

1 **Evidence of unexplained discrepancies between planned and conducted statistical**
2 **analyses: a review of randomized trials**

3

4 Suzie Cro*, Gordon Forbes, Nicholas A Johnson, Brennan C Kahan

5

6 *Correspondence to: Suzie Cro, Imperial Clinical Trials Unit, Imperial College London, 1st

7 Floor, Stadium House, London, UK, W12 7RH, email: s.cro@imperial.ac.uk.

8

9 Suzie Cro, Research Fellow (Statistician), Imperial Clinical Trials Unit, School of Public

10 Health, Imperial College London, London, UK.

11

12 Gordon Forbes, Statistician, Department of Biostatistics and Health Informatics, Kings

13 College London, London, UK.

14

15 Nicholas A Johnson, Statistician, Imperial Clinical Trials Unit, School of Public Health,

16 Imperial College London, London, UK.

17

18 Brennan C Kahan, Senior Research Fellow, MRC Clinical Trials Unit at UCL, UK.

19

20

21

22 **Short title:** Discrepancies between planned and conducted trial analyses

23

24

25

26

27

28 **NOTE:** This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

29 **Abstract**

30 **Background:** Choosing or altering the planned statistical analysis approach after
31 examination of trial data (often referred to as ‘p-hacking’) can bias results of randomized
32 trials. However, the extent of this issue in practice is currently unclear. We conducted a
33 review of published randomized trials to evaluate how often a pre-specified analysis
34 approach is publicly available, and how often the planned analysis is changed.

35

36 **Methods:** A review of randomised trials published between January and April 2018 in six
37 leading general medical journals. For each trial we established whether a pre-specified
38 analysis approach was publicly available in a protocol or statistical analysis plan, and
39 compared this to the trial publication.

40

41 **Results:** Overall, 89 of 101 eligible trials (88%) had a publicly available pre-specified
42 analysis approach. Only 22/89 trials (25%) had no unexplained discrepancies between the
43 pre-specified and conducted analysis. Fifty-four trials (61%) had one or more unexplained
44 discrepancies, and in 13 trials (15%) it was impossible to ascertain whether any unexplained
45 discrepancies occurred due to incomplete reporting of the statistical methods. Unexplained
46 discrepancies were most common for the analysis model (n=31, 35%) and analysis
47 population (n=28, 31%), followed by the use of covariates (n=23, 26%) and the approach for
48 handling missing data (n=16, 18%). Many protocols or statistical analysis plans were dated
49 after the trial had begun, so earlier discrepancies may have been missed.

50

51 **Conclusions:** Unexplained discrepancies in the statistical methods of randomized trials are
52 common. Increased transparency is required for proper evaluation of results.

53

54 **Keywords:** Statistical Analysis, Randomised Controlled Trials, Transparency, Statistical
55 Analysis Plan, P-hacking.

56 **Background**

57

58 The results of a clinical trial depend upon the statistical methods used for analysis. For
59 example, changing the analysis population or method of handling missing data can change
60 the size of the estimated treatment effect or its standard error. In some instances these
61 differences can be large and may affect interpretation of the trial (1-6). If investigators
62 choose the method of analysis based on trial data in order to obtain more favourable results
63 (often referred to as 'p-hacking'), this can cause bias (7). Selective reporting has been
64 identified previously, where outcomes with more favourable results are more likely to be
65 reported than other outcomes (8-19). There is some evidence to suggest this may also be a
66 concern for statistical analyses; pre-specification of the proposed methods is often poor,
67 discrepancies between protocols and publications are common, and in some instances
68 changes may have been made to obtain specific results (5, 8, 10, 13, 20-23).

69

70 Guidelines such as ICH-E9 (24) (International Conference for Harmonisation of Technical
71 Requirements for Pharmaceuticals for Human Use), SPIRIT (25, 26) (Standard Protocol
72 Items: Recommendations for Interventional Trials), and CONSORT(27) (Consolidated
73 Standards of Reporting Trials) require investigators to pre-specify the principle features of
74 their statistical analysis approach in the trial protocol, and report any changes in the trial
75 report. This strategy can reduce bias from analysis being chosen based on trial data, and
76 allows readers to assess whether inappropriate changes were made.

77

78 We conducted a review of trials published in general medical journals to evaluate how often
79 a pre-specified analysis approach was publicly available, how often the planned analysis
80 approach was changed, whether these changes were explained, and the reporting around
81 the timing and blinding status of changes.

82

83 **Methods**

84 ***Search strategy***

85

86 In this review, we examined randomized controlled trials published between January and
87 April 2018 in six general high impact medical journals: Annals of Internal Medicine; The BMJ;
88 Journal of the American Medical Association (JAMA); The Lancet; New England Journal of
89 Medicine (NEJM); and PLOS Medicine. We searched for articles in PubMed with a
90 publication type of “randomized controlled trial” or categorised with the MeSH term “random
91 allocation,” or including the keyword “random*” in the title or abstract, restricted to the
92 aforementioned included journals and publication period. The full search strategy is shown in
93 Appendix 1 in Additional File 1 and was conducted July 2018.

94

95 ***Eligibility***

96

97 Articles were eligible for inclusion if they reported results from a phase 2-4 randomized trial
98 in humans. Exclusion criteria were pilot or feasibility study, phase 1 trial, non-randomized
99 study, secondary analysis of previously published trial, cost-effectiveness as the primary
100 outcome, more than one trial reported in the article, results of an interim analysis, or if the
101 protocol or SAP was not in English.

102 One author screened the title and abstract of each paper for eligibility. The full texts of these
103 articles were then assessed independently by two reviewers to confirm eligibility. For all
104 eligible studies, one author searched the main text, supplementary material, and references
105 to identify whether a protocol and/or SAP was available.

106

107 ***Data extraction***

108

109 Data was extracted onto a pre-piloted standardised data extraction form by two reviewers
110 independently. Disagreements were resolved by discussion, or by a third reviewer where
111 disagreement could not be resolved. Where the trial publication referred to supplementary
112 material, a SAP or protocol, the extractor referred to these documents.

113

114 We extracted data related to the primary analysis of the primary outcome from the trial
115 publication. A single primary outcome was identified as follows; (a) if one outcome was listed
116 as the primary we used this; (b) if no outcomes or multiple outcomes were listed as being
117 primary we used the outcome that the sample size calculation was based on; and (c) if no
118 sample size calculation was performed or sample size was calculated for multiple primary
119 outcomes, we used the first clinical outcome listed in the objectives/outcomes section. We
120 identified the primary analysis as follows; (a) if a single analysis strategy was used, or
121 multiple strategies were used with one being identified as primary, we used this; (b) if
122 multiple strategies were used without one being identified as primary, we used the first one
123 presented in the results section.

124

125 For each article, we extracted general trial characteristics, whether protocols or SAPs were
126 available, including the dates of these documents and, if available, the blinding status of trial
127 statisticians. For articles with a protocol or SAP, we compared the method of analysis in the
128 trial publication against the method specified in the earliest available protocol or SAP which
129 included some information on the analysis of the primary outcome (referred to as the original
130 analysis plan). We assessed the following four analysis elements: (i) analysis population (the
131 set of participants included in the analysis, and which treatment group they were analysed
132 in); (ii) the statistical analysis model; (iii) use of baseline covariates in the analysis; and (iv)
133 the method for handling missing data. We chose these elements as they are specified in the
134 SPIRIT guidelines, and have been used in previous reviews (5, 25).

135

136 We evaluated two types of discrepancies for each analysis element. The first, termed a
137 ‘change’, occurred when the analysis element in the trial publication was different to that
138 specified in the original analysis plan. The following examples would constitute changes: (a)
139 if an intention-to-treat analysis population was originally specified, but a per-protocol analysis
140 was used; (b) if the functional form of the statistical analysis model was changed, such as
141 from a mixed-effects regression model to generalized estimating equations (GEE); (c) if the
142 original analysis plan specified the analysis would not adjust for baseline covariates but the
143 trial publication adjusted for one or more patient characteristic; or (d) if a complete case
144 analysis was originally specified, but multiple imputation was used.

145

146 The second discrepancy, termed an ‘addition’, occurred when the original analysis plan gave
147 the investigators flexibility to subjectively choose the final analysis method after seeing trial
148 data. This could occur if the original analysis plan (i) contained insufficient information about
149 the proposed analysis; or (ii) allowed the investigators to subjectively choose between
150 multiple different potential analyses. The following examples would constitute additions: if
151 the original analysis plan stated that (a) both a per-protocol and intention-to-treat analysis
152 population would be used, without specifying which was the primary analysis (as
153 investigators could then decide during final analysis which was the primary, based on which
154 gave the most favourable result); (b) either parametric or non-parametric methods would be
155 used depending on distributional assumptions, but did not define an objective criteria for
156 assessing distributional assumptions (as the investigators could then present whichever
157 method gave the most favourable result); (c) the analysis would adjust for important baseline
158 covariates, but did not define how these covariates would be chosen (as investigators could
159 choose during final analysis the set of covariates which gave the most favourable result); or
160 (d) multiple imputation would be used, but did not define what the method of imputation
161 would be, or what variables would be included in the imputation model (as this would allow
162 the investigators to run several different imputation models during final analysis and present
163 only the most favourable).

164 We classified each discrepancy as being 'explained' or 'unexplained'. Discrepancies were
165 classified as explained if they had been specified in a subsequent version of the protocol or
166 SAP (with or without a justification or rationale for the discrepancy), or if the trial publication
167 explained that an alteration to the pre-specified analysis approach had been made.
168 Otherwise discrepancies were classified as unexplained.

169

170 ***Outcomes***

171

172 The main outcome measures were (i) the number of trials with a publicly available pre-
173 specified analysis approach for the primary outcome (i.e. whether an original analysis plan
174 was available in a protocol or a SAP); (ii) the number of trials with no unexplained
175 discrepancies from the publicly available pre-specified analysis approach; and (iii) the total
176 number of analysis elements for each trial with an unexplained discrepancy.

177

178 Secondary outcomes were, for each analysis element described earlier, (i) the number of
179 trials with at least one unexplained discrepancy (either change or addition); (ii) the number of
180 trials with at least one unexplained change; and (iii) the number of trials with at least one
181 unexplained addition.

182

183 ***Statistical methods***

184

185 Outcomes were summarised descriptively using frequencies and percentages. We
186 performed two pre-specified subgroup analyses, where we summarised outcomes
187 separately according to trial funding status, and type of intervention. One post-hoc subgroup
188 analysis was performed according to availability of a SAP.

189

190

191 All statistical analyses were performed using Stata version 15 (28).

192

193 Results

194

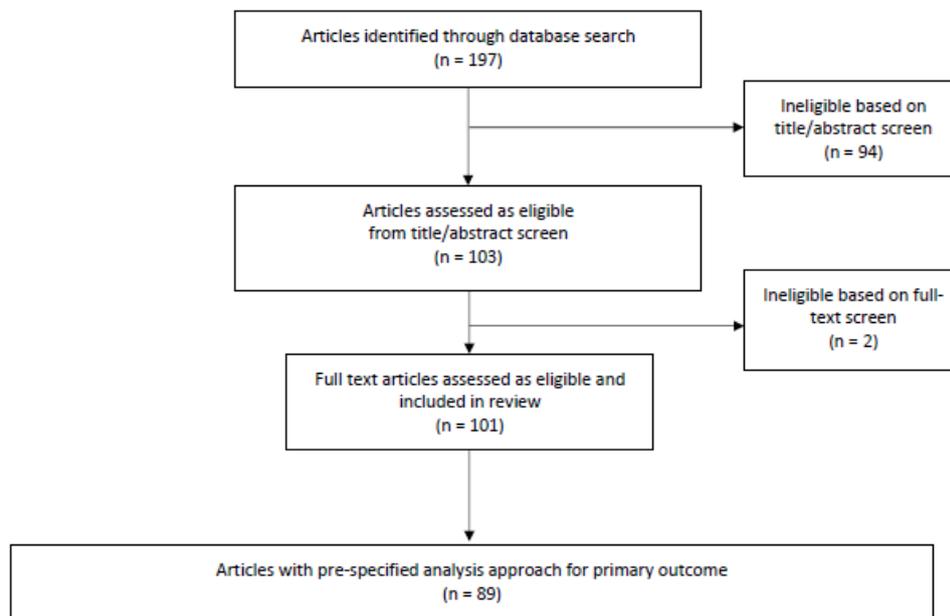
195 *Search results and characteristics of included studies*

196

197 Our search identified 197 articles, of which 101 were eligible (see Fig 1 and for a list of
198 eligible trials Appendix 2 in Additional File 1). General trial characteristics are shown in Table
199 1.

200

201 **Table 1 – Characteristics of eligible trials (N=101)**



202

203 **Fig 1: Flow chart of article selection**

204

205 Protocols were available for 90 trials (89%) (48 published, 70 as supplementary material with
206 publication, 5 on a website). SAPs were available for 46 trials (46%) (3 published, 43 as
207 supplementary material with publication, 2 on a website). Of 90 trials with an available
208 protocol, the earliest version available was dated before recruitment began for 45 (50%)
209 trials, 19 (21%) were dated during recruitment, 8 (9%) were dated after recruitment ended,
210 and 18 (20%) did not have a date. Of 46 trial with an available SAP, the earliest version of
211 the SAP was dated before recruitment began for 9 (20%) trials, 13 (28%) were dated during
212 recruitment, 13 (28%) were dated after recruitment ended, and 11 (24%) did not have a
213 date.

214

215 Overall, only 11 trials (11%) stated in the trial publication, protocol, or SAP that the
216 statistician was blinded until the SAP was signed off and 10 (10%) stated the statistician was
217 blinded until the database was locked.

218

219 ***Availability of pre-specified analysis approach***

220

221 Overall, 89 of 101 trials (88%) had a publicly available pre-specified analysis approach for
222 the primary outcome. Eleven trials did not have an available protocol or SAP, and one trial
223 had a protocol with no information on the analysis and no SAP. The document containing the
224 original analysis plan (83 in a protocol, 6 in a SAP) was dated before the start of recruitment
225 for 41 of 89 (46%) trials, during recruitment in 19 (21%) trials (median 19 months post-
226 recruitment beginning, IQR 9 to 46), and after the end of recruitment in 8 (9%) trials (median
227 7 months post-recruitment completion, IQR 4 to 13). In 21 trials (24%) no date was available.

228

229 ***Comparison of pre-specified and conducted statistical analysis approach***

230

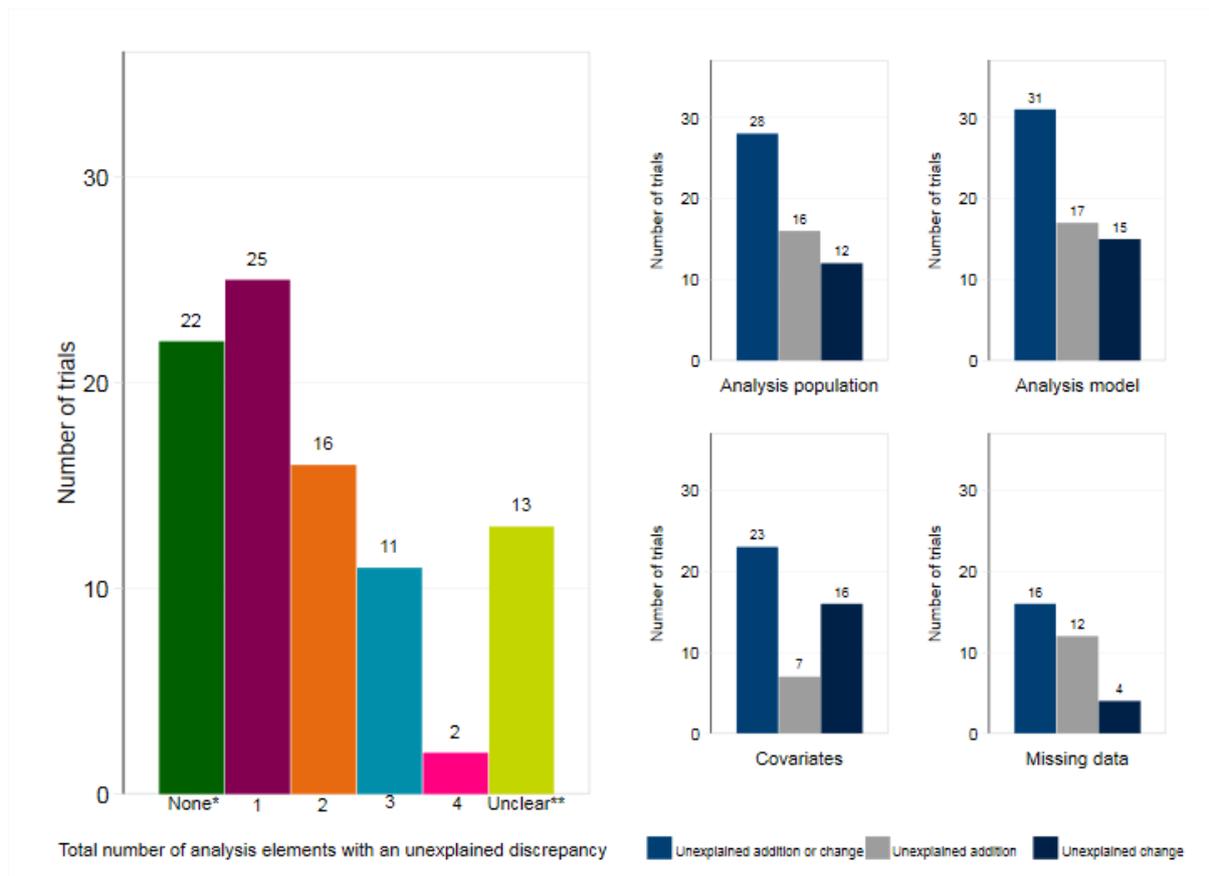
231 Of the 89 trials with an available pre-specified analysis approach, only 22 (25%) did not have
232 any unexplained discrepancies (no discrepancies n=5, explained discrepancies only n=17).

233 A further 54 trials (61%) had one or more unexplained discrepancies (see Fig 2). In 13 trials
234 (15%) it was unclear whether an unexplained discrepancy occurred due to poor reporting of
235 statistical methods (unclear whether discrepancy occurred n=11, unclear whether
236 discrepancy explained n=2).

237

238 Most trials had one (n=25, 28%) or two (n=16, 18%) unexplained discrepancies. Only 11
239 (12%) had three and 2 (2%) had four unexplained discrepancies. Unexplained discrepancies
240 were most common for the statistical analysis model (n=31, 35%) and analysis population
241 (n=28, 31%), followed by the use of covariates (n=23, 26%) and handling of missing data
242 (n=16, 18%). Table 2 provides a description of the unexplained discrepancies.

243



244

245 **Fig 2 – Number of trials with unexplained discrepancies (Total N=89)** *Of the n=22 trials
 246 with none; no discrepancies (n=5), explained discrepancies only (n=17). **Unclear if
 247 discrepancy occurred (n=11), unclear if discrepancy explained (n=2). One trial had both a
 248 change and an addition for the analysis model.

249

250

251 **Table 2 – Description of unexplained discrepancies (N=89)**

252

253

254 Overall, 29 trials (33%) had at least one explained discrepancy. Most discrepancies were
 255 explained in a later version of the protocol or SAP; only 2 trials explained a discrepancy in
 256 the trial publication. Of the 29 trials with an explained discrepancy, only 6 (21%) stated that

257 the statistician was blinded until the SAP was signed off, and 4 (14%) until the database was
258 locked.

259

260

261 Subgroup analyses

262 A total of 43/61 (66%) trials that were not for profit only had at least one unexplained
263 discrepancy, compared to 11/28 (45%) trials that were for profit only. Fewer trials with a SAP
264 available had unexplained discrepancies than trials without an available SAP, though this
265 figure was still high (SAP available 22/46 [48%] with ≥ 1 unexplained discrepancy vs. no SAP
266 32/43 [74%]). Trials with a SAP still had a relatively high number of additions to the analysis
267 method, indicating that methods were not being adequately pre-specified within these SAPs
268 (range 7-15% across analysis elements). See Additional File 1, Appendix 3 and 4 for
269 additional results.

270

271

272 Discussion

273

274 In our review of 101 trials published in high impact general medical journals, we found that
275 most had a pre-specified analysis approach for the primary outcome available in either a
276 protocol or SAP. This is essential to allow transparent assessment of whether inappropriate
277 changes were made to the statistical methods. However, most pre-specified statistical
278 analysis approaches were available in a document that was dated after the trial had begun,
279 or had no date available. It is therefore possible that the analysis approach in these
280 documents may have already been changed from the pre-trial version.

281

282 Only 25% of trials did not have any unexplained discrepancies between the trial publication
283 and the pre-specified analysis approach, and only 6% had no discrepancies at all. Most trials
284 had at least one unexplained discrepancy (61%), with 32% of trials having two or more. In
285 15% of trials it was impossible to assess whether there were unexplained discrepancies due
286 to poor reporting of the statistical methods used. Of note, 33% of trials had one or more
287 explained discrepancies; however, less than a quarter of these trials reported that the
288 statistician was blinded to treatment allocation until the analysis plan was finalised or the
289 database was locked. These alterations may therefore have been made based on unblinded
290 trial data, despite being explained. It was also surprising that only two trials explained a
291 discrepancy in the trial publication, despite requirements by the CONSORT (29) statement to
292 do so.

293

294 Our results are broadly consistent with previous reviews. Spence *et al* (30) evaluated the
295 availability of protocols and SAPs for trials published in high impact medical journals, and
296 found similar rates of availability. However, the rates of discrepancies we found were
297 generally lower than those previously reported (8, 10, 20, 21). For example, Chan *et al*
298 compared publications to protocols for 70 trials that received ethical approval by the
299 scientific-ethics committees for Copenhagen and Frederiksberg, Denmark in 1994-5 (21).
300 Overall, 44% of trials had unexplained discrepancies in the analysis population, 60% in the
301 analysis model, 82% in the use of covariates, and 80% for handling of missing data. There
302 are several potential explanations for these differences. The introduction of the SPIRIT
303 guidelines in 2013 (25, 26) may have led to better reporting of statistical methods in trial
304 protocols. We also accessed statistical analysis plans in almost half of trials, which
305 increased the number of explained discrepancies. Finally, we evaluated a different
306 population of trials; most of the high impact general medical journals in our review required

307 submission of the trial protocol alongside the article, and may have been less likely to accept
308 trials with extreme discrepancies.

309

310 The key issues we identified in this study were: (i) low availability of pre-trial protocols and
311 analysis plans; (ii) poor pre-specification of statistical methods within protocols and analysis
312 plans; (iii) frequent unexplained discrepancies in the final trial publication; (iv) poor reporting
313 of the blinding status of statisticians in relation to modifications of analysis methods or
314 access to trial data; and (v) poor descriptions of the actual analysis methods used in the final
315 publication. Increased adherence to guidelines such as SPIRIT, CONSORT, and the
316 guidelines for Statistical Analysis Plans (6, 26, 27) would help, though alternative
317 approaches to increase transparency around the statistical methods are also required. Two
318 simple proposals that would greatly improve the situation are (a) journals could require
319 authors to submit the first and last version of their protocol and SAP alongside the results
320 article, and publish these as supplementary material; this would allow transparent evaluation
321 of modifications to the analysis approach and be more effective than relying on authors to
322 publish these documents; and (b) journals could require that authors include the statistical
323 code used to perform their analysis alongside the article as supplementary content to allow a
324 complete and transparent comparison of the planned methods vs the final methods (31).

325

326 Our study had some limitations. We only included articles from six high impact medical
327 journals; it is likely that trials published in other journals may have lower availability of
328 protocols and SAPs, and higher rates of unexplained discrepancies. Comparisons were
329 based on the first available protocol or SAP, however many were dated after the trial had
330 begun, so there may have been discrepancies before this that we missed.

331

332

333 **Conclusions**

334 In conclusion, unexplained discrepancies in the statistical methods of randomized trials are
335 common. Increased transparency around the statistical methods used in randomized trials is
336 required for proper evaluation of trial results.

337

338

339 **List of abbreviations**

340

341 CONSORT Consolidated Standards of Reporting Trials

342 GEE Generalized Estimating Equations

343 IQR Interquartile Range

344 SAP Statistical Analysis Plan

345 SPIRIT Standard Protocol Items: Recommendations for Interventional Trials

346

347

348

349 **Declarations**

350

351 **Ethical approval and consent to participate**

352 No ethical approval or consent was required for this review of previously published trials.

353

354 **Consent for publication**

355 Not applicable

356

357 **Availability of data and materials**

358 The datasets used and/or analysed during the current study are available from the
359 corresponding author on reasonable request.

360

361 **Competing Interests**

362 The authors declare that they have no competing interests.

363

364 **Funding**

365 No specific funding was obtained for this research. Brennan Kahan is grateful for support
366 from the UK Medical Research Council, grant MC_UU_12023/21.

367

368 **Author contributions**

369 The corresponding author affirms that all listed authors meet authorship criteria and that no
370 others meeting the criteria have been omitted. SC and BK had access to all the data in the
371 study and take full responsibility for the work and the conduct of the study and controlled the
372 decision to publish.

373 Study concept: BK

374 Study design: BK, SC and GF

375 Acquisition and interpretation of data: All authors.

376 Statistical analysis: SC

377 Drafting of the manuscript: SC and BK

378 Critical revision of the manuscript for important intellectual content: All authors

379

380 **Acknowledgements**

381 Brennan Kahan is grateful for support from the UK Medical Research Council, grant

382 MC_UU_12023/21.

383

384 **References**

385

- 386 1. Prakash A, Risser RC, Mallinckrodt CH. The impact of analytic method on interpretation of
387 outcomes in longitudinal clinical trials. *International journal of clinical practice*. 2008;62(8):1147-58.
- 388 2. Saquib N, Saquib J, Ioannidis JPA. Practices and impact of primary outcome adjustment in
389 randomized controlled trials: meta-epidemiologic study. *BMJ : British Medical Journal*.
390 2013;347:f4313.
- 391 3. Porta N, Bonet C, Cobo E. Discordance between reported intention-to-treat and per protocol
392 analyses. *J Clin Epidemiol*. 2007;60(7):663-9.
- 393 4. Melander H, Ahlqvist-Rastad J, Meijer G, Beermann B. Evidence b(i)ased medicine--selective
394 reporting from studies sponsored by pharmaceutical industry: review of studies in new drug
395 applications. *Bmj*. 2003;326(7400):1171-3.
- 396 5. Greenberg L, Jairath V, Pearse R, Kahan BC. Pre-specification of statistical analysis
397 approaches in published clinical trial protocols was inadequate. *Journal of Clinical Epidemiology*.
398 2018;101:53-60.
- 399 6. Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Dore C, et al. Guidelines for the Content
400 of Statistical Analysis Plans in Clinical Trials. *Jama*. 2017;318(23):2337-43.
- 401 7. Page MJ, McKenzie JE, Forbes A. Many scenarios exist for selective inclusion and reporting of
402 results in randomized trials and systematic reviews. *Journal of Clinical Epidemiology*. 2013;66(5):524-
403 37.
- 404 8. Li G, Abbade LPF, Nwosu I, Jin Y, Leenus A, Maaz M, et al. A systematic review of
405 comparisons between protocols or registrations and full reports in primary biomedical research.
406 *BMC medical research methodology*. 2018;18(1):9-.
- 407 9. Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for
408 selective reporting of outcomes in randomized trials: comparison of protocols to published articles.
409 *Jama*. 2004;291(20):2457-65.
- 410 10. Dwan K, Altman DG, Cresswell L, Blundell M, Gamble CL, Williamson PR. Comparison of
411 protocols and registry entries to published reports for randomised controlled trials. *Cochrane*
412 *Database of Systematic Reviews*. 2011(1).
- 413 11. Chan AW, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed:
414 review of publications and survey of authors. *Bmj*. 2005;330(7494):753.
- 415 12. Dwan K, Altman DG, Arnaiz JA, Bloom J, Chan AW, Cronin E, et al. Systematic review of the
416 empirical evidence of study publication bias and outcome reporting bias. *PLoS one*. 2008;3(8):e3081.

- 417 13. Hahn S, Williamson PR, Hutton JL. Investigation of within-study selective reporting in clinical
418 research: follow-up of applications submitted to a local research ethics committee. *Journal of*
419 *evaluation in clinical practice*. 2002;8(3):353-9.
- 420 14. Ramagopalan S, Skingsley AP, Handunnetthi L, Klingel M, Magnus D, Pakpoor J, et al.
421 Prevalence of primary outcome changes in clinical trials registered on ClinicalTrials.gov: a cross-
422 sectional study. *F1000Research*. 2014;3:77.
- 423 15. Rising K, Bacchetti P, Bero L. Reporting Bias in Drug Trials Submitted to the Food and Drug
424 Administration: Review of Publication and Presentation. *PLOS Medicine*. 2008;5(11):e217.
- 425 16. Vedula SS, Bero L, Scherer RW, Dickersin K. Outcome Reporting in Industry-Sponsored Trials
426 of Gabapentin for Off-Label Use. *New England Journal of Medicine*. 2009;361(20):1963-71.
- 427 17. Williamson PR, Gamble C, Altman DG, Hutton JL. Outcome selection bias in meta-analysis.
428 *Statistical methods in medical research*. 2005;14(5):515-24.
- 429 18. Goldacre B, Drysdale H, Dale A, Milosevic I, Slade E, Hartley P, et al. COMPare: a prospective
430 cohort study correcting and monitoring 58 misreported trials in real time. *Trials*. 2019;20(1):118.
- 431 19. Chen T, Li C, Qin R, Wang Y, Yu D, Dodd J, et al. Comparison of Clinical Trial Changes in
432 Primary Outcome and Reported Intervention Effect Size Between Trial Registration and Publication.
433 *JAMA Network Open*. 2019;2(7):e197242-e.
- 434 20. Dwan K, Altman DG, Clarke M, Gamble C, Higgins JPT, Sterne JAC, et al. Evidence for the
435 Selective Reporting of Analyses and Discrepancies in Clinical Trials: A Systematic Review of Cohort
436 Studies of Clinical Trials. *PLOS Medicine*. 2014;11(6):e1001666.
- 437 21. Chan A-W, Hróbjartsson A, Jørgensen KJ, Gøtzsche PC, Altman DG. Discrepancies in sample
438 size calculations and data analyses reported in randomised trials: comparison of publications with
439 protocols. *BMJ*. 2008;337:a2299.
- 440 22. Editorial Review of Protocols for Clinical Trials. *New England Journal of Medicine*.
441 1990;323(19):1355-.
- 442 23. Powers JH, Dixon CA, Goldberger MJ. Voriconazole versus Liposomal Amphotericin B in
443 Patients with Neutropenia and Persistent Fever. *New England Journal of Medicine*. 2002;346(4):289-
444 90.
- 445 24. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for
446 Human Use. ICH Harmonised Tripartite Guideline: Statistical Principles for Clinical Trials E9.: London,
447 England: European Medicines Agency; 1998.
- 448 25. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gotzsche PC, Krleza-Jeric K, et al. SPIRIT 2013
449 statement: defining standard protocol items for clinical trials. *Ann Intern Med*. 2013;158(3):200-7.
- 450 26. Chan A-W, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013
451 explanation and elaboration: guidance for protocols of clinical trials. *BMJ : British Medical Journal*.
452 2013;346:e7586.
- 453 27. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT
454 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials.
455 *BMJ*. 2010;340:c869.
- 456 28. StataCorp. *Stata Statistical Software: Release 15*. College Station, TX: Stata Corp LLC.
457 2017.
- 458 29. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised
459 CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med*.
460 2001;134(8):663-94.
- 461 30. Spence OM, Hong K, Onwuchekwa Uba R, Doshi P. Availability of study protocols for
462 randomized trials published in high-impact medical journals: A cross-sectional analysis. *Clinical*
463 *Trials*.0(0):1740774519868310.
- 464 31. Goldacre B, Morton CE, DeVito NJ. Why researchers should share their analytic code. *BMJ*.
465 2019;367:l6365.

466

467

468 **Additional Material**

469

470 Additional File 1.doc - Supplementary material – contains additional methods and results.

471 Additional File 2.doc - Protocol and data extraction form – contains protocol and data

472 extraction form used for this study.

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

Table 1 – Characteristics of eligible trials (N=101)

Characteristic	N (%)
Journal (n, %)	
Annals of Internal Medicine	3 (3%)
The BMJ	3 (3%)
JAMA	19 (19%)
Lancet	28 (28%)
NEJM	42 (42%)
PLOS Medicine	6 (6%)
Funding (n, %)	
Pharmaceutical	21 (21%)
Other for profit medical company	8 (8%)
Government	37 (37%)
Charity	5 (5%)
Multiple including pharmaceutical/other for profit medical	4 (4%)
Multiple excluding pharmaceutical/other for profit medical	22 (22%)
Other	4 (4%)
Type of intervention (n, %)	
Pharmacologic	52 (51%)
Surgical	13 (13%)
Psychosocial/behavioural/educational	9 (9%)
Other	24 (24%)
Multiple types	3 (3%)
Cluster trial (n, %)	14 (14%)
Factorial trial (n, %)	3 (3%)

Crossover trial (n, %)	3 (3%)
Non-inferiority trial (n, %)	20 (20%)
No. of treatment arms (n, %)	
Two	85 (84%)
Three or more	16 (16%)
Sample size	
Median, IQR	758 (306, 2129)
Min, Max	(36, 415357)

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

Table 2 – Description of unexplained discrepancies (N=89)

Unexplained changes	N (%)
Analysis population	
Changed set of patients included by specifying additional exclusions	12 (13%)
Analysis model	
Changed model	13 (15%)
Changed method of selecting analysis model	2 (2%)
Covariates	
Changed from unadjusted to adjusted analysis	7 (8%)
Changed from adjusted to unadjusted analysis	4 (4%)
Changed set of covariates included in analysis	5 (6%)
Missing data	
Changed from complete case to multiple imputation	1 (1%)
Changed imputation strategy	2 (2%)
Unexplained additions	
Analysis population	
Not mentioned in Original Analysis Plan	8 (9%)
Incomplete detail given in Original Analysis Plan	7 (8%)
Allowed analyst to subjectively choose analysis population based on trial dataset	1 (1%)
Analysis model	
Not mentioned in Original Analysis Plan	5 (6%)
Incomplete detail given in Original Analysis Plan	5 (6%)
Allowed analyst to subjectively choose analysis model based on trial dataset	7 (8%)

Covariates	
Not mentioned in Original Analysis Plan	2 (2%)
Incomplete detail given in Original Analysis Plan	2 (2%)
Allowed analyst to subjectively choose covariates based on trial dataset	3 (3%)
Missing data	
Not mentioned in Original Analysis Plan	9 (10%)
Incomplete detail given in Original Analysis Plan	3 (3%)
Allowed analyst to subjectively choose missing data approach based on trial dataset	1 (1%)

507

508

509