of Hospital-Based Comprehensive Medication Reviews Including Postdischarge Follow-up on Older Patients' Use of Health Care: A Cluster Randomized Clinical Trial. *JAMA Netw Open*. 2021;4(4):e216303.
29. Ho HJ, Tan YR, Cook AR, Koh G, et al. Increasing Influenza and Pneumococcal Vaccination Uptake in Seniors Using Point-of-Care Informational Interventions in Primary Care in Singapore: A Pragmatic, Cluster-Randomized Crossover Trial. *American journal of public health*. 2019;109(12):1776-83.
30. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *Bmj*. 2008;337:a2390.
31. McIntyre LA, Fergusson DA, McArdle T, Fox-Robichaud A, et al. FLUID trial: a hospital-wide open-label cluster cross-over pragmatic comparative effectiveness randomised pilot trial comparing normal saline to Ringer's lactate. *BMJ Open*. 2023;13(2):e067142.
32. Hoffmann TC, Glasziou PP, Boutron I, Milne R, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *Bmj*. 2014;348:g1687.
33. Pourrat X, Leyrat C, Allenet B, Bouzige B, et al. Effectiveness of a multicomponent pharmacist intervention at hospital discharge for drug-related problems: A cluster randomised cross-over trial. *Br J Clin Pharmacol*. 2020;86(12):2441-54.
34. Butcher NJ, Monsour A, Mew EJ, Chan A-W, et al. Guidelines for Reporting Outcomes in Trial Reports: The CONSORT-Outcomes 2022 Extension. *JAMA*. 2022;328(22):2252-64.
35. Sterne JAC, Savovic J, Page MJ, Elbers RG, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *Bmj*. 2019;366:l4898.
36. Group CS, Sidhu VS, Kelly TL, Pratt N, et al. Effect of Aspirin vs Enoxaparin on Symptomatic Venous Thromboembolism in Patients Undergoing Hip or Knee Arthroplasty: The CRISTAL Randomized Trial. *JAMA*. 2022;328(8):719-27.
37. Lamberink HJ, Otte WM, Sinke MRT, Lakens D, et al. Statistical power of clinical trials increased while effect size remained stable: an empirical analysis of 136,212 clinical trials between 1975 and 2014. *Journal of clinical epidemiology*. 2018;102:123-8.
38. Cook JA, Julious SA, Sones W, Hampson LV, et al. DELTA(2) guidance on choosing the target difference and undertaking and reporting the sample size calculation for a randomised controlled trial. *Bmj*. 2018;363:k3750.
39. Forbes AB, Akram M, Pilcher D, Cooper J, et al. Cluster randomised crossover trials with binary data and unbalanced cluster sizes: application to studies of near-universal interventions in intensive care. *Clin Trials*. 2015;12(1):34-44.

40. Giraudeau B, Ravaud P, Donner A. Sample size calculation for cluster randomized cross-over trials. *Statistics in medicine*. 2008;27(27):5578-85.
41. Hemming K, Kasza J, Hooper R, Forbes A, et al. A tutorial on sample size calculation for multiple-period cluster randomized parallel, cross-over and stepped-wedge trials using the Shiny CRT Calculator. *Int J Epidemiol*. 2020;49(3):979-95.
42. Campbell M, Grimshaw J, Steen N. Sample size calculations for cluster randomised trials. *Changing Professional Practice in Europe Group (EU BIOMED II Concerted Action)*. *J Health Serv Res Policy*. 2000;5(1):12-6.
43. Korevaar E, Kasza J, Taljaard M, Hemming K, et al. Intra-cluster correlations from the CLustered OUtcome Dataset bank to inform the design of longitudinal cluster trials. *Clin Trials*. 2021;18(5):529-40.
44. Goodman SN, Berlin JA. The use of predicted confidence intervals when planning experiments and the misuse of power when interpreting results. *Ann Intern Med*. 1994;121(3):200-6.
45. Peptic Investigators for the Australian New Zealand Intensive Care Society Clinical Trials Group, Alberta Health Services Critical Care Strategic Clinical Network,, the Irish Critical Care Trials Group,, Young PJ, Bagshaw SM, Forbes AB, et al. Effect of Stress Ulcer Prophylaxis With Proton Pump Inhibitors vs Histamine-2 Receptor Blockers on In-Hospital Mortality Among ICU Patients Receiving Invasive Mechanical Ventilation: The PEPTIC Randomized Clinical Trial. *JAMA*. 2020;323(7):616-26.
46. Moulton LH. Covariate-based constrained randomization of group-randomized trials. *Clinical Trials*. 2004;1(3):297-305.
47. Scott NW, McPherson GC, Ramsay CR, Campbell MK. The method of minimization for allocation to clinical trials. a review. *Controlled clinical trials*. 2002;23(6):662-74.
48. Donner A, Klar N. Pitfalls of and controversies in cluster randomization trials. *American journal of public health*. 2004;94(3):416-22.
49. Rosa RG, Falavigna M, da Silva DB, Sganzerla D, et al. Effect of Flexible Family Visitation on Delirium Among Patients in the Intensive Care Unit: The ICU Visits Randomized Clinical Trial. *JAMA*. 2019;322(3):216-28.
50. Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. *Lancet*. 2002;359(9306):614-8.
51. Eldridge S, Campbell MK, Campbell MJ, Dahota AK, et al. Revised Cochrane risk of bias tool for randomized trials (RoB 2) Additional considerations for cluster-randomized trials (RoB 2 CRT). <https://www.riskofbias.info/welcome/rob-2-0-tool/rob-2-for-cluster-randomized-trials>; 2021.
52. Hemming K, Taljaard M, McKenzie JE, Hooper R, et al. The CONSORT extension for Stepped-Wedge Cluster Randomised Trials. *Bmj*. 2018;Accepted March 2018.
53. Caille A, Kerry S, Tavernier E, Leyrat C, et al. Timeline cluster: a graphical tool to identify risk of bias in cluster randomised trials. *Bmj*. 2016;354:i4291.
54. Guy RJ, Ward J, Causer LM, Natoli L, et al. Molecular point-of-care testing for chlamydia and gonorrhoea in Indigenous Australians attending remote primary health services (TTANGO): a cluster-randomised, controlled, crossover trial. *Lancet Infect Dis*. 2018;18(10):1117-26.
55. Estcourt CS, Stirrup O, Copas A, Low N, et al. Accelerated partner therapy contact tracing for people with chlamydia (LUSTRUM): a crossover cluster-randomised controlled trial. *Lancet Public Health*. 2022;7(10):e853-e65.
56. Eldridge S, Kerry S, Torgerson DJ. Bias in identifying and recruiting participants in cluster randomised trials: what can be done? *Bmj*. 2009;339:b4006.
57. Crawshaw AF, Maung TM, Shafique M, Sint N, et al. Acceptability of insecticide-treated clothing for malaria prevention among migrant rubber tappers in Myanmar: a cluster-randomized non-inferiority crossover trial. *Malar J*. 2017;16(1):92.
58. Nix HP, Weijer C, Brehaut JC, Forster D, et al. Informed consent in cluster randomised trials: a guide for the perplexed. *BMJ Open*. 2021;11(9):e054213.

59. Hutton JL. Are distinctive ethical principles required for cluster randomized controlled trials? *Statist Med.* 2001;20(3):473-88.
60. Edwards SJ, Braunholtz DA, Lilford RJ, Stevens AJ. Ethical issues in the design and conduct of cluster randomised controlled trials. *Bmj.* 1999;318(7195):1407-9.
61. Taljaard M, Weijer C, Grimshaw JM, Eccles MP, et al. The Ottawa Statement on the ethical design and conduct of cluster randomised trials: precis for researchers and research ethics committees. *Bmj.* 2013;346:f2838.
62. Giraudeau B, Ravaud P. Preventing bias in cluster randomised trials. *PLoS Med.* 2009;6(5):e1000065.
63. Eldridge SM, Ashby D, Feder GS. Informed patient consent to participation in cluster randomized trials: an empirical exploration of trials in primary care. *Clin Trials.* 2005;2(2):91-8.
64. Rolfzen ML, Wick A, Mascha EJ, Shah K, et al. Best Practice Alerts Informed by Inpatient Opioid Intake to Reduce Opioid Prescribing after Surgery (PRIOR): A Cluster Randomized Multiple Crossover Trial. *Anesthesiology.* 2023;139(2):186-96.
65. Beard E, Lewis JJ, Copas A, Davey C, et al. Stepped wedge randomised controlled trials: systematic review of studies published between 2010 and 2014. *Trials.* 2015;16(1):353.
66. Taljaard M, Hemming K, Shah L, Giraudeau B, et al. Inadequacy of ethical conduct and reporting of stepped wedge cluster randomized trials: Results from a systematic review. *Clin Trials.* 2017:1740774517703057.
67. Ornowska M, Wong H, Ouyang Y, Mitra A, et al. Control of Line Complications with KiteLock (CLiCK) in the critical care unit: study protocol for a multi-center, cluster-randomized, double-blinded, crossover trial investigating the effect of a novel locking fluid on central line complications in the critical care population. *Trials.* 2022;23(1):719.
68. ICMJE International Committee of Medical Journal Editors. Recommendations for the Conduct, Reporting, Editing and Publication of Scholarly Work in Medical Journals.; Updated January 2024. <https://www.icmje.org/>.
69. Kahan BC, Forbes G, Cro S. How to design a pre-specified statistical analysis approach to limit p-hacking in clinical trials: the Pre-SPEC framework. *BMC Med.* 2020;18(1):253.
70. Page MJ, McKenzie JE, Forbes A. Many scenarios exist for selective inclusion and reporting of results in randomized trials and systematic reviews. *Journal of clinical epidemiology.* 2013;66(5):524-37.
71. Committee for Medicinal Products for Human Use. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials Step 5 EMA/CHMP/ICH/436221/2017; 2020 [Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical-principles_en.pdf].
72. Kahan BC, Li F, Copas AJ, Harhay MO. Estimands in cluster-randomized trials: choosing analyses that answer the right question. *International Journal of Epidemiology.* 2022.
73. Kahan BC, Blette BS, Harhay MO, Halpern SD, et al. Demystifying estimands in cluster-randomised trials. *Statistical methods in medical research.* 2024.
74. Kasza J, Hemming K, Hooper R, Matthews J, et al. Impact of non-uniform correlation structure on sample size and power in multiple-period cluster randomised trials. *Statistical methods in medical research.* 2017:962280217734981.
75. Heagerty PJ, Kurland BF. Misspecified maximum likelihood estimates and generalised linear mixed models. *Biometrika.* 2001;88(4):973-85.
76. Freeman PR. The performance of the two-stage analysis of two-treatment, two-period crossover trials. *Statistics in medicine.* 1989;8(12):1421-32.
77. Jones B. Design and analysis of cross-over trials. 2nd ed. ed. Kenward MG, editor. Boca Raton, Fla.: Boca Raton, Fla. : Chapman & Hall/CRC; 2003.
78. Leyrat C, Morgan KE, Leurent B, Kahan BC. Cluster randomized trials with a small number of clusters: which analyses should be used? *Int J Epidemiol.* 2018.

79. Taljaard M, Donner A, Klar N. Imputation strategies for missing continuous outcomes in cluster randomized trials. *Biom J*. 2008;50(3):329 - 45.
80. Caille A, Leyrat C, Giraudeau B. A comparison of imputation strategies in cluster randomized trials with missing binary outcomes. *Statistical methods in medical research*. 2016;25(6):2650-69.
81. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. *Bmj*. 2011;342:d40.
82. Gallis JA, Turner EL. Relative Measures of Association for Binary Outcomes: Challenges and Recommendations for the Global Health Researcher. *Ann Glob Health*. 2019;85(1):137.
83. Mukka S, Hailer NP, Moller M, Gordon M, et al. Study protocol: The DAICY trial-dual versus single-antibiotic impregnated cement in primary hemiarthroplasty for femoral neck fracture-a register-based cluster-randomized crossover-controlled trial. *Acta Orthop*. 2022;93:794-800.
84. Morris TP, Kahan BC, White IR. Choosing sensitivity analyses for randomised trials: principles. *BMC medical research methodology*. 2014;14:11.
85. van Duijn PJ, Verbrugghe W, Jorens PG, Spohr F, et al. The effects of antibiotic cycling and mixing on antibiotic resistance in intensive care units: a cluster-randomised crossover trial. *Lancet Infect Dis*. 2018;18(4):401-9.
86. Dwan K, Gamble C, Williamson PR, Kirkham JJ, et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. *PLoS One*. 2013;8(7):e66844.
87. Hayes RJ, Bennett S. Simple sample size calculation for cluster-randomized trials. *Int J Epidemiol*. 1999;28(2):319-26.
88. Eldridge SM, Ukoumunne OC, Carlin JB. The intra-cluster correlation coefficient in cluster randomized trials: A review of definitions. *International Statistical Review / Revue Internationale de Statistique*. 2009;77(3):378-94.
89. Deeks JJ. Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. *Statistics in medicine*. 2002;21(11):1575-600.
90. Turner EL, Platt AC, Gallis JA, Tetreault K, et al. Completeness of reporting and risks of overstating impact in cluster randomised trials: a systematic review. *Lancet Glob Health*. 2021;9(8):e1163-e8.
91. Gamble C, Krishan A, Stocken D, Lewis S, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. *JAMA*. 2017;318(23):2337-43.
92. Eldridge SM, Chan CL, Campbell MJ, Bond CM, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *Bmj*. 2016;355:i5239.
93. United Nations Educational Scientific and Cultural Organization. UNESCO Recommendation on Open Science. 2021.
94. Kozlov M. NIH issues a seismic mandate: share data publicly. *Nature*. 2022;602:558-9.
95. Government of Canada. Tri-Agency Research Data Management Policy 2021 [Available from: https://www.ic.gc.ca/eic/site/063.nsf/eng/h_97610.html].
96. European Commission. The EU's open science policy [Available from: https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/open-science_en].
97. Ross JS. Clinical research data sharing: what an open science world means for researchers involved in evidence synthesis. *Systematic Reviews*. 2016;5(1):159.
98. Hostler TJ. The Invisible Workload of Open Research. *Journal of Trial & Error*. 2023;<https://doi.org/10.36850/mr5>.
99. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data*. 2016;3:160018.
100. Baker M. 1,500 scientists lift the lid on reproducibility. *Nature*. 2016;533(7604):452-4.
101. Naudet F, Sakarovitch C, Janiaud P, Cristea I, et al. Data sharing and reanalysis of randomized controlled trials in leading biomedical journals with a full data sharing policy: survey of studies published in The BMJ and PLOS Medicine. *Bmj*. 2018;360:k400.

102. Larregue J, Vincent-Lamarre P, Lebaron F, Lariviere V. LSE [Internet]. LSE2020. [cited 2022]. Available from: <https://blogs.lse.ac.uk/impactofsocialsciences/2020/11/30/covid-19-where-is-the-data/>.
103. INVOLVE. Briefing notes for researchers: involving the public in NHS, public health and social care research.; 2012.
104. Gamble C, Dudley L, Allam A, Bell P, et al. Patient and public involvement in the early stages of clinical trial development: a systematic cohort investigation. *BMJ Open*. 2014;4(7):e005234.
105. Staniszewska S, Brett J, Simera I, Seers K, et al. GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research. *Bmj*. 2017;358:j3453.
106. Gray-Burrows KA, Willis TA, Foy R, Rathfelder M, et al. Role of patient and public involvement in implementation research: a consensus study. *BMJ Quality & Safety*. 2018;27(10):858.
107. South A, Hanley B, Gafos M, Cromarty B, et al. Models and impact of patient and public involvement in studies carried out by the Medical Research Council Clinical Trials Unit at University College London: findings from ten case studies. *Trials*. 2016;17(1):376-.
108. Aldcroft A. *BMJ Open*2018. Available from: <https://blogs.bmj.com/bmjopen/2018/03/23/new-requirements-for-patient-and-public-involvement-statements-in-bmj-open/>.
109. Fergusson D, Monfaredi Z, Pussegoda K, Garritty C, et al. The prevalence of patient engagement in published trials: a systematic review. *Research Involvement and Engagement*. 2018;4(1):17.
110. Estcourt CS, Howarth AR, Copas A, Low N, et al. Accelerated partner therapy (APT) partner notification for people with Chlamydia trachomatis: protocol for the Limiting Undetected Sexually Transmitted infections to Reduce Morbidity (LUSTRUM) APT cross-over cluster randomised controlled trial. *BMJ Open*. 2020;10(3):e034806.
111. Medeiros M, Love TR, Slobogean GP, Sprague S, et al. Patient and stakeholder engagement learnings: PREP-IT as a case study. *J Comp Eff Res*. 2021;10(6):439-42.
112. Taichman DB, Sahni P, Pinborg A, Peiperl L, et al. Data Sharing Statements for Clinical Trials: A Requirement of the International Committee of Medical Journal Editors. *Ann Intern Med*. 2017;167(1):63-5.
113. Vickerstaff V, Omar RZ, Ambler G. Methods to adjust for multiple comparisons in the analysis and sample size calculation of randomised controlled trials with multiple primary outcomes. *BMC medical research methodology*. 2019;19(1):129.
114. Li F, Forbes AB, Turner EL, Preisser JS. Power and sample size requirements for GEE analyses of cluster randomized crossover trials. *Statistics in medicine*. 2018.
115. Cro S, Kahan BC, Patel A, Henley A, et al. Starting a conversation about estimands with public partners involved in clinical trials: a co-developed tool. *Trials*. 2023;24(1):443.
116. Anderson DJ, Chen LF, Weber DJ, Moehring RW, et al. Enhanced terminal room disinfection and acquisition and infection caused by multidrug-resistant organisms and *Clostridium difficile* (the Benefits of Enhanced Terminal Room Disinfection study): a cluster-randomised, multicentre, crossover study. *Lancet*. 2017;389(10071):805-14.