OPEN ACCESS

Check for updates

CONSORT 2010 statement: extension to randomised crossover trials

Kerry Dwan,[1] Tianjing Li,[2] Douglas G Altman,[3] Diana Elbourne[4]

[1]Review Production and Quality Unit, Editorial and Methods Department, Cochrane Central Executive, Cochrane, St Alban's House, London SW1Y 4QX, UK

[2]Center for Clinical Trials and Evidence Synthesis, Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

[3]Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK

[4]London School of Hygiene and Tropical Medicine, Department of Medical Statistics, London, UK

Correspondence to: K Dwan
kdwan@cochrane.org

(ORCID 0000-0001-6918-1215)

Cite this as: *BMJ* 2019;366:l4378
http://dx.doi.org/10.1136/bmj.l4378

Accepted: 7 June 2019

Evidence shows the quality of reporting of randomised controlled trials is not optimal. The lack of transparent reporting impedes readers from judging the reliability and validity of trial findings, prevents researchers from extracting information for systematic reviews, and results in research waste. The Consolidated Standards of Reporting Trials (CONSORT) statement was developed to improve the reporting of randomised controlled trials. The primary focus of the statement was on parallel group trials with two treatment groups. Crossover trials are a particular type of trial for chronic conditions in which participants are randomised to a sequence of interventions. They are a useful and efficient design because participants act as their own control. However, the reporting of crossover trials has been variable and incomplete, which hinders their use in clinical decision making and by future researchers. We present the CONSORT extension to randomised crossover trials, which aims to facilitate better reporting of crossover trials. The CONSORT 2010 checklist is revised for crossover designs, and introduces a modified flowchart and baseline table to enhance transparency. Examples of good reporting and evidence based rationale for CONSORT crossover checklist items are provided.

SUMMARY POINTS

- The Consolidated Standards of Reporting Trials (CONSORT) statement provides a minimum set of 25 items to be reported with rationale and exemplars for all randomised trials
- CONSORT extension to crossover trials extends 14 items of the CONSORT statement
- The use of the CONSORT extension to crossover trials will improve reporting of randomised crossover trials

Inadequate reporting of randomised controlled trials (RCTs) is associated with bias in the estimation of treatment effects[1 2]; it also impairs the critical appraisal of the quality of randomised trials, which is important when assessing the validity of the results of the individual trial and when conducting systematic reviews. To attempt to address this issue, the Consolidated Standards of Reporting Trials (CONSORT) statement was developed and includes a set of recommendations for the reporting of RCTs.[3] The statement comprises a checklist of essential items that should be included in reports of RCTs and a diagram to document the flow of participants through the trial from before group assignment through to the final analysis. These items are evidence based when possible. An explanation and elaboration of the rationale for the checklist items are provided in an accompanying article.[4] Many journals now require that reports of RCTs conform to the recommendations in the CONSORT statement.[5]

The primary focus of the CONSORT statement is the most common type of RCT with two treatment groups (two "arms") using an individually randomised, parallel group, superiority design.[3] Almost all the elements of the CONSORT statement apply equally to RCTs with other designs, but some elements need adaptation, and in some cases additional issues need to be discussed. Members of the CONSORT group have published several extension papers that augment the CONSORT statement for different types of interventions and data. Extensions of CONSORT 2010 to different trial designs have also been published, including cluster randomised trials,[6] non-inferiority and equivalence trials,[7] N-of-1 trials,[8] pragmatic trials,[9] and within person trials.[10] As part of that series, in this paper we extend the CONSORT 2010 recommendations to simple crossover RCTs in which participants receive two treatments sequentially over two periods and the order in which treatments are received is randomised.

Scope of this paper

Firstly, we summarise the key methodological features of crossover trials. Secondly, we consider the empirical evidence about how common crossover trials are and review published studies about the quality of reporting of such trials. After these literature reviews, we make suggestions for amendments to the CONSORT checklist adapted for crossover trials and give illustrative examples of good reporting. In this guideline we

BMJ: first published as 10.1136/bmj.l4378 on 31 July 2019. Downloaded from http://www.bmj.com/ on 20 April 2024 by guest. Protected by copyright.

focus on the simplest and most common form of the randomised crossover trial in which all participants receive two interventions in one of two sequences (known as the 2×2 or AB/BA design). Most of the recommendations also apply to the more complicated designs (more than two interventions, periods, or sequences). In a separate section, we briefly discuss specific issues that arise in trials comparing more than two interventions.

Methodological features of randomised crossover trials

In contrast to a parallel group trial, each individual in a crossover trial receives multiple interventions but in a random order; that is, participants are randomised to sequences of interventions. In this way, each participant acts as his or her own control. Such prespecified designs should not be confused with trials in which some individuals "cross over" through non-compliance or use of rescue medication, or in which all participants in the control group are given the chance to "cross over" to the experimental treatment at the end of the main trial. Zeng and colleagues found that almost one quarter of records (n=17/72) labelled as "crossover assignment" did not use a randomised crossover design to randomise participants to a sequence; instead, these trials allowed participants to change intervention during the course of the trial.[11]

Randomised crossover trials present particular challenges. One challenge is the potential for a "carry over effect"; that is, the effect of the first intervention persists into the second period so that the observed difference between the treatments depends on the order in which they were received (see box 1 for glossary of terms). A carry over effect could have a range of causes. In addition to the obvious problem of a drug or other treatment remaining in the system, participants' later responses can be affected by previous side effects or other reactions to previous treatment. It is recommended that crossover trials should include a sufficient "washout" between the end of the first intervention and the start of the second intervention, so that any effects from the first intervention will not be "carried over" to the measurement of outcome in the second intervention period.

Box 1: Glossary

Period: a length of time when one treatment was received.

Sequence: treatment sequence (AB, BA), participants allocated to the AB study arm receive treatment A first, followed by treatment B, and vice versa in the BA arm.

Within participant variability: the expected standard deviation of the within participant differences.

Washout: a length of time between treatment periods when no treatment is received to allow the treatment to wear off.

Carry over effect: when the effect of the first intervention persists into the second period.

Period effect: the outcome of interest changes with time irrespective of treatment effect.

Within participant comparison: a within participant comparison takes into account the correlation between measurements for each participant because they act as their own control, therefore measurements are not independent.

Another issue is the "period effect," which occurs when the outcome of interest changes with time irrespective of treatment effect; for example, the condition might not be stable or the effect of treatment is seasonal.

A further issue is the possibility of participants dropping out of the trial if the first intervention is either very successful or unsuccessful; the results for these participants cannot be included in the analysis.

Design

The particular strength of the simple AB/BA crossover design is that both interventions are evaluated using the same participant, which allows comparison at the individual rather than the group level. In addition, participants in a crossover trial can express preferences by comparing their experiences of the two interventions, which is not possible in a parallel group design because participants will only receive one intervention.[12]

A crucial methodological question is whether the use of the crossover design is justified. Crossover trials are most appropriate for symptomatic treatment (that is, treatment for symptoms, such as pain) of conditions or diseases that are chronic or relatively stable (for example, multiple sclerosis or rheumatoid arthritis), at least over the time period under study; additionally, when the treatment effects are reversible and short lived. The crossover design is inappropriate when the condition of interest can be cured or when participants will probably die during the trial period. The design is commonly used, however, in less appropriate circumstances. For example, pregnancy is an intended outcome of subfertility treatment. If a woman becomes pregnant during the first period of the trial (that is, before crossover), she will be excluded from subsequent phases of the trial. Nevertheless the crossover design is defended in the field[13] (for instance, it has been suggested that pregnancies can be treated statistically as "missing at random"[14]), and remains common despite criticism.[15]

The sample size calculation for such trials is based on the within participant variability in responses. The crossover design is much more efficient than the parallel design when there is a high positive correlation between participants' responses to the different treatments. Compared with a parallel group design, fewer participants are required for a crossover trial to obtain the same power for a target effect size and type 1 error rate.

Crossover trials have certain weaknesses. In particular, there can be carry over effects as previously discussed. Participants could drop out after the first treatment and so not receive the second treatment. Withdrawal might be related to side effects.

Analysis

The analysis of a crossover trial should be based on paired data.[16-18] The estimation approaches should account for the correlation of repeated measurements in the same individual. The tests for significance should use procedures such as the paired t test (assuming no

BMJ: first published as 10.1136/bmj.l4378 on 31 July 2019. Downloaded from http://www.bmj.com/ on 20 April 2024 by guest. Protected by copyright.

carry over or period effect), which is based on within participant differences for a continuous response and the Mainland-Gart test for a binary response.[19 20]

A previously recommended but criticised method for analysing crossover trials was to test for carry over, and if this was statistically significant, to discard the second period data and analyse only the data from the first period. In other words, the first period's data are analysed as if they were from a parallel group trial. Freeman[21] showed that this strategy is flawed and leads to biased answers (which is generally the case when the choice between two analyses is based on the result of a preliminary hypothesis test). Senn[17] and others have argued that the use of the two period, two treatment crossover design is effectively built on the assumption that there is minimal carry over effect.

The other statistical issue specific to crossover studies is the need for adjustment for possible period and carry over effects. Parameters can be included for carry over effect in the statistical model. In the AB/BA crossover design, the terms "carry over" and "treatment by period" interaction are sometime used interchangeably because the effects of "carry over" and "treatment by period interaction" are not separately identifiable in the data. Although the carry over effect can be estimated, Senn[17] and others have argued that there is little value in using the carry over effect to adjust the treatment effect. This is because such adjustment relies on assumptions about the nature of the possible carry over effect and reduces the statistical efficiency for estimating the main treatment effect.

A period effect can be dealt with and adjusted for in the analysis. In the AB/BA crossover design, when equal numbers of participants are allocated to each sequence, then on average the period effect will not bias the estimate of treatment effect. However, a period effect will affect the variance estimate because it interferes with how much of the treatment effect might be attributed to random variation. It is important for authors to present data to help readers understand the extent of the period effect and communicate clearly whether the period effect was adjusted for or not adjusted for in the analysis, and whether such a decision was made a priori.

How common are randomised crossover trials?

A detailed review of all PubMed indexed RCTs published in December 2000 found that 74% (383/519) of trials used a parallel design and 22% (116/519) were crossover trials.[22] Of the trials indexed in Medline in December 2000, 22% (116/526) were crossover trials and most used two treatments (72%) and had two periods (64%).[23] A review of all PubMed indexed RCTs published in December 2006 found 77% (477/616) of trials used a parallel design and 16% (100/616) were crossover trials.[24] A review of intervention studies registered with ClinicalTrials.gov between 2007 and 2010 found that 11.2% (4351/38 969) were crossover trials.[25] A more recent review of PubMed, in December 2012, found that 8.7% (98/1122) of RCTs had a crossover design.[26]

What is the quality of reporting of randomised crossover trials?

Although articles on the quality of reporting of RCTs in relation to CONSORT are relatively common, few articles have specifically examined the quality of reporting of crossover trials. Mills and colleagues found that randomised crossover trials indexed in Medline in December 2000 frequently omitted details on design, analysis, and interpretation.[23] However, most trials reported and defended a washout period (69%, 87/127) and reported use of paired data in the analysis (95%, 121/127). Gewandter and colleagues investigated 124 crossover clinical trials of drug treatments for chronic pain published between 1993 and 2013. They found that 28% (35/124) of trials reported baseline and post washout pain levels, and only 31% (23/75) reported a sample size calculation that specifically indicated that it was based on within participant variability.[27] Straube and colleagues considered 98 crossover trials on chronic painful conditions published between 1990 and 2014 and indexed on PubMed. They found that adverse events were poorly reported in the abstracts of the trial reports and also infrequently reported in the full article, and only 23% (23/98) presented a breakdown by treatment period.[28] Zeng and colleagues found that of 54 phase III randomised crossover trials analysed from ClinicalTrials.gov in September 2014, nearly two thirds had a simple AB/BA design, with most trials (87%, 47/54) providing sufficient information for the participant flow throughout the trial.[11] Baseline characteristics were most often reported for all participants as a single group (59%, 32/54), and primary outcomes and adverse events were most commonly reported "per intervention" (81%, 44/54 and 83%, 45/54, respectively). The reporting of results in baseline characteristics, outcome measures, and adverse events generally did not appear to fully reflect the crossover design.

Several studies have considered the reporting of randomised crossover trials in relation to meta-analyses.[29-31] They found that data were frequently reported inappropriately to allow them to be included in a meta-analysis.

These studies show that the problems have not improved over several years and most of these studies call for guidance on reporting of randomised crossover trials.

Methods used to develop this CONSORT extension

In May 2002, several CONSORT authors met in Arlington, Virginia, USA to consider extensions to the 2001 CONSORT statement for a range of different designs. The first drafts of a paper extending the statement to crossover trials were developed by Doug Altman and Diana Elbourne in 2002-03. In 2010, the CONSORT statement was updated. Work on the extension to crossover trials progressed in 2014 when Kerry Dwan and then Tianjing Li joined the group. The checklist and explanatory text were informed by reviews of published randomised trials (as cited above)

BMJ: first published as 10.1136/bmj.l4378 on 31 July 2019. Downloaded from http://www.bmj.com/ on 20 April 2024 by guest. Protected by copyright.

and completed through numerous teleconferences between the authors from 2014 to 2018. We followed guidance of the CONSORT group to include a member of the CONSORT Group Executive (Doug Altman), who was also chair of the EQUATOR Steering Group. A draft paper was distributed to the wider CONSORT group and other selected individuals, and the paper was revised to take account of their feedback, and approved by the Executive.

CONSORT checklist for randomised crossover RCTs

We discuss the checklist items and focus on any changes to the standard CONSORT items for randomised crossover trials. We explain the background, and provide one or more examples of good reporting. We also discuss other checklist items for which we do not suggest any modifications but where implementation requires specific considerations for crossover RCTs. Table 1 shows the suggested modifications to the standard CONSORT checklist for randomised crossover trials.

Title and abstract

Item 1a: Title

Identification as a randomised crossover trial in the title.

Standard CONSORT item—Identification as a randomised trial in the title.

Example 1—"Effect of Ginkgo biloba on visual field and contrast sensitivity in Chinese patients with normal tension glaucoma: a randomized, crossover clinical trial".[33]

Example 2—"Effects of unfermented and fermented whole grain rye crisp breads served as part of a standardized breakfast, on appetite and postprandial glucose and insulin responses: a randomized cross-over trial".[34]

Explanation—The primary reason for identifying the design in the title is to help readers to identify the study design. Identification of the trial as a randomised crossover trial also ensures that readers will start thinking of the implications of the design in relation to sample size and analysis.

Item 1b: Abstract

Specify a crossover design and report all information outlined in table 2.

Standard CONSORT item—Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts[3]).

Example

CONTEXT: The relationship between sildenafil citrate use and reported adverse cardiovascular events in men with coronary artery disease (CAD) is unclear.

OBJECTIVE: To evaluate the cardiovascular effects of sildenafil during exercise in men with CAD.

DESIGN, SETTING, AND SUBJECTS: Randomised, double blind, placebo controlled *two period* crossover trial conducted March to October 2000 at a US ambulatory care referral centre among 105 men (*55 to receive sildenafil first, and 55 to receive placebo first*) with a mean (SD) age of 66 (9) years who had erectile dysfunction and known or highly suspected CAD.

INTERVENTIONS: All patients underwent two symptom limited supine bicycle echocardiograms separated by an interval of one to three days after receiving a single dose of sildenafil (50 or 100 mg) or placebo one hour before each exercise test.

MAIN OUTCOME MEASURES: Haemodynamic effects of sildenafil during exercise (onset, extent, and severity of ischemia) assessed by exercise echocardiography.

RESULTS: The difference between mean change after sildenafil and placebo use was 4.3 (95% CI 0.9 to 7.7; P=0.01). Exercise capacity was similar with sildenafil use and placebo use (mean difference 0.07; 95% CI −0.06 to 0.19; P=0.29). Exercise blood pressure and heart rate increments were similar. Dyspnoea or angina developed in 69 patients who took sildenafil and 70 patients who took placebo (P=0.89); exercise electrocardiography was positive in 12 patients (11%) who took sildenafil and 17 patients (16%) who took placebo (P=0.09). Exercise induced wall motion abnormalities developed in similar numbers of patients after sildenafil and placebo use (84 and 86 patients, respectively; P=0.53). Wall motion score index at peak exercise was similar after sildenafil and placebo use (mean difference 0.01; 95% CI −0.01 to 0.03; P=0.40).

CONCLUSION: In men with stable CAD, sildenafil had no effect on symptoms, exercise duration, or presence or extent of exercise induced ischaemia, as assessed by exercise echocardiography. (Adapted from Arruda-Olson and colleagues.[36])

Explanation—Clear, transparent, and sufficiently detailed abstracts are important. Readers might only have access to the abstract, and many others will skim it before deciding whether to read further. A well written abstract also helps retrieval of relevant reports from electronic databases. In 2008 a CONSORT extension on reporting abstracts of randomised trials was published,[35] and those recommendations were incorporated into CONSORT 2010.[3] Abstracts for crossover RCTs should indicate the design of the trial and therefore the randomisation to sequence and analysis by taking into account the within participant comparisons. Table 2 shows information to be included in the abstract of a crossover trial.

We were not able to find examples of good reporting tackling all the items required. We have therefore adapted a published abstract (see example).

Methods

Item 3a: Trial design

Rationale for a crossover design. Description of the design features including allocation ratio, especially the number and duration of periods, duration of washout period, and consideration of carry over effect.

Standard CONSORT item—Description of trial design (such as parallel, factorial) including allocation ratio.

Example 1—"The trial was a randomised double-blind, placebo controlled, crossover design of 15 months' duration ... randomisation (1 month);

BMJ: first published as 10.1136/bmj.l4378 on 31 July 2019. Downloaded from http://www.bmj.com/ on 20 April 2024 by guest. Protected by copyright.

Table 1 | CONSORT checklist of information to include when reporting randomised crossover trials

Section/topic	Item No	Description	Page No*
Title†	1a	Identification as a randomised crossover trial in the title	
Abstract†	1b	Specify a crossover design and report all information outlined in table 2	
Introduction:			
Background‡	2a	Scientific background and explanation of rationale	
Objectives‡	2b	Specific objectives or hypotheses	
Methods:			
Trial design†	3a	Rationale for a crossover design. Description of the design features including allocation ratio, especially the number and duration of periods, duration of washout period, and consideration of carry over effect	
Change from protocol‡	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	
Participants‡	4a	Eligibility criteria for participants	
Settings and location‡	4b	Settings and locations where the data were collected	
Interventions†	5	The interventions with sufficient details to allow replication, including how and when they were actually administered	
Outcomes‡	6a	Completely defined prespecified primary and secondary outcome measures, including how and when they were assessed	
Changes to outcomes‡	6b	Any changes to trial outcomes after the trial commenced, with reasons	
Sample size†	7a	How sample size was determined, accounting for within participant variability	
Interim analyses and stopping guidelines‡	7b	When applicable, explanation of any interim analyses and stopping guidelines	
Randomisation:			
Sequence generation‡	8a	Method used to generate the random allocation sequence	
Sequence generation‡	8b	Type of randomisation; details of any restriction (such as blocking and block size)	
Allocation concealment mechanism‡	9	Mechanism used to implement the random allocation sequence§ (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	
Implementation†	10	Who generated the random allocation sequence,§ who enrolled participants, and who assigned participants to the sequence of interventions	
Blinding‡	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	
Similarity of interventions‡	11b	If relevant, description of the similarity of interventions	
Statistical methods†	12a	Statistical methods used to compare groups for primary and secondary outcomes which are appropriate for crossover design (that is, based on within participant comparison)	
Additional analyses‡	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	
Results			
Participant flow (a diagram is strongly recommended)†	13a	The numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome, separately for each sequence and period	
Losses and exclusions†	13b	No of participants excluded at each stage, with reasons, separately for each sequence and period	
Recruitment‡	14a	Dates defining the periods of recruitment and follow-up	
Trial end‡	14b	Why the trial ended or was stopped	
Baseline data†	15	A table showing baseline demographic and clinical characteristics by sequence and period	
Numbers analysed†	16	Number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	
Outcomes and estimation†	17a	For each primary and secondary outcome, results including estimated effect size and its precision (such as 95% confidence interval) should be based on within participant comparisons.¶ In addition, results for each intervention in each period are recommended	
Binary outcomes‡	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 and adjusted analyses, distinguishing prespecified from exploratory	
Harms†	19	Describe all important harms or untended effects in a way that accounts for the design (for specific guidance, see CONSORT for harms[32])	
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	
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	
Protocol‡	24	Where the full trial protocol can be accessed, if available	
Funding‡	25	Sources of funding and other support (such as supply of drugs), role of funders	

CONSORT=Consolidated Standards of Reporting Trials.
*Note: page numbers are optional depending on journal requirements.
†Modified original CONSORT item.
‡Unmodified CONSORT item.
§Random sequence here refers to a list of random orders, typically generated through a computer program. This should not be confused with the sequence of interventions in a randomised crossover trial, for example receiving intervention A before B for an individual trial participant.
¶A within participant comparison takes into account the correlation between measurements for each participant because they act as their own control, therefore measurements are not independent.

BMJ: first published as 10.1136/bmj.l4378 on 31 July 2019. Downloaded from http://www.bmj.com/ on 20 April 2024 by guest. Protected by copyright.

treatment period one (6 months); washout (2 months); and finally treatment period two (6 months) ... Patients were randomly assigned azithromycin in treatment period one, followed by placebo in treatment period two, or placebo in treatment period one followed by azithromycin in treatment period two."[37]

Example 2—"A crossover design was chosen for this study instead of the more traditional randomized, parallel-group design because the within-patient variation is less than the between patient variation and thus required fewer patients. In addition, some of the known disadvantages of the crossover design (e.g. larger dropout rate, instability of the patient's condition, and a potential carryover effect) were not expected in this study."[38]

Example 3—"Each treatment period was separated by a 2-week washout, equating to five or more half-lives for either treatment, to allow the effective systemic elimination of the drug before initiation of subsequent treatment."[39]

Example 4—"We did not include a medicine-free period between treatments to increase patient safety. In addition, we believed the 8-week treatment period was sufficient to allow for the washout of the first treatment before the efficacy measurements at the end of period 2."[40]

Explanation—The methods should contain a rationale for the use of a crossover design in the given setting. In particular, given that a carry over effect can neither be identified with sufficient power, nor can adjustment be made for such an effect in the 2×2 crossover design, the assumption needs to be made that any carry over effects are negligible and some justification presented for this. The description of the design should make clear how many interventions were tested, through how many periods, including information on the length of the treatment, run in, and washout periods (if any).

Item 3b: Changes to methods

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

No change from standard CONSORT item.

Explanation—A test for carry over is not recommended. However, if a test for carry over is performed and as a result the authors use only the first period data, then this should be reported. The use of the test should also be discussed under item 12a (Statistical methods). The reason for the presence of a carry over should also be discussed.

Item 5: Interventions

The interventions with sufficient details to allow replication, including how and when they were actually administered.

Standard CONSORTitem—The interventions for each group with sufficient details to allow replication, including how and when they were actually administered.

Explanation—For this item, "for each group" was deleted for the extension as in a 2×2 randomised crossover trial; the intention is that all participants receive both of the interventions.

Item 7a: Sample size

How sample size was determined, accounting for within participant variability.

Standard CONSORT item—How sample size was determined.

Example—"Earlier research of the Cambridge study site (unpublished data) with the Apathy Evaluation Scale [AES] showed a mean score of 31 points (standard deviation SD=15.6). If we define a clinical significant improvement on the AES-I as a 35% reduction of the mean score, this leads to an absolute effect size of 0.35*31 points=10.85 points. Thus a conservative estimate of 10 units is used for sample size estimation.

Table 2 | Information to include in abstract of report of randomised crossover trial: extension of CONSORT for abstracts checklist[35]

Item	Description
Title*	Identification of study as a randomised crossover trial
Trial design*	Description of the trial design (crossover trial and number of periods)
Methods:	
Participants†	Eligibility criteria for participants and the settings where the data were collected
Interventions*	Interventions intended for all participants
Objective†	Specific objective or hypothesis
Outcome†	Clearly defined primary outcome for this report
Randomisation*	How participants were allocated to sequences
Blinding (masking)*	Whether or not participants, care givers, and those assessing the outcomes were blinded to intervention
Results:	
Numbers randomised*	Number of participants randomised to each sequence
Recruitment†	Trial status‡
Numbers analysed*	Number of participants analysed
Outcome*	For the primary outcome, the estimated effect size and its precision based on within participant comparisons
Harms†	Important adverse events or side effects
Conclusions†	General interpretation of the results
Trial registration†	Registration number and name of trial register
Funding†	Source of funding

CONSORT=Consolidated Standards of Reporting Trials.
*Modified original CONSORT item.
†Unmodified CONSORT item.
‡This is applicable to conference abstracts.

BMJ: first published as 10.1136/bmj.l4378 on 31 July 2019. Downloaded from http://www.bmj.com/ on 20 April 2024 by guest. Protected by copyright.

Furthermore a within subjects SD=15.0 is assumed. When the sample size in each sequence group is 19, (a total sample size of 38) a 2×2 crossover design will have 80% power to detect a difference in means of 10.000 (the difference between a Treatment 1 mean, μ1, of 31 and a Treatment 2 mean, μ2, of 21) assuming that the crossover ANOVA [analysis of variance] √MSE [mean square error] is 15.000 (the Standard deviation of differences, sd, is 21.213) using a two group t test (Crossover ANOVA) with a 0.050 two-sided significance level. In order to account for potential drop-outs 40 patients will be randomized. Sample size calculation was performed with nQuery 7.0."[41]

Explanation—A key advantage of the crossover design is that, for a given significance level, power, and effect size, a smaller sample size is required compared with a parallel design in which each participant receives only one treatment. This is because each participant acts as his or her own control (each participant receives the experimental and control intervention), so the within participant variability is removed.

It is important that trial authors report the usual quantities required for sample size calculation, including significance level and power, but also for continuous variables the within participant variability as shown in the example. It is often difficult to get the necessary within participant information to inform the sample size calculation. Published reports of crossover trials should clarify how the sample size was determined, and ideally should indicate that an appropriate estimate of within participant variability was used. For crossover trials with a continuous outcome, it is the expected standard deviation of the within participant differences that must be incorporated into the sample size estimation. In practice, for many trials it is unlikely that there will be data to support a realistic estimate of this value; however, ignoring it could result in an overestimation of the sample size for a crossover trial and is thus conservative.[42] Some attempt should be made to estimate the standard deviation of the within participant differences (or allow for the correlation).

Likewise, with a binary outcome, not considering the paired nature of the data will result in an unnecessarily large sample size due to failure to account for the within participant comparison arising from the paired design. Authors are expected to give appropriate details so that the sample size calculation can be replicated.

Any allowance in the sample calculation for losses to follow-up should also be reported.

Item 8a: Sequence generation

Method used to generate the random allocation sequence.

No change from standard CONSORT item.

Example 1—"After a 4-week placebo run-in, eligible patients were randomly assigned, according to a computer generated allocation schedule, to 1 of 2 treatment sequences: montelukast and placebo-matching salmeterol or salmeterol and placebo-matching montelukast. After a 2-week washout, patients crossed over to the other treatment."[43]

Example 2—"Eligible subjects were randomized in a 1:1 allocation to one of two treatment sequences—denosumab/alendronate or alendronate/denosumab—and received each treatment for 1 year."[44]

Explanation—In crossover RCTs, allocation sequence refers to the order in which interventions are received. The allocation might be to sequence one, in which participants have A followed by B, or to sequence two, in which participants have B followed by A.

Item 10: Implementation

Who generated the random allocation sequence, who enrolled participants, and who assigned participants to the sequence of interventions.

Standard CONSORT item—Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions.

Explanation—For this item, "the sequence of" was included before interventions as participants are randomised to a sequence of interventions rather than one intervention.

Item 12a: Statistical methods

Statistical methods used to compare groups for primary and secondary outcomes which are appropriate for crossover design (that is, based on within participant comparison).

Standard CONSORT item—Statistical methods used to compare groups for primary and secondary outcomes.

Example 1—"Cross-over analyses for health related quality of life scores averaged the between-treatment difference for each patient within each sequence and then across both sequences, providing an estimate of treatment effect. The estimated treatment difference, 95% CI and P value were adjusted for period and sequence effects in the *analysis of variance model*" (emphasis added).[39]

Example 2—"*A generalized linear mixed models* approach was used to estimate differences between periods of electrical stimulation and no stimulation while accounting for within-subject correlations arising from the crossover design" (emphasis added).[45]

Example 3—"Statistical analysis allowed for the comparison of both treatment groups with respect to baseline information and subsequent comparison at 2 and 4 weeks for treatment effect. The investigator's assessment and patient's assessment of treatment were analysed using *Gart's test* for binary responses, which takes treatment order [strictly period] into account" (emphasis added).[46]

Example 4—"Side effects and patient preferences were analyzed descriptively and using *McNemar's test*" (emphasis added).[47]

Example 5—"*Prescott's test* was used to analyze the primary end point to test the significance of difference between the two treatments in the presence of period effects" (emphasis added).[39]

BMJ: first published as 10.1136/bmj.l4378 on 31 July 2019. Downloaded from http://www.bmj.com/ on 20 April 2024 by guest. Protected by copyright.

Explanation—In line with recommendations made by the International Committee for Medical Journal Editors and the CONSORT group, analytical methods should be described "with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results" (http://www.icmje.org/recommendations/browse/manuscript-preparation/preparing-for-submission.html#d). Identification of the crossover design and the statistical methods used allows readers to evaluate the methods of analysis.

The analysis of a crossover trial should respect the within participant nature of the comparisons. The Methods section should specify which method of analysis was used. This should clearly show how the within participant analysis has been constructed, for example using t tests on within participant differences, or analysis of variance with participant, period, and treatment effects. If period and carry over effects have been modelled, then this should be reported. Similarly, for a binary outcome, conditional logistic regression provides an alternative way of conducting the Mainland-Gart test. The consequences of an analysis not accounting for a within participant comparison could overestimate the variance for the treatment effect.

In some crossover trials participants are measured on the outcome variable at the beginning and at the end of both periods, and the treatment effect is estimated using the change score from each period. This intuitive approach is claimed to eliminate carry over effect; however it could produce a less precise and even biased estimate of treatment effect,[48 49] and therefore should be discouraged.

While missing data raise the same generic issues in crossover trials as in other designs, the specifics are more complicated. The analysis model, in the absence of missing data, should be identified and the role of baseline data needs to be carefully considered because often baseline adjustment increases the standard error. A mixed model of all available data (eg, in this context, with a mixture of fixed and random effects) is typically the preferred first step, with the contextually appropriate adjustment for within subject dependence, and is valid under Rubin's "missing at random" assumption. Broadly, this states that the distribution of later outcome data, given treatment sequence and earlier data, is the same whether or not those data are observed. Analysis of the complete records gives a valid intention to treat estimate by assuming that the distribution of the outcomes given baseline and treatment sequence is the same, whether or not they are observed (that is, missing at random). We can explore the robustness of the conclusion to this untestable assumption by multiply imputing the data and forcing the distribution of imputed outcomes to differ from the observed ones given baseline and treatment sequence. The use of multiple imputation, imputing from subsets of patients (rather than single mean imputation, last value carried forward, or best/worst imputation) is welcome because the imputed data are contextually plausible and appropriately reflect the variability.[50]

Results

Item 13a: Participant flow

The numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome, separately for each sequence and period (a flow diagram is strongly recommended; see fig 1).

Standard CONSORT item—For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome.

Example 1—See figure 2 (adapted from Chen et al[51]).

Example 2—See figure 3 (adapted from Marchetti et al[52]).

Explanation—The flow diagram is a key element of the CONSORT statement and has been widely adopted. For crossover trials it is important to understand the flow of participants across periods. Although we recommend a flow diagram for communicating the flow of participants throughout the study, the exact form and content can vary in relation to the specific features of a trial. We recommend using vertical alignment and including a timescale.

Item 13b: Losses and exclusions

Number of participants excluded at each stage, with reasons, separately for each sequence and period.

Standard CONSORT item—For each group, losses and exclusions after randomisation, together with reasons.

Example 1—"One subject assigned to receive active placebo first withdrew because of a scheduling conflict before taking any study medication. Two subjects assigned to receive pregabalin first withdrew in the first period because of adverse events. The remaining 26 subjects completed the study."[53]

Example 2—"Of the 23 patients who provided consent, 17 were randomized to a treatment sequence (9 to pancrelipase then placebo, 8 to placebo then pancrelipase). Sixteen patients completed the study; 1 patient (pancrelipase/placebo sequence) withdrew consent on day 2 of the first treatment period."[54]

Explanation—Participants who drop out part way through the trial will have their outcome assessed for only one intervention. Dropping out might be informative; for example, they could be dissatisfied with the treatment they were given and so do not wish to try any other treatments. This could bias the results.

Authors should indicate the loss of participants for each intervention, separately for each sequence and period, possibly within the flow diagram with reasons if possible.

There are statistical methods to deal with incomplete data (see Item 12a).

Item 15: Baseline data

A table showing baseline demographic and clinical characteristics by sequence and period.

Standard CONSORT item—A table showing baseline demographic and clinical characteristics for each group.

Example 1—See table 3 (adapted from Fogel et al[43]).

BMJ: first published as 10.1136/bmj.l4378 on 31 July 2019. Downloaded from http://www.bmj.com/ on 20 April 2024 by guest. Protected by copyright.

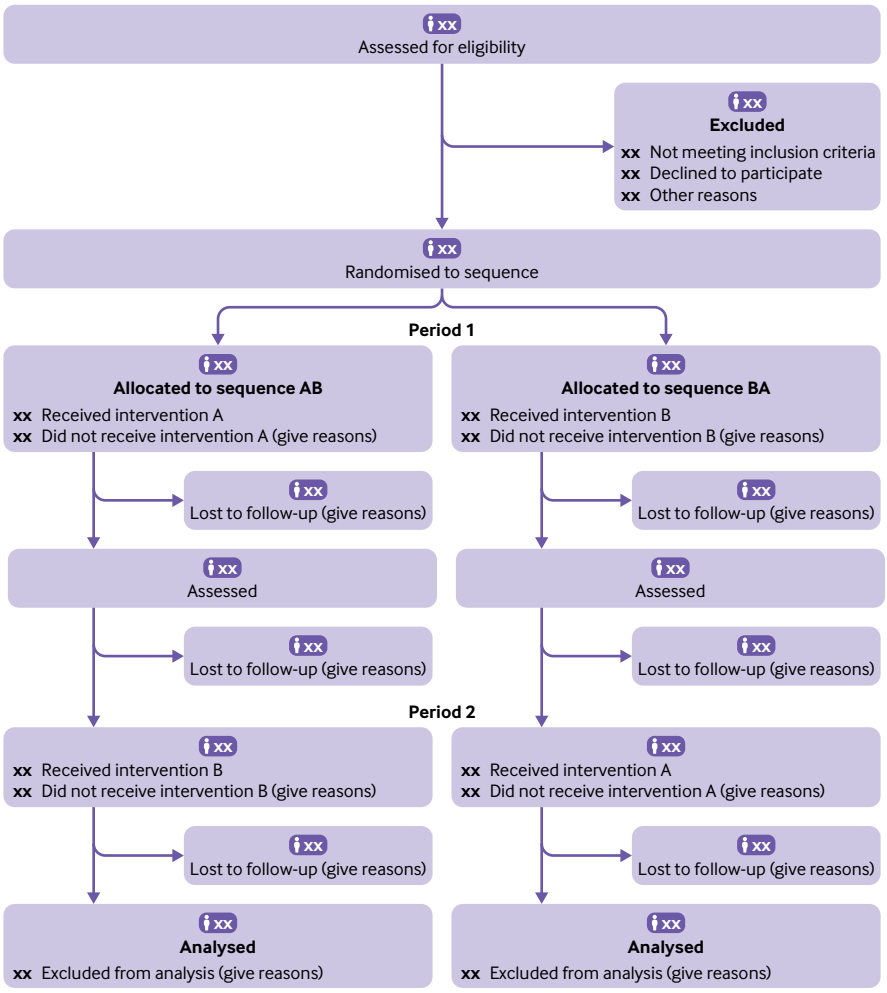

Fig 1 | CONSORT flow diagram for crossover trials (would need modification if more than two groups and/or periods). A=intervention; B=control or other intervention; CONSORT=Consolidated Standards of Reporting Trials. In many cases a washout period can be shown

Example 2—See table 4 (adapted from Valentino et al[55]).

Explanation—Random assignment by individual ensures that any differences in group characteristics at baseline are the result of chance rather than some systematic bias.[2] For randomised crossover trials, it is desirable to know whether baseline characteristics that can be affected by the intervention have returned to their initial state at the beginning of the second period. The by sequence information is needed to assess whether randomisation has achieved balance between the sequences for important variables at the start of the trial. The by period information is helpful for readers to understand whether the treatment effect in the next period is confounded by the changing participant characteristics between periods. Characteristics that remain the same at the start of the two periods, such as sex and age, can be presented once; however, unstable prognostic factors and baseline value of the main outcome must be checked at the beginning of each period. If the characteristic can change over time, then a baseline table by sequence only precludes inference of differences between period (that is, treatment).

Item 16: Numbers analysed

Number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups.

Standard CONSORT item—For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups.

Explanation—The number of participants who contribute to the analysis of a trial is essential to interpreting the results. The analysis of crossover trials has to account for the paired nature of the design; the numbers analysed for each outcome should be equal to the numbers of within participant differences or contrasts that were possible. However, not all participants might contribute to the analysis of each outcome. In a crossover trial, when participants do not contribute to the analysis from one period, the corresponding period may be lost. Assuming no carry over or period effect, if imputation is undertaken the data could be salvaged and when no imputation is undertaken the data are lost, and this becomes a power issue. As the sample size and hence the power

BMJ: first published as 10.1136/bmj.l4378 on 31 July 2019. Downloaded from http://www.bmj.com/ on 20 April 2024 by guest. Protected by copyright.

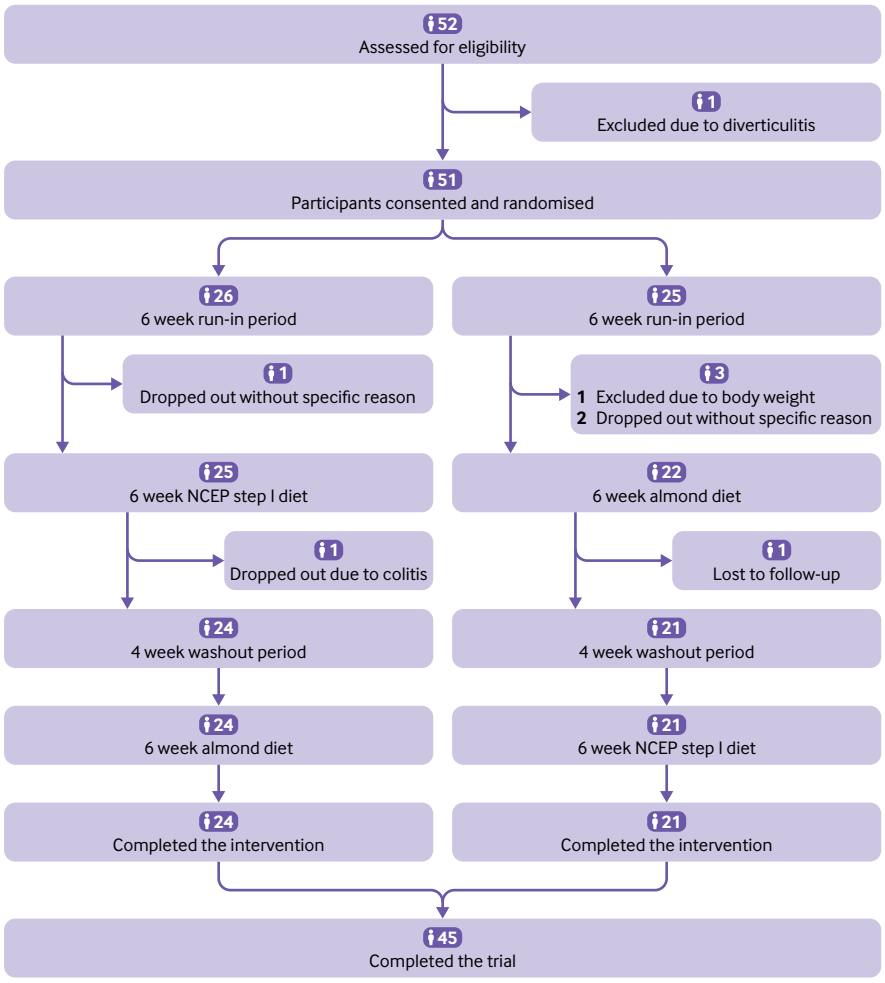

Fig 2 | Example of flow diagram. NCEP=National Cholesterol Education Program

of the study is calculated on the assumption that all participants will provide information, the number of participants contributing to a particular analysis should be reported so that any potential drop in statistical power can be assessed. When there is carry over or a period effect, missing data will result in a biased estimate. In addition, and as explained in detail in the CONSORT 2010 guideline,[2] it should be specified whether a per protocol or an intention to treat analysis was followed.

Item 17a: Outcomes and estimation

For each primary and secondary outcome, results including estimated effect size and its precision (such as 95% confidence interval) should be based on within participant comparisons. In addition, results for each intervention in each period are recommended.

Standard CONSORT item—For each primary and secondary outcome, results for each group, and estimated effect size and its precision (such as 95% confidence interval).

Example 1—See table 5 (adapted from Graff et al[54]).

Example 2—See table 6 (adapted from Rubio-Aurioles et al[38]).

Example 3—"Eighty patients (70%) preferred pazopanib; the most common reasons included better overall quality of life (QoL) and less fatigue. Twenty-five patients (22%) preferred sunitinib; the most common reasons included less diarrhoea and better overall QoL. Physician preferences were consistent with patient preferences. More physicians preferred to continue their patients on pazopanib (61%) than on sunitinib (22%), with 17% stating no preference."[39]

Example 4—See table 7 (adapted from O'Connor et al[56]).

Explanation—When reporting the results of randomised crossover trials, point estimates with confidence intervals should be reported for primary and secondary outcomes; this is the same as the standard CONSORT guideline except that these results should be based on the appropriate within participant analysis. Results should not be presented as though they are from a parallel group trial or by double counting the participants. Ideally, as the correlation impacts on the power of the study, the correlation coefficient for each primary outcome being analysed should also be provided to help with the planning of future crossover trials.

BMJ: first published as 10.1136/bmj.l4378 on 31 July 2019. Downloaded from http://www.bmj.com/ on 20 April 2024 by guest. Protected by copyright.

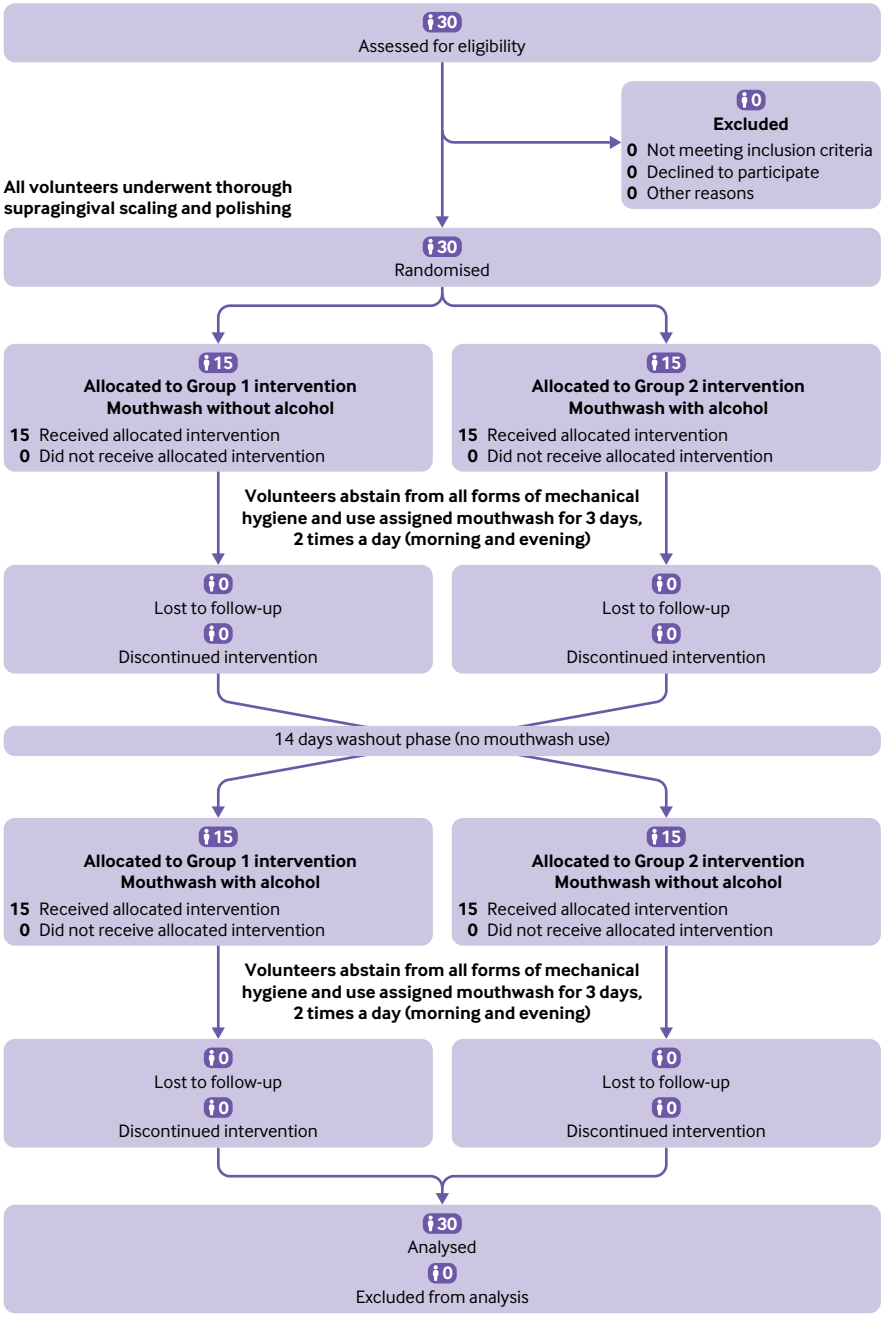

Fig 3 | Example of flow diagram

For binary outcomes a presentation using a matched tabulation format is desirable because it allows the reader to see the concordant and discordant pairs. The matched tabulation facilitates the use of such trials in future meta-analyses because it allows appropriate formulas to be used to adjust the between treatment variance downwards by accounting for the within participant correlation, even when not available.[57-59] Presentation of the 2×2 table of results from a crossover design in a parallel trial format does not allow for appropriate adjustments of the between treatment variance.[57] The paired presentation is also helpful for future sample size calculations. However, in many circumstances the data will be analysed by a model that accounts for the design and is displayed as shown in example 4.

Presentation of the results for each intervention in each period is recommended because these can be used to help understand any treatment by period interaction, regardless of how the trial investigators handled it in their analysis (see table 7 of Li et al[30]).

Ideally, participant preference outcomes should also be reported at the participant level. For example, the participants should be split according to those

BMJ: first published as 10.1136/bmj.l4378 on 31 July 2019. Downloaded from http://www.bmj.com/ on 20 April 2024 by guest. Protected by copyright.

Table 3 | Baseline demographic and clinical characteristics by sequence only

Characteristics	Montelukast–salmeterol (n=78)	Salmeterol–montelukast (n=76)
Sex:		
Male	43 (55.1)	30 (39.5)
Female	35 (44.9)	30 (39.5)
Race:		
Asian	1 (1.3)	0 (0.0)
Black	11 (4.1)	7 (9.2)
White	38 (48.7)	41 (53.9)
Other	28 (35.9)	28 (36.8)
Mean (SD) age (years)	10.2 (2.0)	9.8 (2.0)
Mean (SD) pre-exercise FEV1 (L)	2.30 (1.1)	2.2 (0.6)
Mean (SD) pre-exercise FEV1 (% predicted)	96.3 (31.8)	92.8 (12.4)
Mean (SD) maximum percentage decrease in FEV1 after exercise	24.8 (10.3)	25.4 (9.0)
Mean (SD) AUC0–20min (%·min)	320.1 (208.6)	317.7 (165.7)
Mean* (SD) time to recovery (min)	23.5 (10.5)	21.5 (8.3)
Mean (SD) maximum FEV1 (% predicted)	99.9 (32.5)	100.5 (15.6)
Mean (SD) average percentage change in FEV1 after first SABA use	1.4 (11.0)	4.8 (10.9)
Need for rescue medication after challenge:		
No	77 (98.7)	75 (98.7)
Yes	1 (1.3)	1 (1.3)
Asthma exacerbations limit normal physical activity:		
Not at all	2 (2.6)	4 (5.3)
Slightly	21 (26.9)	20 (26.3)
Moderately	46 (59.0)	44 (57.9)
Severely	9 (11.5)	8 (10.5)

AUC0–20min=area under the curve for the first 20 minutes after exercise; FEV1=forced expiratory volume in 1 second; SABA=short acting β agonist.
Data are number (%) unless stated otherwise.
*Based on the number of patients who returned to within 5% of the baseline FEV1 value.

who prefer intervention A and those who prefer intervention B, and analysed using McNemar's test or, if allowing for period, the Mainland-Gart test or Prescott's test.

Item 19: Harms

Describe all important harms or unintended effects in a way that accounts for the design (for specific guidance see CONSORT for harms[32]).

Standard CONSORT item—All important harms or unintended effects in each group (for specific guidance see CONSORT for harms[32]).

Example—See table 8 (this example is fictional).

Explanation—The types of adverse events and the overall frequency under each intervention should be described. In addition, for crossover trials, presenting concordant and discordant pairs of adverse events or providing estimates of effect and precision (when between group comparisons were made) will inform the relative safety of the interventions tested. The table provides an example of how to tabulate adverse events.

Discussion

Item 20: Limitations

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

Standard CONSORT item—Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses.

Example 1—"The 24-hour washout period may have been insufficient to eliminate the effects of stimulation.

Table 4 | Baseline demographic and clinical characteristics by sequence and by total

Characteristic	Treatment sequence		
	100 IU kg^{-1} once weekly to 50 IU kg^{-1} twice weekly (n=22)	50 IU kg^{-1} twice weekly to 100 IU kg^{-1} once weekly (n=25)	Total (n=50)*
Mean (SD) age (years)	31.7 (13.4)	25.1 (14.4)	27.7 (13.9)
Male sex	22 (100.0)	25 (100.0)	50 (100.0)
Ethnicity:			
White	21 (95.5)	25 (100.0)	49 (98.0)
Black	1 (4.5)	0	1 (2.0)
Hispanic or Latino	5 (22.7)	2 (8.0)	7 (14.0)
Non-Hispanic or non-Latino	17 (77.3)	23 (92.0)	43 (86.0)
Mean (SD) weight (kg)	72.3 (14.2)	64.6 (26.0)	69.2 (21.3)
Target joints†	20 (90.9)	19 (76.0)	42 (84.0)
Haemophilic arthropathy†	20 (90.9)	17 (68.0)	40 (80.0)
Decreased movement due to haemophilic arthropathy†	18 (81.8)	14 (56.0)	34 (68.0)

SD=standard deviation.
Data are number (%) unless stated otherwise.
*Includes three subjects who received study drug in first on demand period, but were not randomised.
†At study entry.

BMJ: first published as 10.1136/bmj.l4378 on 31 July 2019. Downloaded from http://www.bmj.com/ on 20 April 2024 by guest. Protected by copyright.

Table 5 | Coefficient of fat absorption results by treatment and severity of exocrine pancreatic insufficiency

Variable	Pancrelipase (n=16)	Placebo (n=16)	Treatment difference (pancrelipase–placebo) (n=16)	P
CFA (%):				
LS mean (SE)	82.8 (2.7)	47.4 (2.7)	35.4 (3.8)	<0.001
95% CI	77.0 to 88.6	41.6 to 53.2	27.2 to 43.6	—
CFA by severity of EPI (%):				
Placebo CFA≤50%	n=10	n=10	n=10	<0.001
LS mean (SE)	81.8 (1.7)	37.3 (1.7)	44.5 (2.4)	—
95% CI	77.9 to 85.7	33.4 to 41.2	39.0 to 50.0	
Placebo CFA>50%	n=6	n=6	n=6	0.008
LS mean (SE)	84.5 (2.9)	64.3 (2.9)	20.2 (4.1)	—
95% CI	76.5 to 92.5	55.3 to 72.3	8.9 to 31.6	

95% CI=95% confidence interval; CFA=coefficient of fat absorption; EPI=exocrine pancreatic insufficiency; LS, least square; SE=standard error.

Potential carryover effects should be addressed by the use of alternative study designs (eg, parallel groups, longer study/washout periods, stepped-wedge designs)."[45]

Example 2—"Strengths of this study include blinding of study treatments and a cross-over design, where patients were exposed to both treatments in similar health states. This allowed for detection of differences in tolerability not confounded by differences in health states and for each patient to act as their own control. In addition, the 2-week washout period and random assignment minimized possible effects of the order of treatment and carryover."[39]

Example 3—"Finally, it is possible that the crossover design could have obscured differences in the period on and off HCQ [hydroxychloroquine]. While allowing for a washout period may have helped rule out such a possibility, the pilot study suggested no such washout period was required."[60]

Explanation—A limitation with the crossover design is that the treatment from the first period might affect the results from the second period, either to improve the outcome with the opposite treatment or to suppress the effect. This carry over effect could potentially render a crossover trial invalid and reporting of such a limitation is unlikely to be found given that it would invalidate the trial results. Possible limitations that should be reported include losses to follow-up before the second intervention is applied, and mixing up the interventions so that the sequence applied was not the one to which the participant was randomised. The appropriateness of a crossover design in terms of the stability of the disease over the duration of the trial could also be discussed.

More complicated trial designs

In the previous sections we discussed reporting of the simple 2×2 trial design where each participant is randomised to one of two sequences in which to receive the two competing interventions. More complicated variations of the crossover design include comparing three or more interventions (please see the CONSORT extension for multiarm trials[61]) and cluster crossover randomised trials. In a cluster crossover RCT, each cluster receives multiple interventions in a randomised sequence.[62] A recent review found that there is a need to ensure an appropriate analysis is undertaken and reporting needs to be improved.[63]

Table 6 | Treatment comparisons and changes between baseline and treatment endpoint for secondary outcomes

	Changes between baseline and endpoint*			Treatment comparisons†		
Secondary outcome	Sildenafil prn	Tadalafil OaD	Tadalafil prn	Tadalafil OaD–sildenafil prn	Tadalafil OaD–tadalafil prn	Tadalafil prn–sildenafil prn
SEAR scale	25.40 (1.36)	25.56 (1.36)	26.92 (1.35)	0.23 (1.11)	−1.47 (1.11)	1.71 (1.10)
	n=347	n=348	n=355	[−1.95 to 2.42; P=0.834]	[−3.65 to 0.70; P=0.185]	[−0.46 to 3.87; P=0.123]
Sexual relationship	19.50 (1.31)	19.40 (1.31)	20.42 (1.30)	−0.07 (1.07)	−1.12 (1.06)	1.06 (1.06)
	n=347	n=349	n=355	[−2.17 to 2.04; P=0.951]	[−3.22 to 0.97; P=0.291]	[−1.03 to 3.15; P=0.320]
Confidence total	22.87 (1.29)	22.94 (1.29)	24.13 (1.29)	0.11 (1.050)	−1.30 (1.040)	1.42 (1.04)
	n=347	n=348	n=355	[−1.95 to 2.17; P=0.915]	[−3.35 to 0.74; P=0.212]	[−0.63 to 3.46; P=0.174]
IIEF-EF domain score	9.70 (0.36)	8.68 (0.36)	9.54 (0.36)	−0.85 (0.30)	−0.80 (0.29)	−0.05 (0.29)
	n=348	n=350	n=355	[−1.43 to −0.27; P=0.004]	[−1.37 to −0.22; P=0.007]	[−0.62 to 0.53; P=0.866]
EDITS score	75.68 (1.32)	75.81 (1.31)	79.50 (1.31)	0.12 (1.28)	−3.55 (1.27)	3.66 (1.27)
	n=348	n=351	n=355	[−2.40 to 2.64; P=0.926]	[−6.05 to −1.04; P=0.006]	[1.16 to 6.17; P=0.004]
Morning erection frequency	0.11 (0.02)	0.26 (0.02)	0.20 (0.02)	0.15 (0.01)	0.06 (0.01)	0.09 (0.01)
	n=347	n=352	n=355	[0.12 to 0.18; P<0.001]	[0.03 to 0.09; P<0.001]	[0.06 to 0.12; P<0.001]

EDITS=Erectile Dysfunction Inventory of Treatment Satisfaction; IIEF-EF=International Index of Erectile Function-Erectile Function Domain; prn=as required; OaD=once a day; SEAR=Self-Esteem and Relationship.
*Mean (standard error).
†Least square mean difference (standard error) [95% confidence interval; P value].

BMJ: first published as 10.1136/bmj.l4378 on 31 July 2019. Downloaded from http://www.bmj.com/ on 20 April 2024 by guest. Protected by copyright.

Table 7 | Incidence rate ratios (IRRs) with 95% confidence intervals for effects of treatment, time, and treatment–time interaction on behaviour and affect scores after taking account of age, gender, and dementia severity

		Behaviour		Affect			
				Positive		Negative	
Treatment and time		IRR	P value	IRR	P value	IRR	P value
Treatment	Lavender compared with control	0.884 (0.778 to 1.004)	0.057	1.072 (0.848 to 1.355)	0.56	0.891 (0.504 to 1.573)	0.690
Time	First 30 min post exposure compared with pre exposure	0.899 (0.793 to 1.020)	0.097	0.900 (0.706 to 1.147)	0.393	0.960 (0.550 to 1.675)	0.887
	Second 30 min post exposure compared with pre exposure	0.858 (0.755 to 0.974)	0.018	0.865 (0.67 to 1.106)	0.248	0.641 (0.348 to 1.179)	0.153
Treatment–time interactions	Lavender × first 30 min post exposure	0.961 (0.798 to 1.157)	0.672	1.020 (0.726 to 1.433)	0.910	0.848 (0.371 to 1.938)	0.696
	Lavender × second 30 min post exposure	1.045 (0.869 to 1.259)	0.636	0.954 (0.675 to 1.348)	0.790	0.687 (0.269 to 1.750)	0.431

The development of an extension of CONSORT to cluster crossover trials is underway (Joanne McKenzie, personal communication).

There could also be issues of repeated measurements (that is, measurements taken at several time points) or multiplicity within participants in crossover trials (for example, both eyes are assessed within participants). Other, less frequently used versions of the crossover design include bioequivalence studies, Balaam's design, extra period designs, n-of-1 designs, and an incomplete block design.[17]

Comment

Reports of RCTs should include key information on the methods and findings to allow readers to accurately interpret the results. This information is particularly important for meta-analysts attempting to extract data from such reports. The CONSORT 2010 statement provides the latest recommendations from the CONSORT group on essential items to be included in the report of a RCT. In this paper we introduce and explain corresponding updates in an extension of the CONSORT checklist specific to reporting randomised crossover trials.

Use of the CONSORT statement for the reporting of two group parallel trials is associated with improved reporting quality.[64] We believe that the routine use of this proposed extension to the CONSORT statement will eventually result in improvements to crossover designs. When reporting a randomised crossover trial, authors should address all 25 items on the CONSORT checklist by using this document in conjunction with the main CONSORT guidelines.[3] Authors might also find it useful to consult the CONSORT extensions for other trial designs (available at www.consort-statement.org/extensions).

Table 8 | Adverse events

Adverse event		No of adverse events
Vomiting	No adverse event under either NSAIDS or placebo	108
	No adverse event under NSAIDS but adverse events observed under placebo	7
	Adverse event observed under NSAIDS but not under placebo	13
	Adverse events observed under both NSAIDS and placebo	3

NSAID=non-steroidal anti-inflammatory drug.

The CONSORT statement can help researchers to design trials in the future and can guide peer reviewers and editors in their evaluation of manuscripts. Many journals recommend authors adhere to the CONSORT recommendations in their instructions to authors. We encourage them to direct authors to this and to other extensions of CONSORT for specific trial designs. The most up to date versions of all CONSORT recommendations can be found online (www.consort-statement.org).

We thank James Carpenter and Mike Kenward for their input into how missing data are handled in crossover trials. We also thank Sally Hopewell, Nikolaos Pandis, and Drummond Rennie, and other members of the CONSORT Executive for helpful comments, as well as the peer reviewers Stephen Senn, Francois Curtin, and Karla Hemming. Note that small parts of the text in this manuscript are necessarily similar to other CONSORT articles.

Contributors: DGA and DE initiated the work. KD, TL, DE, and DGA drafted the manuscript and all authors reviewed it. KD, TL, and DE approved the final version (DGA died in June 2018). KD is the guarantor. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: This study received no specific funding.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: TL reports grants from National Eye Institute, National Institutes of Health, and grants from National Library of Medicine, National Institutes of Health during the conduct of the study.

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

1 Page MJ, Higgins JP, Clayton G, Sterne JA, Hróbjartsson A, Savović J. Empirical evidence of study design biases in randomized trials: systematic review of meta-epidemiological studies. *PLoS One* 2016;11:e0159267. doi:10.1371/journal.pone.0159267
2 Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408-12. doi:10.1001/jama.1995.03520290060030
3 Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c332. doi:10.1136/bmj.c332
4 Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c869. doi:10.1136/bmj.c869
5 Shamseer L, Hopewell S, Altman DG, Moher D, Schulz KF. Update on the endorsement of CONSORT by high impact factor journals: a survey of journal "Instructions to Authors" in 2014. *Trials* 2016;17:301. doi:10.1186/s13063-016-1408-z
6 Campbell MK, Piaggio G, Elbourne DR, Altman DG, CONSORT Group. Consort 2010 statement: extension to cluster

BMJ: first published as 10.1136/bmj.l4378 on 31 July 2019. Downloaded from http://www.bmj.com/ on 20 April 2024 by guest. Protected by copyright.

randomised trials. *BMJ* 2012;345:e5661. doi:10.1136/bmj.e5661
7 Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG, CONSORT Group. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA* 2012;308:2594-604. doi:10.1001/jama.2012.87802
8 Vohra S, Shamseer L, Sampson M, et al, CENT Group. CONSORT extension for reporting N-of-1 trials (CENT) 2015 Statement. *BMJ* 2015;350:h1738. doi:10.1136/bmj.h1738
9 Zwarenstein M, Treweek S, Gagnier JJ, et al, CONSORT Group, Pragmatic Trials in Healthcare (Practihc) Group. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 2008;337:a2390. doi:10.1136/bmj.a2390
10 Pandis N, Chung B, Scherer RW, Elbourne D, Altman DG. CONSORT 2010 statement: extension checklist for reporting within person randomised trials. *BMJ* 2017;357:j2835. doi:10.1136/bmj.j2835
11 Zeng, L., L.A. Drye, T. Li, Characterizing current registration of phase 3 crossover trials on ClinicalTrials.gov. 2015. http://jhir.library.jhu.edu/handle/1774.2/38151.
12 Hui D, Zhukovsky DS, Bruera E. Which treatment is better? Ascertaining patient preferences with crossover randomized controlled trials. *J Pain Symptom Manage* 2015;49:625-31. doi:10.1016/j.jpainsymman.2014.11.294
13 Cohlen BJ, te Velde ER, Looman CW, Eijckemans R, Habbema JD. Crossover or parallel design in infertility trials? The discussion continues. *Fertil Steril* 1998;70:40-5. doi:10.1016/S0015-0282(98)00114-9
14 Makubate B, Senn S. Planning and analysis of cross-over trials in infertility. *Stat Med* 2010;29:3203-10. doi:10.1002/sim.3981
15 Daya S. Differences between crossover and parallel study designs—debate?*Fertil Steril* 1999;71:771-3.
16 Jones B, Kenward MG. *Design and Analysis of Cross-over Trials*. 2nd edition ed. 2003, London: Chapman&Hall/CRC.
17 Senn S. *Crossover-trials in Clinical Research*. 2nd ed. Wiley, 2002. doi:10.1002/0470854596.
18 Wellek S, Blettner M. On the proper use of the crossover design in clinical trials: part 18 of a series on evaluation of scientific publications. *Dtsch Arztebl Int* 2012;109:276-81.
19 Gart JJ. An exact test for comparing matched proportions in cross-over designs. *Biometrika* 1969;56:75-80. doi:10.1093/biomet/56.1.75
20 Mainland D. *Elementary Medical Statistics*. W.B. Saunders, 1963.
21 Freeman PR. The performance of the two-stage analysis of two-treatment, two-period crossover trials. *Stat Med* 1989;8:1421-32. doi:10.1002/sim.4780081202
22 Chan AW, Altman DG. Epidemiology and reporting of randomised trials published in PubMed journals. *Lancet* 2005;365:1159-62. doi:10.1016/S0140-6736(05)71879-1
23 Mills EJ, Chan AW, Wu P, Vail A, Guyatt GH, Altman DG. Design, analysis, and presentation of crossover trials. *Trials* 2009;10:27. doi:10.1186/1745-6215-10-27
24 Yu LM, Chan AW, Hopewell S, Deeks JJ, Altman DG. Reporting on covariate adjustment in randomised controlled trials before and after revision of the 2001 CONSORT statement: a literature review. *Trials* 2010;11:59. doi:10.1186/1745-6215-11-59
25 Inrig JK, Califf RM, Tasneem A, et al. The landscape of clinical trials in nephrology: a systematic review of Clinicaltrials.gov. *Am J Kidney Dis* 2014;63:771-80. doi:10.1053/j.ajkd.2013.10.043
26 Odutayo A, Emdin CA, Hsiao AJ, et al. Association between trial registration and positive study findings: cross sectional study (Epidemiological Study of Randomized Trials-ESORT). *BMJ* 2017;356:j917. doi:10.1136/bmj.j917
27 Gewandter JS, McDermott MP, McKeown A, et al. Reporting of cross-over clinical trials of analgesic treatments for chronic pain: Analgesic, Anesthetic, and Addiction Clinical Trial Translations, Innovations, Opportunities, and Networks systematic review and recommendations. *Pain* 2016;157:2544-51. doi:10.1097/j.pain.0000000000000673
28 Straube S, Werny B, Friede T. A systematic review identifies shortcomings in the reporting of crossover trials in chronic painful conditions. *J Clin Epidemiol* 2015;68:1496-503. doi:10.1016/j.jclinepi.2015.04.006
29 Lathyris DN, Trikalinos TA, Ioannidis JP. Evidence from crossover trials: empirical evaluation and comparison against parallel arm trials. *Int J Epidemiol* 2007;36:422-30. doi:10.1093/ije/dym001
30 Li T, Yu T, Hawkins BS, Dickersin K. Design, analysis, and reporting of crossover trials for inclusion in a meta-analysis. *PLoS One* 2015;10:e0133023. doi:10.1371/journal.pone.0133023
31 Nolan SJ, Hambleton I, Dwan K. The use and reporting of the cross-over study design in clinical trials and systematic reviews: a systematic assessment. *PLoS One* 2016;11:e0159014. doi:10.1371/journal.pone.0159014
32 Ioannidis JP, Evans SJ, Gøtzsche PC, et al. CONSORT Group. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004;141:781-8. doi:10.7326/0003-4819-141-10-200411160-00009
33 Guo X, Kong X, Huang R, et al. Effect of Ginkgo biloba on visual field and contrast sensitivity in Chinese patients with normal tension glaucoma: a randomized, crossover clinical trial. *Invest Ophthalmol Vis Sci* 2014;55:110-6. doi:10.1167/iovs.13-13168
34 Johansson DP, Lee I, Risérus U, Langton M, Landberg R. Effects of unfermented and fermented whole grain rye crisp breads served as part of a standardized breakfast, on appetite and postprandial glucose and insulin responses: a randomized cross-over trial. *PLoS One* 2015;10:e0122241. doi:10.1371/journal.pone.0122241
35 Hopewell S, Clarke M, Moher D, et al, CONSORT Group. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet* 2008;371:281-3. doi:10.1016/S0140-6736(07)61835-2
36 Arruda-Olson AM, Mahoney DW, Nehra A, Leckel M, Pellikka PA. Cardiovascular effects of sildenafil during exercise in men with known or probable coronary artery disease: a randomized crossover trial. *JAMA* 2002;287:719-25. doi:10.1001/jama.287.6.719
37 Equi A, Balfour-Lynn IM, Bush A, Rosenthal M. Long term azithromycin in children with cystic fibrosis: a randomised, placebo-controlled crossover trial. *Lancet* 2002;360:978-84. doi:10.1016/S0140-6736(02)11081-6
38 Rubio-Aurioles E, Porst H, Kim ED, et al. A randomized open-label trial with a crossover comparison of sexual self-confidence and other treatment outcomes following tadalafil once a day vs. tadalafil or sildenafil on-demand in men with erectile dysfunction. *J Sex Med* 2012;9:1418-29. doi:10.1111/j.1743-6109.2012.02667.x
39 Escudier B, Porta C, Bono P, et al. Randomized, controlled, double-blind, cross-over trial assessing treatment preference for pazopanib versus sunitinib in patients with metastatic renal cell carcinoma: PISCES Study. *J Clin Oncol* 2014;32:1412-8. doi:10.1200/JCO.2013.50.8267
40 Konstas AG, Lake S, Economou AI, Kaltsos K, Jenkins JN, Stewart WC. 24-Hour control with a latanoprost-timolol fixed combination vs timolol alone. *Arch Ophthalmol* 2006;124:1553-7. doi:10.1001/archopht.124.11.1553
41 Gelderblom H, Wüstenberg T, McLean T, et al. Bupropion for the treatment of apathy in Huntington's disease: a multicenter, randomised, double-blind, placebo-controlled, prospective crossover trial. *PLoS One* 2017;12:e0173872. doi:10.1371/journal.pone.0173872
42 Julious SA, Campbell MJ, Altman DG. Estimating sample sizes for continuous, binary, and ordinal outcomes in paired comparisons: practical hints. *J Biopharm Stat* 1999;9:241-51. doi:10.1081/BIP-100101174
43 Fogel RB, Rosario N, Aristizabal G, et al. Effect of montelukast or salmeterol added to inhaled fluticasone on exercise-induced bronchoconstriction in children. *Ann Allergy Asthma Immunol* 2010;104:511-7. doi:10.1016/j.anai.2009.12.011
44 Freemantle N, Satram-Hoang S, Tang ET, et al, DAPS Investigators. Final results of the DAPS (Denosumab Adherence Preference Satisfaction) study: a 24-month, randomized, crossover comparison with alendronate in postmenopausal women. *Osteoporos Int* 2012;23:317-26. doi:10.1007/s00198-011-1780-1
45 Abell TL, Johnson WD, Kedar A, et al. A double-masked, randomized, placebo-controlled trial of temporary endoscopic mucosal gastric electrical stimulation for gastroparesis. *Gastrointest Endosc* 2011;74:496-503.e3.
46 Hill J, Bird HA, Fenn GC, Lee CE, Woodward M, Wright V. A double-blind crossover study to compare lysine acetyl salicylate (Aspergesic) with ibuprofen in the treatment of rheumatoid arthritis. *J Clin Pharm Ther* 1990;15:205-11. doi:10.1111/j.1365-2710.1990.tb00376.x
47 Bonten TN, Snoep JD, Assendelft WJ, et al. Time-dependent effects of aspirin on blood pressure and morning platelet reactivity: a randomized cross-over trial. *Hypertension* 2015;65:743-50. doi:10.1161/HYPERTENSIONAHA.114.04980
48 Fleiss JL, Wallenstein S, Rosenfeld R. Adjusting for baseline measurements in the two-period crossover study: a cautionary note. *Control Clin Trials* 1985;6:192-7. doi:10.1016/0197-2456(85)90002-9
49 Willan AR, Pater JL. Using baseline measurements in the two-period crossover clinical trial. *Control Clin Trials* 1986;7:282-9. doi:10.1016/0197-2456(86)90036-X
50 Carpenter JR, Kenward MG. Sensitivity Analysis with Multiple Imputation. In: Molenberghs G, et al. eds. *Handbook of Missing Data Methodology*. CRC Press New York, 2015: 446.
51 Chen CY, Holbrook M, Duess MA, et al. Effect of almond consumption on vascular function in patients with coronary artery disease: a randomized, controlled, cross-over trial. *Nutr J* 2015;14:61. doi:10.1186/s12937-015-0049-5
52 Marchetti E, Mummolo S, Di Mattia J, et al. Efficacy of essential oil mouthwash with and without alcohol: a 3-day plaque accumulation model. *Trials* 2011;12:262. doi:10.1186/1745-6215-12-262
53 Markman JD, Frazer ME, Rast SA, et al. Double-blind, randomized, controlled, crossover trial of pregabalin for neurogenic

BMJ: first published as 10.1136/bmj.l4378 on 31 July 2019. Downloaded from http://www.bmj.com/ on 20 April 2024 by guest. Protected by copyright.

claudication. *Neurology* 2015;84:265-72. doi:10.1212/WNL.0000000000001168
54 Graff GR, Maguiness K, McNamara J, et al. Efficacy and tolerability of a new formulation of pancrelipase delayed-release capsules in children aged 7 to 11 years with exocrine pancreatic insufficiency and cystic fibrosis: a multicenter, randomized, double-blind, placebo-controlled, two-period crossover, superiority study. *Clin Ther* 2010;32:89-103. doi:10.1016/j.clinthera.2010.01.012
55 Valentino LA, Rusen L, Elezovic I, Smith LM, Korth-Bradley JM, Rendo P. Multicentre, randomized, open-label study of on-demand treatment with two prophylaxis regimens of recombinant coagulation factor IX in haemophilia B subjects. *Haemophilia* 2014;20:398-406. doi:10.1111/hae.12344
56 O'Connor DW, Eppingstall B, Taffe J, van der Ploeg ES. A randomized, controlled cross-over trial of dermally-applied lavender (Lavandula angustifolia) oil as a treatment of agitated behaviour in dementia. *BMC Complement Altern Med* 2013;13:315. doi:10.1186/1472-6882-13-315
57 Elbourne DR, Altman DG, Higgins JP, Curtin F, Worthington HV, Vail A. Meta-analyses involving cross-over trials: methodological issues. *Int J Epidemiol* 2002;31:140-9. doi:10.1093/ije/31.1.140
58 Stedman MR, Curtin F, Elbourne DR, Kesselheim AS, Brookhart MA. Meta-analyses involving cross-over trials: methodological issues. *Int J Epidemiol* 2011;40:1732-4. doi:10.1093/ije/dyp345
59 Curtin F, Elbourne D, Altman DG. Meta-analysis combining parallel and cross-over clinical trials. II: Binary outcomes. *Stat Med* 2002;21:2145-59. doi:10.1002/sim.1206
60 Solomon DH, Garg R, Lu B, et al. Effect of hydroxychloroquine on insulin sensitivity and lipid parameters in rheumatoid arthritis patients without diabetes mellitus: a randomized, blinded crossover trial. *Arthritis Care Res (Hoboken)* 2014;66:1246-51. doi:10.1002/acr.22285
61 Juszczak E, Altman DG, Hopewell S, Schulz K. Reporting of Multi-Arm Parallel-Group Randomized Trials: Extension of the CONSORT 2010 Statement. *JAMA* 2019;321:1610–20. doi:10.1001/jama.2019.3087
62 Rietbergen C, Moerbeck M. The Design of Cluster Randomized Crossover Trials. *J Educ Behav Stat* 2011;36:472-90. doi:10.3102/1076998610379136.
63 Arnup SJ, Forbes AB, Kahan BC, Morgan KE, McKenzie JE. Appropriate statistical methods were infrequently used in cluster-randomized crossover trials. *J Clin Epidemiol* 2016;74:40-50. doi:10.1016/j.jclinepi.2015.11.013
64 Turner L, Shamseer L, Altman DG, Schulz KF, Moher D. Does use of the CONSORT Statement impact the completeness of reporting of randomised controlled trials published in medical journals? A Cochrane review. *Syst Rev* 2012;1:60. doi:10.1186/2046-4053-1-60

BMJ: first published as 10.1136/bmj.l4378 on 31 July 2019. Downloaded from http://www.bmj.com/ on 20 April 2024 by guest. Protected by copyright.

No commercial reuse: See rights and reprints http://www.bmj.com/permissions Subscribe: http://www.bmj.com/subscribe