Journal of Clinical Epidemiology 101 (2018) 53–60

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Pre-specification of statistical analysis approaches in published clinical trial protocols was inadequate

Lauren Greenberg[a], Vipul Jairath[b,c], Rupert Pearse[d], Brennan C. Kahan[a,*]

[a]*Pragmatic Clinical Trials Unit, Queen Mary University of London, UK*
[b]*Department of Medicine, Western University, London, Ontario, Canada*
[c]*Department of Epidemiology and Biostatistics, Western University, London, Ontario, Canada*
[d]*Adult Critical Care Unit, Royal London Hospital, London E1 1BB, UK*

Accepted 25 May 2018; Published online 2 June 2018

Abstract

Objectives: Results from randomized trials can depend on the statistical analysis approach used. It is important to prespecify the analysis approach in the trial protocol to avoid selective reporting of analyses based on those which provide the most favourable results. We undertook a review of published trial protocols to assess how often the statistical analysis of the primary outcome was adequately prespecified.

Methods: We searched protocols of randomized trials indexed in PubMed in November 2016. We identified whether the following aspects of the statistical analysis approach for the primary outcome were adequately prespecified: (1) analysis population; (2) analysis model; (3) use of covariates; and (4) method of handling missing data.

Results: We identified 99 eligible protocols. Very few protocols adequately prespecified the analysis population (8/99, 8%), analysis model (27/99, 27%), covariates (40/99, 40%), or approach to handling missing data (10/99, 10%). Most protocols did not adequately predefine any of these four aspects of their statistical analysis approach (39%) or predefined only one aspect (36%). No protocols adequately predefined all four aspects of the analysis.

Conclusion: The statistical analysis approach is rarely prespecified in published trial protocols. This may allow selective reporting of results based on different analyses.

Keywords: Randomized controlled trial; Clinical trial; Statistical analysis; Pre-specification; Analysis switching

1. Introduction

Well-designed clinical trials are the gold standard for evaluating the efficacy and safety of health care interventions. It is widely agreed that the trial methodology should be prespecified in the protocol to avoid issues such as selective reporting of results [1,2]. Previous research has shown that failure to adequately prespecify trial outcomes can lead to ‘‘outcome switching,’’ where statistically significant outcomes are more likely to be reported than nonsignificant ones, leading to exaggerated treatment effect sizes and misleading conclusions [3–13].

Similar issues are faced when specifying a statistical analysis plan (SAP) for the trial [14–17]. The analysis approach should be chosen to address the study research question and involves a series of decisions, including identifying the participants to be included in the analysis, the statistical model to be used, and the method of handling missing data [1,2]. Different approaches could lead to different results and hence influence the interpretation of the trial. It is therefore important that these decisions are prespecified before seeing the trial data because lack of prespecification may affect the trial’s validity by allowing investigators to selectively report the analysis approach that provides the most favorable results [18].

The International Conference for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH)-E9 guidelines state that the trial protocol should contain ‘‘all the principal features of the proposed confirmatory analysis of the primary variable(s)’’ [19]. Similarly, the SPIRIT (Standard Protocol Items: Recommendations for Interventional

Funding: None.
Conflict of interest: None.

* Corresponding author. Yvonne Carter Building, 58 Turner Street, London, Ontario E1 2AB, Canada. Tel.: 0207 882 6307.
E-mail address: b.kahan@qmul.ac.uk (B.C. Kahan).

https://doi.org/10.1016/j.jclinepi.2018.05.023
0895-4356/

What is New?

- The chosen statistical analysis approach can affect results from randomized trials. Pre-specification of the analysis approach can guard against selective reporting of analyses; however, it is not known how often the statistical analysis approach is adequately prespecified in trial protocols.
- Our review found that no protocols adequately prespecified their entire statistical analysis approach for the primary outcome. The analysis population and the approach to handling missing data had the lowest rates of pre-specification; however, the analysis model and the use of covariates were also poorly prespecified.
- An exploratory re-analysis of two trials found that changing the analysis approach based on the trial data could lead to either statistically significant or nonsignificant results, depending on what the investigator wished to show.
- Many trials may be at risk of selective reporting of statistical analyses, which could affect the interpretation of study results.

Trials) guidelines state that ''The protocol should prespecify the main (''primary'') analysis of the primary outcome, including the analysis methods to be used for statistical comparisons; precisely which trial participants will be included; and how missing data will be handled'' [2]. The aims of this study were to evaluate whether statistical analysis approaches for the primary outcome were being adequately prespecified in published trial protocols, with a particular focus on the analysis population, analysis model, use of covariates, and handling of missing data.

2. Methods

2.1. Review of published protocols

We conducted a review of published trial protocols to assess how well statistical analysis approaches were being prespecified. Protocols of randomized controlled trials conducted in humans and published in English were eligible for inclusion, regardless of therapeutic area or nature of the intervention. The main exclusion criteria were pilot and feasibility trials, and phase 1 or phase 2 trials. This was because we wanted to focus on large, phase III trials that could affect clinical practice. We also excluded articles with a primary outcome of cost-effectiveness and any articles with published results.

We identified articles in a PubMed search of titles and abstracts using the terms ''protocol'' or ''randomi*'' and excluding articles that included the terms ''pilot,'' ''feasibility,'' ''phase 1,'' ''phase one,'' ''phase i,'' ''phase 2,'' ''phase two,'' and ''phase ii'' in the title. We restricted the search to articles published in November 2016. One author (L.G.) initially screened abstracts to identify appropriate full-text articles. All full-text articles were screened independently and in duplicate by two authors (L.G. and B.C.K.) to ensure they met the inclusion criteria.

Two authors (L.G. and B.C.K.) independently extracted data for all included protocols onto a standardized, prepiloted form. We extracted information on whether the following elements were adequately predefined in relation to the primary outcome: (1) the analysis population to be used; (2) the analysis model to be used; (3) the covariates to be included in the model; and (4) the method of handling missing data. Further details on these elements are available in Table 1. Discrepancies between extractors were resolved by discussion.

For protocols that did not specify a primary outcome or specified multiple primary outcomes, we used the outcome used in the sample size calculation. If no sample size calculation was reported, or if the sample size calculation was performed for multiple primary outcomes, we used the first outcome listed in the protocol abstract.

We classified each element as either (1) adequately predefined; (2) incompletely predefined; or (3) not mentioned. Elements were classified as adequately predefined if they contained sufficient detail to allow replication by a third party and would not allow the analyst to choose the analysis approach subjectively based on the trial data. Elements were classified as incompletely predefined if some detail was included but not enough to allow replication by a third party (eg, if a per-protocol population was specified without defining under which circumstances patients would be excluded from the analysis) or if it allowed the analyst to choose the analysis approach subjectively based on the data (eg, if the analyst was to choose between multiple analysis models based on the fit of the data, but no objective or reproducible method for choosing was given). Elements were classified as not mentioned if they were not addressed at all in the text.

2.2. Exploratory re-analysis of the OPTIMISE and TRIGGER trials

We also conducted an exploratory re-analysis of two randomized trials that were recently completed by two authors (R.P. and V.J.) in order to assess the impact that changing the analysis approach could have on results. Specifically, we wished to see how extreme the difference in results for each trial could be if the analyst was choosing the analysis approach based on the trial data to obtain a specific result (to demonstrate either as large or as small of an effect as possible).

For each trial, we chose an initial reference method of analysis. We then varied different aspects of the analysis in turn, to obtain either a larger or smaller effect than that

Table 1. Requirements for adequate pre-specification of each statistical element

Analysis element	Definition	Requirements for adequate pre-specification
Analysis population	The set of patients that will be included in the analysis.	The exact criteria for determining whether each patient will be included in the analysis should be defined. Generic labels such as intention-to-treat or per-protocol without further elaboration are not sufficient as these terms are not used consistently. If multiple populations will be analyzed (eg, all patients vs. only patients who received at least one dose of study treatment), it should be specified which of these is considered the primary analysis population.
Analysis model	The statistical method that will be used to generate the treatment effect, confidence interval, or *P*-value.	The statistical model to be used for analysis should be specified, with sufficient detail to allow replication. For instance, if generalized estimating equations were to be used for the analysis, the working correlation matrix and whether robust standard errors will be used should also be specified. If multiple statistical models will be used, it should be specified which of these is considered the primary analysis model. If the statistical model will be chosen based on characteristics of the data, an objective way of choosing the final model (which does not allow analysts to choose the model that provides the most favorable result) should be specified.
Adjustment for covariates	Whether covariate adjustment will be used, and if so, the set of covariates to be included, and the method of including each one.	The set of covariates to be used for adjustment should be specified (if covariate adjustment is to be used). The method of including each covariate should also be specified; for instance, continuous covariates could be dichotomized or included as a continuous variable. If multiple sets of covariates will be used (eg, both an adjusted and unadjusted analysis will be performed), it should be specified which of these is considered the primary analysis. If the set of covariates will be chosen based on characteristics of the data, an objective way of choosing (which does not allow analysts to choose the set of covariates that provides the most favorable result) should be specified.
Handling of missing data	The approach that will be used for patients with missing outcome data.	The method of handling missing outcome data should be specified, with sufficient detail to allow replication. For instance, if multiple imputation were to be used for the analysis, the imputation model, method of imputing each variable, number of imputations, and method for combining results across imputed data sets should be specified, alongside any other important details. If patients with missing outcomes will be excluded from the analysis, this should be explicitly stated, rather than assumed. If multiple approaches will be used (eg, both a complete case analysis and multiple imputation), it should be specified which of these is considered as the primary analysis. If the approach will be chosen based on characteristics of the data, an objective way of choosing (which does not allow analysts to choose the method that provides the most favorable result) should be specified. For instance, under multiple imputation, the analyst may wish to include variables associated with missingness and/or the outcome in the imputation model; if this is not known in advance and will be estimated from the data, an objective approach for choosing variables to include in the imputation model should be given.

shown from the reference method. We varied the analysis population [20−22] (the set of participants included in the analysis), the statistical model used, whether the analysis was adjusted for baseline covariates [23−25], and the method of handling missing data [26−28].

We note that this analysis is exploratory and is only intended to assess the discrepancy in results that could have occurred in these two trials, had the investigators chosen the analysis method based on the data to obtain a desired result; it is not intended to reflect common statistical practice or be generalizable across all trials.

2.2.1. OPTIMISE

The Optimisation of Cardiovascular Management to Improve Surgical Outcome (OPTIMISE) trial was a multicenter randomized controlled trial comparing a cardiac output-guided hemodynamic therapy algorithm plus dopexamine to usual care for high-risk participants undergoing major gastrointestinal surgery [29]. The primary outcome was a composite outcome of predefined 30-day moderate or major complications and mortality.

The reference analysis was based on the intention-to-treat population where all randomized participants with a recorded outcome were included and analyzed according to the group they were randomized to. We used a logistic regression model, unadjusted for any baseline covariates. Participants with missing outcome data were excluded from the analysis. Based on this reference analysis, we then varied the analysis population, statistical model, and use of baseline covariates in turn. We did not vary the method of handling missing data because only 4/734 (0.5%) participants were missing the primary outcome.

First, we altered the analysis population by implementing a per-protocol analysis, where participants who were given an incorrect dopexamine dose after surgery were excluded from the analysis. We then varied the statistical model by using Fisher's exact test instead of a logistic regression model. Finally, we varied the use of baseline covariates by using a logistic regression model adjusted for three different sets of covariates. The model was first adjusted for the subset of covariates that were prespecified in the OPTIMISE protocol. The second set of covariates was chosen by selectively adding covariates to the model and only keeping those that reduced the size of the *P*-value. The third set of covariates was chosen in a similar manner, but only covariates that increased the size of the *P*-value were kept.

2.2.2. TRIGGER

The Transfusion in Gastrointestinal Bleeding (TRIGGER) trial was a feasibility, cluster-randomized controlled trial comparing two blood transfusion strategies in participants admitted to hospital with acute upper gastrointestinal bleeding [30,31]. This was a feasibility trial and not intended to make claims regarding effectiveness of the interventions under study. However, we include it here as it makes for an interesting case study on the impact of the choice of analysis on results in certain situations. Because this was a feasibility trial, the main outcome measures were of a feasibility nature. However, the primary clinical outcome was specified as further bleeding up to day 28, and that is what we use here.

The reference analysis was based on the intention-to-treat population where all enrolled participants with a recorded outcome were included and analyzed according to their allocated treatment group. We used a logistic regression model, with generalized estimating equations with an exchangeable correlation structure across hospitals and robust standard errors [32]. The model was unadjusted for any covariates, and participants with missing outcome data were excluded from the analysis. We note that this reference analysis is different to the analysis approach for the trial, which was based on a cluster-level summary approach in each cluster; this was because a reference approach based on analyzing individual-level data (rather than aggregate cluster-level summaries) would make it easier to vary different aspects of the analysis, such as the analysis model, or the use of covariate adjustment.

Based on reference analysis listed previously, we then varied the analysis population, statistical model, use of baseline covariates, and handling of missing data in turn. We varied the analysis population by implementing a per-protocol analysis, where participants who received a transfusion against protocol guidelines were excluded from the analysis. We then changed the statistical model by using a mixed-effects model with a random intercept for hospital. We varied the use of baseline covariates by using a logistic regression model with generalized estimating equations adjusted for three different sets of covariates; these three sets of covariates were chosen in the same way as for the OPTIMISE trial. Finally, we varied the method of handling missing outcome data by considering a scenario where participants who were lost to follow-up were imputed as having no event (ie, not experiencing further bleeding) and a scenario where those participants were imputed as having an event (ie, experiencing further bleeding). Overall, 31/936 patients (3%) had missing outcome data.

3. Results

3.1. Review of published protocols

Our search identified 277 articles, 178 of which were ineligible and thus excluded. This left 99 eligible protocols that were included (Fig. 1).

Results are shown in Tables 2 and 3. Most articles did not adequately predefine any of the four aspects of their statistical analysis approach (39/99, 39%) or predefined only one aspect (36/99, 36%). None of the trials adequately predefined all aspects of the analysis.

Only 8/99 (8%) of protocols adequately predefined the analysis population, and 10/99 (10%) adequately predefined their approach to handling missing data. Pre-specification of

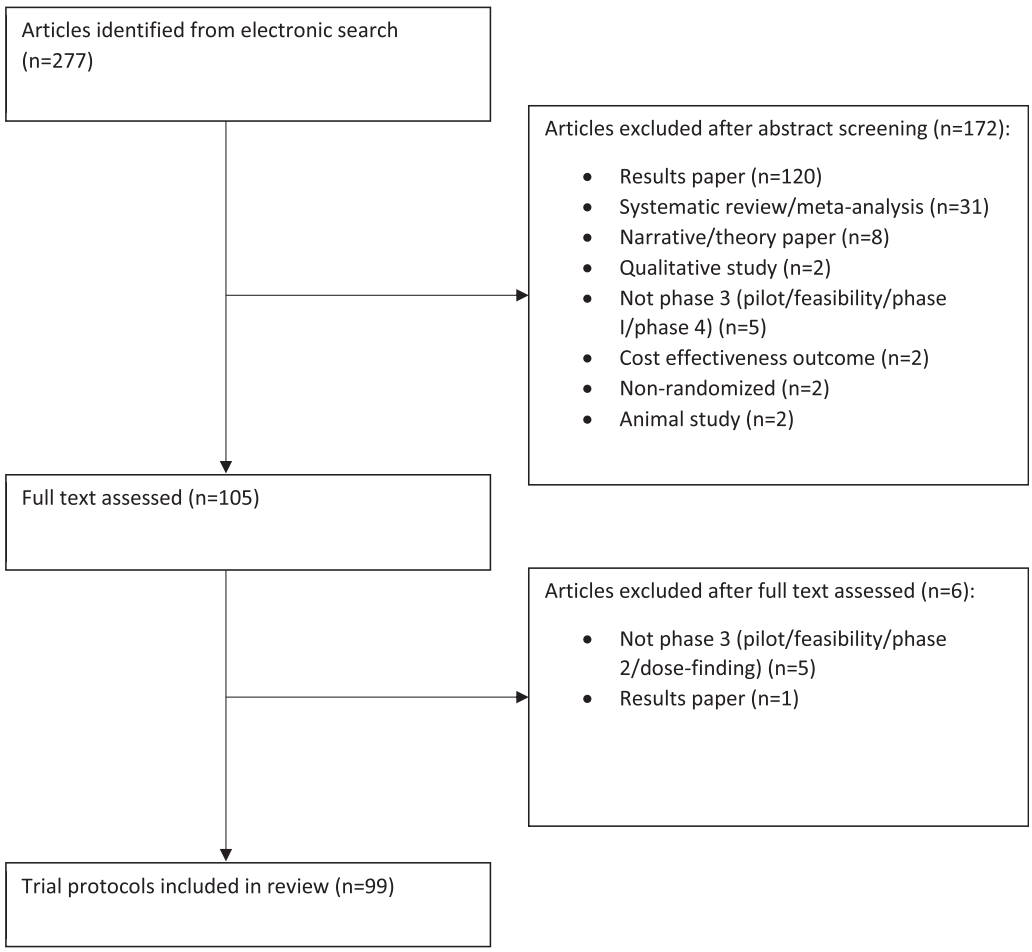

Fig. 1. Flow diagram of protocol selection.

the analysis model (27/99, 27%) and use of covariates (40/99, 40%) was higher, but still insufficient.

Many of the protocols that did not adequately predefine their analysis approach did mention the relevant aspect in the article but did not provide sufficient explanation, for example, stating that a per-protocol analysis would be undertaken with no explanation of which patients would be specifically excluded. The one exception was the handling of missing data, which most articles did not mention.

3.2. Exploratory re-analysis of the OPTIMISE and TRIGGER trials

Results are shown in Table 4. Re-analysis of the OPTIMISE trial found that the choice of statistical analysis approach had little impact on the estimated treatment effect (range odds ratio [OR] 0.73–0.78) but had a large impact on the significance of the results (range *P*-value 0.045–0.10). It was possible to obtain both significant and nonsignificant results by varying either the patient population included or the set of covariates used in the analysis model.

Re-analysis of the TRIGGER trial found that the estimated OR ranged from 0.45 to 1.09 across the different analytical approaches, while the *P*-value ranged from $<$ 0.0001 to 0.80. Changes to the analysis population included led to a large change in the estimated OR (from 0.52 to 1.09), whereas varying the set of covariates included in the model led to *P*-values between $<$0.0001 and 0.49. It was possible to obtain both significant and nonsignificant results by varying either the set of covariates used in the analysis model or the method of handling missing data.

4. Discussion

Results from randomized trials can depend on the statistical methods used to analyze the data. If the data are analyzed in multiple ways, investigators may only present the most favorable results, which can provide a distorted view of the evidence. It is therefore important that the analysis approach is prespecified in sufficient detail to prevent selective reporting of results based on different analysis methods [2–12]. Despite recommendations from both the ICH-E9 and SPIRIT guidelines, our review of published protocols found that very few trial protocols adequately prespecified the method of analysis for the primary outcome. None of the protocols adequately predefined all

aspects of the analysis, and most either did not predefine any aspect (39%) or predefined only one aspect (36%). The analysis population was particularly poorly reported. Furthermore, the method for handling missing data was rarely mentioned [26,28].

Our exploratory re-analysis of two published trials found that the specific analysis approach used impacted the study results. For both trials, it was possible to obtain both significant and nonsignificant results, and in one trial, the estimated odds ratio varied between 0.45 and 1.09 depending on the method of analysis.

Our review had some limitations. We only included published protocols indexed on PubMed. Protocols that are not published or those appearing in journals not indexed on PubMed may report the statistical analysis section differently. We excluded early-phase trials, such as phase II or feasibility trials, which are often used to determine whether a subsequent larger trial will take place. Lack of a prespecified statistical analysis approach may make these trials look more promising than they are (or hide harms), which could divert resources into follow-up trials that are unlikely to show treatment benefit.

Table 2. Number of protocols adequately predefining each aspect of the statistical analysis approach for the primary outcome

Analysis element	Number (%)
Analysis population	
Adequately predefined	8 (8)
Incompletely predefined[a]	64 (65)
Not mentioned	27 (27)
Analysis model	
Adequately predefined	27 (27)
Incompletely predefined[a]	61 (62)
Not mentioned	11 (11)
Adjustment for covariates	
Adequately predefined	40 (40)
Incompletely predefined[a]	32 (32)
Not mentioned	27 (27)
Handling of missing data	
Adequately predefined	10 (10)
Incompletely predefined[a]	24 (24)
Not mentioned	65 (66)
Number of aspects adequately predefined	
0	39 (39)
1	36 (36)
2	23 (23)
3	1 (1)
4	0 (0)

Elements were classified as adequately predefined if there was sufficient detail to allow replication by a third party and could not allow the user to choose the approach subjectively based on the trial data. Elements were classified as incompletely predefined if some detail was included but not enough to allow replication by a third party. Elements were classified as not mentioned if they were not addressed in the text.

[a] Reasons that elements were incompletely prespecified are listed in Table 3.

Table 3. Reasons for incomplete pre-specification of each element of the statistical analysis approach for the primary outcome [number (%)]

Reason for incomplete pre-specification[a]	Protocols
Analysis population incompletely predefined (n = 64)	
Listed multiple populations to be analyzed, but not which is primary	11 (17)
Did not adequately define the population	64 (100)
No definition	42/64 (66)
Omitted essential details	22/64 (34)
Analysis model incompletely predefined (n = 61)	
Specified the model but did not provide enough detail to replicate the analysis	42 (69)
Listed multiple models to be used but did not specify which would be the primary	11 (18)
Specified a list of potential models they would pick from but did not provide an objective way of choosing	19 (31)
Other[b]	4 (7)
Covariates incompletely predefined (n = 32)	
Specified covariate adjustment but did not list all covariates which would be adjusted for	23 (72)
Did not list any of the covariates to be adjusted for	9/23 (40)
Listed examples of covariates that may be included in the model but did not provide an exhaustive list	14/23 (61)
Specified covariate adjustment but did not include adequate detail of an algorithm for inclusion of the covariates	17 (53)
Specified a list of potential covariates they would pick from but did not specify how these would be selected for inclusion	9/17 (52)
Specified a list of potential covariates they would pick from but used subjective criteria for inclusion	8/17 (47)
Other[c]	11 (34)
Handling of missing data incompletely predefined (n = 24)	
Specified they will choose a method to handle missing data but did not provide detail for choosing the method	3 (13)
Multiple imputation specified but did not provide enough detail to replicate the analysis	15 (63)
Other[d]	9 (38)

[a] Categories are not mutually exclusive.

[b] Unclear which analysis model relates to the primary outcome (n = 4).

[c] Listed multiple sets of covariates to adjust for but did not specify which set would be the primary (n = 9); unclear description of approach (n = 2).

[d] Stated missing data would be accounted for in the analysis but did not say how (n = 4); stated inverse probability weighting would be used, but provided no further details (n = 2); listed multiple methods of handling missing data but did not specify which would be the primary (n = 2); stated multiple imputation would be used as a sensitivity analysis but did not state what would be used for the primary analysis (n = 1).

Table 4. Re-analysis of the OPTIMISE and TRIGGER results

	OPTIMISE		TRIGGER	
Analysis method	**Odds ratio (95% CI)**	***P*-value**	**Odds ratio (95% CI)**	***P*-value**
Analysis population				
ITT[a]	0.75 (0.56, 1.01)	0.061	0.52 (0.25, 1.08)	0.08
PP[b]	0.73 (0.54, 0.99)	0.045	1.09 (0.56, 2.10)	0.80
Analysis model				
Model 1[a,c]	0.75 (0.56, 1.01)	0.061	0.52 (0.25, 1.08)	0.08
Model 2[d]	–	0.070	0.53 (0.23, 1.20)	0.13
Adjustment for covariates				
Unadjusted[a]	0.75 (0.56, 1.01)	0.061	0.52 (0.25, 1.08)	0.08
Adjusted for prespecified variables[e]	0.73 (0.53, 1.00)	0.052	0.50 (0.32, 0.78)	0.002
Adjustment based on data (v1)[f]	0.74 (0.55, 1.00)	0.048	0.45 (0.31, 0.65)	<0.0001
Adjustment based on data (v2)[g]	0.78 (0.58, 1.05)	0.10	0.73 (0.30, 1.78)	0.49
Handling of missing data				
Complete case[a]	–	–	0.52 (0.25, 1.08)	0.08
Imputed as no event	–	–	0.53 (0.26, 1.10)	0.09
Imputed as an event	–	–	0.55 (0.32, 0.93)	0.03

Abbreviations: ITT, intention-to-treat; PP, per-protocol.
[a] Denotes the reference analysis.
[b] OPTIMISE: participants excluded if dopexamine dose given incorrectly; TRIGGER: participants excluded if they were transfused against protocol guidelines.
[c] OPTIMISE: logistic regression, TRIGGER: generalized estimating equations.
[d] OPTIMISE: Fisher's exact test, TRIGGER: mixed-effects model.
[e] Prespecified in the trial protocol.
[f] Chosen to minimize the *P*-value.
[g] Chosen to maximize the *P*-value.

Furthermore, it is possible that some investigators opted to prespecify the statistical analysis approach in an SAP rather than in the trial protocol. The key principle to maintaining trial integrity is that the statistical analysis approach is fully prespecified before the trial begins and made publicly available and that any changes to the analysis approach are documented. It can be argued that, provided the SAP fulfills these requirements, the protocol itself does not need to contain sufficient detail on the planned analysis approach for the primary outcome.

However, there are some limitations with this approach in practice. First, in some cases, the SAP may only be completed after the trial has begun. Investigators may therefore change the approach specified in the SAP based on early looks at the data. Second, SAPs are rarely publicly accessible to readers because they are infrequently published in their own right. They are sometimes required to be submitted to journals or regulatory agencies; however, they are not always released alongside the trial results. We also note that many of the issues we found in our review were unlikely to be resolved by the use of SAPs. For instance, many issues were based on inadequate explanations of what was intended (eg, an inadequate explanation of which patients would be excluded from a per-protocol analysis), or investigators listing multiple analysis approaches without specifying which was the primary (eg, listing both intention-to-treat and per-protocol without stating which was primary).

We therefore suggest that investigators should follow the SPIRIT and ICH-E9 guidelines and prespecify their statistical analysis approach in the trial protocol. Furthermore, prospectively registering the analysis approach on a trial registry web site or another publicly available independent platform before the trial begins would ensure this information is publicly available, regardless of whether the protocol or SAP has been published, thereby allowing a comparison between the planned and final analysis approach [13].

5. Conclusions

The method of analysis can have a large impact on the results of a trial. However, the statistical analysis approach is rarely prespecified in trial protocols in the detail recommended by guidelines. Investigators should routinely prespecify the analysis methods to be used in the trial protocol to prevent issues such as selective reporting of results based on different analyses.

Acknowledgments

Authors' contributions: L.G. developed review, developed search strategy, performed search, performed data extraction, performed statistical analyses, wrote first draft of the article. V.J. contributed to statistical analyses and

contributed to the article. R.P. contributed to statistical analyses and contributed to the article. B.C.K. developed the concept, developed review, developed search strategy, performed data extraction, and contributed to the article.

References

[1] Chan A, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. Spirit 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013;158:200−7.

[2] Chan AW, Tetzlaff JM, Gotzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ 2013;346:e7586.

[3] Chan A, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA 2004;291:2457−65.

[4] Dwan K, Altman DG, Cresswell L, Blundell M, Gamble CL, Williamson PR. Comparison of protocols and registry entries to published reports for randomised controlled trials. Cochrane Database Syst Rev 2011;(1):MR000031.

[5] Contopoulos-Ioannidis DG, Alexiou GA, Gouvias TC, Ioannidis JPA. An empirical evaluation of multifarious outcomes in pharmacogenetics: beta-2 adrenoceptor gene polymorphisms in asthma treatment. Pharmacogenet Genomics 2006;16(10):705−11.

[6] Chan A-W, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. BMJ 2005;330:753.

[7] Dwan K, Altman DG, Arnaiz JA, Bloom J, Chan A-W, Cronin E, et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. PLoS One 2008;3:e3081.

[8] Hahn S, Williamson PR, Hutton JL. Investigation of within-study selective reporting in clinical research: follow-up of applications submitted to a local research ethics committee. J Eval Clin Pract 2002;8:353−9.

[9] Ramagopalan S, Skingsley AP, Handunnetthi L, Klingel M, Magnus D, Pakpoor J, et al. Prevalence of primary outcome changes in clinical trials registered on ClinicalTrials.gov: a cross-sectional study. F1000Res 2014;3:77.

[10] Rising K, Bacchetti P, Bero L. Reporting bias in drug trials submitted to the Food and Drug Administration: review of publication and presentation. PLoS Med 2008;5(11):e217.

[11] Vedula SS, Bero L, Scherer RW, Dickersin K. Outcome reporting in industry-sponsored trials of gabapentin for off-label use. N Engl J Med 2009;361(20):1963−71.

[12] Williamson PR, Gamble C, Altman DG, Hutton JL. Outcome selection bias in meta-analysis. Stat Methods Med Res 2005;14(5):515−24.

[13] Grant S, Booth M, Khodyakov D. Lack of pre-registered analysis plan allows unacceptable data mining for and selective reporting of consensus in Delphi studies. J Clin Epidemiol 2018;99:96−105.

[14] Chan AW, Hrobjartsson A, Jorgensen KJ, Gotzsche PC, Altman DG. Discrepancies in sample size calculations and data analyses reported in randomised trials: comparison of publications with protocols. BMJ 2008;337:a2299.

[15] Dwan K, Altman DG, Clarke M, Gamble C, Higgins JP, Sterne JA, et al. Evidence for the selective reporting of analyses and discrepancies in clinical trials: a systematic review of cohort studies of clinical trials. PLoS Med 2014;11(6):e1001666.

[16] Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Dore C, et al. Guidelines for the content of statistical analysis plans in clinical trials. JAMA 2017;318:2337−43.

[17] Dworkin JD, McKeown A, Farrar JT, Gilron I, Hunsinger M, Kerns RD, et al. Deficiencies in reporting of statistical methodology in recent randomized trials of nonpharmacologic pain treatments: ACTTION systematic review. J Clin Epidemiol 2016;72:56−65.

[18] Kahan BC, Jairath V, Murphy MF, Doré CJ. Update on the transfusion in gastrointestinal bleeding (TRIGGER) trial: statistical analysis plan for a cluster-randomised feasibility trial. Trials 2013;14(1):206.

[19] ICH Harmonised Tripartite Guideline. Statistical principles for clinical trials. International Conference on Harmonisation E9 Expert Working Group. Stat Med 1999;18:1905−42.

[20] Abraha I, Cherubini A, Cozzolino F, De Florio R, Luchetta ML, Rimland JM, et al. Deviation from intention to treat analysis in randomised trials and treatment effect estimates: meta-epidemiological study. BMJ 2015;350:h2445.

[21] Matilde Sanchez M, Chen X. Choosing the analysis population in non-inferiority studies: per protocol or intent-to-treat. Stat Med 2006;25:1169−81.

[22] Nüesch E, Trelle S, Reichenbach S, Rutjes AWS, Bürgi E, Scherer M, et al. The effects of excluding patients from the analysis in randomised controlled trials: meta-epidemiological study. BMJ 2009;339:b3244.

[23] Raab GM, Day S, Sales J. How to select covariates to include in the analysis of a clinical trial. Controlled Clin Trials 2000;21:330−42.

[24] Kahan BC, Jairath V, Doré CJ, Morris TP. The risks and rewards of covariate adjustment in randomized trials: an assessment of 12 outcomes from 8 studies. Trials 2014;15(1):139.

[25] Kahan BC, Rushton H, Morris TP, Daniel RM. A comparison of methods to adjust for continuous covariates in the analysis of randomised trials. BMC Med Res Methodol 2016;16:42.

[26] 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.

[27] Akl EA, Briel M, You JJ, Sun X, Johnston BC, Busse JW, et al. Potential impact on estimated treatment effects of information lost to follow-up in randomised controlled trials (LOST-IT): systematic review. BMJ 2012;344:e2809.

[28] Saha C, Jones MP. Bias in the last observation carried forward method under informative dropout. J Stat Plann Inference 2009;139(2):246−55.

[29] Pearse RM, Harrison DA, MacDonald N, Gillies MA, Blunt M, Ackland G, et al. Effect of a perioperative, cardiac output−guided hemodynamic therapy algorithm on outcomes following major gastrointestinal surgery: a randomized clinical trial and systematic review. JAMA 2014;311:2181−90.

[30] Jairath V, Kahan BC, Gray A, Doré CJ, Mora A, James MW, et al. Restrictive versus liberal blood transfusion for acute upper gastrointestinal bleeding (TRIGGER): a pragmatic, open-label, cluster randomised feasibility trial. Lancet 2015;386:137−44.

[31] Kahan BC, Dore CJ, Murphy MF, Jairath V. Bias was reduced in an open-label trial through the removal of subjective elements from the outcome definition. J Clin Epidemiol 2016;77:38−43.

[32] Liang K-Y, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika 1986;73(1):13−22.