Journal of Clinical Epidemiology 108 (2019) 17−25

Journal of Clinical Epidemiology

REVIEW

Justification and reporting of subgroup analyses were lacking or inadequate in randomized controlled trials

Jingchun Fan[a,b], Fujian Song[b,*], Max O. Bachmann[b]

[a]*Epidemiology and Evidence Based Medicine, School of Public Health, Gansu University of Chinese Medicine, Lanzhou City, Gansu Province, P.R. China*
[b]*Public Health and Health Services Research, Norwich Medical School, University of East Anglia, Norwich, Norfolk, UK*

Accepted 11 December 2018; Published online 15 December 2018

Abstract

Objectives: The aim of the article was to assess the appropriateness and rationales of subgroup analyses planned in protocols of randomized controlled trials and reported in subsequent corresponding trial publications.

Study Design and Setting: We searched PubMed to identify trial protocols published in journals during 2006−2017. From a total of 3,774 initially identified records, we included a random sample of 479 protocols and identified 280 trial publications corresponding to the included protocols.

Results: Subgroup analyses were specified in 19% of the protocols and reported in 21% of the trial publications. Of the 94 protocols with planned subgroup analyses, 32% mentioned testing for interaction, and only three considered statistical power. Subgroup analyses were not prespecified in 56% of the 59 trial publications with subgroup analyses. Subgroup analyses were stated as prespecified in nine trial publications, without support evidence from the corresponding protocols. Subgroup analyses were often reported insufficiently for assessing the consistency of subgroup effects across studies. Justifications for subgroup analyses were provided in only four trial protocols and seven trial publications.

Conclusion: Inappropriate specification and reporting of subgroup analyses remain problematic in protocols and reports of randomized controlled trials. Justifications or rationales for subgroup analyses were only rarely provided in trial protocols and reports.

Keywords: Subgroup analysis; Subgroup effect; Prespecification; Clinical trial; Study protocol; Study reporting

1. Introduction

Randomized controlled trials (RCTs) usually report results as the overall average of relative effects of health care interventions, such as hazard ratios or mean differences. However, treatment effects are unlikely to be homogeneous across study participants, and averaging effects may mislead clinicians in the care of individual patients [1,2]. To maximize health gain from limited resources, it is a clinically meaningful question about whether the effect of a treatment differs by differing characteristics of patients [3].

Subgroup analysis in clinical trials aims to detect subgroup effects, defined as the difference in treatment effect between subgroups because of causal interactions [4], to inform clinical decisions about which patients should or should not receive a treatment [5,6]. However, there are well-known limitations with subgroup analyses, including inflated false positives due to multiple testing and high false negatives due to inadequate statistical power [7,8]. Although issues related to subgroup analysis have been debated for decades and numerous guidance on subgroup analyses has been advocated, controversy remains

Funding Statement: J.F., as a visiting scholar to the University of East Anglia, was supported by funding from China Scholarship Council (No. 201608625033), P. R. China. No other external funding received for this study.

Ethical Approval: This study is using data from published literature, and no ethical approval is required.

Data Sharing: No additional data available.

Conflict of interest statement: None declared.

Authors Contributions: F.S. developed the initial idea, had full access to all of the data, and take responsibility for the integrity and accuracy of the data. F.S., J.C.F., and M.O.B. finalized the study protocol. F.S. conducted literature search and randomly selected eligible trial protocols. J.C.F. collected full text publications of the included trial protocols and corresponding trial reports. J.C.F. extracted data from the trial protocols and trial reports. F.S. checked the data extracted, analyzed data, and drafted the article. J.C.F. and M.O.B. critically commented on the draft article. All authors approved the final article.

* Correspondence author. Norwich Medical School, Faculty of Medicine & Health Science, University of East Anglia, Norwich, NR4 7TJ UK. Tel.: 44 1603 591253.

E-mail address: fujian.song@uea.ac.uk (F. Song).

https://doi.org/10.1016/j.jclinepi.2018.12.009

What is new?

Key findings

- Inappropriate specification and reporting of subgroup analyses remain problematic in recently published protocols and reports of randomized controlled trials.
- Subgroup analyses were often reported insufficiently for assessing the consistency of results of subgroup analyses across studies, and rationales for subgroup analyses were very rarely provided in trial protocols and trial reports.

What this adds to what was known?

- The present study summarized empirical evidence about what rationales were used for subgroup analyses in a large number of randomized controlled trials. Furthermore, this is the first to examine whether results of subgroup analyses were sufficiently reported to facilitate the checking of consistency in subgroup effects reported.

What is the implication and what should change now?

- Further research are required to accumulate more empirical evidence about how and why subgroup analyses are planned or conducted in clinical trials.
- Methodological and reporting guidance needs to focus more on sufficient reporting of prespecified or post hoc subgroup analyses so as to facilitate checking of consistency in results of subgroup analyses across clinical trials.

regarding the conduct, reporting, and interpretation of subgroup analyses in clinical trials [9−12]. Recent studies found that there has been no improvement in the reporting of subgroup analyses [13,14], and that there were discrepancies between subgroup analyses planned in protocols and journal publications of clinical trials [15].

Many journals have been accepting and publishing clinical trial protocols so as to promote public access to them [16,17]. Using a sample of trial protocols published in journals, this study aimed to assess the appropriateness and rationales of subgroup analyses planned in trial protocols and those reported in trial publications.

2. Methods

We defined subgroup analyses as those based on participant characteristics at baseline. We excluded subgroup analyses defined by intervention characteristics (such as dosages, differences in health care providers), patient response to treatment, or other events after the start of a clinical trial.

2.1. Identification and inclusion of trial protocols

We included trial protocols published in journals during 2006−2017 and excluded those that met any of the following exclusion criteria: (1) preclinical trials that included healthy people, (2) nonrandomized, (3) published in languages other than English, (4) published before 2006, (5) nonhuman, and (6) the full text not openly available.

We searched PubMed on March 5, 2018, using key terms ‘‘study protocol,’’ random*, and ‘‘trial’’ in the title and located a total of 3,774 records. The located records were managed by Microsoft Excel (2016) and separated into five groups by year of publication: protocols published during 2006−2009, 2010−2011, 2012−2013, 2014−2015, and 2016−2017. Then we used Excel RAND command to generate a random number (from 0 to 1) for each of the identified records and ordered them by the assigned random numbers from the smallest to the largest. From each of the 5 year groups, we included the first 100 ranked protocols (starting from the smallest random number assigned). Because of the small number of protocols published before 2010, all 83 protocols published during 2006−2009 were included. Therefore, we initially included a total of 483 randomly selected trial protocols.

2.2. Data extraction and synthesis

We finalized a data extraction form (Supplementary Appendix 1) after pilot testing by two reviewers (J.C.F. and F.S.) using 10 of the sampled trial protocols. One reviewer (J.C.F.) then extracted information from the remaining sampled protocols, and a second reviewer (F.S.) thoroughly checked the extracted data.

We used the unique numbers of clinical trial registration to identify any subsequent publications reporting the results of trials that corresponded with the included trial protocols (up to July 31, 2018). We extracted data from the journal publications by one reviewer (J.C.F.) and checked by a second reviewer (F.S.) in line with items used in a previous study (Supplementary Appendix 1) [15]. The reporting of subgroup analyses in a trial publication was judged to be sufficient for cumulative subgroup analyses [18] if the trial report provided estimated subgroup effects with standard errors, confidence intervals, or interaction *P* values [19]. A subgroup effect was considered to be as claimed if the trial authors explicitly stated in the abstract or discussion/conclusion that the effect of an intervention was different between subgroups or that a clear benefit or harm was seen in one or more subgroups [20].

We calculated the proportion of trial protocols with planned subgroup analyses, the proportion of trial

publications reported subgroup analyses, explored the association of planned or reported subgroup analyses with study characteristics, and compared subgroup analyses planned in protocols and reported in trial publications. Data regarding sample size, journal impact factor, and number of subgroup analyses were categorized based on percentile distributions. We used the chi-square test for statistically analyzing categorical data and nonparametric trend test for ordinal data (such as ordered groups by year of publication and sample size).

3. Results

3.1. Characteristics of trial protocols and reports

Four of the 483 initially included protocols were eventually excluded for the following reasons: publication in non-English language, nonrandomized trial, animal study, and only the qualitative component of a planned trial. We identified 280 trial publications corresponding to the included protocols (see Supplementary Appendix 2 for references of included trial protocols and publications). The median number of years between the publication of the protocol and trial report was three (ranging from 0 to 10 years).

The main characteristics of the 479 included protocols and the 280 trial publications are shown in Table 1. More than half of the protocols (53.0%) were published in Trials, and 83.5% were published in journals with a journal impact factor ranging from 2.000 to 4.999 in 2017. Most protocols concerned trials in high-income countries (85.4%) and with nonindustry funding (85.4%). Of the 479 protocols, 86.2% were superiority design, 60.5% were multicenter studies, 19.0% were cluster trials, 23.8% had a sample size more than 500, and 50.1% used objectively measured primary outcomes (Table 1).

Subgroup analyses were planned in 94 (19.6%) of the 479 protocols (Table 1). The proportion of protocols with planned subgroup analyses decreased over time ($P = 0.029$), from 25.6% in protocols published during 2006−2009 to 13.3% in those published during 2016−2017 (Table 1). In addition, the specification of subgroup analyses in protocols was positively associated with larger sample size ($P < 0.001$), multicenter design ($P = 0.009$), and nonindustry funding ($P = 0.036$). There was no statistically significant association between specification of subgroup analyses and other trial characteristics in Table 1 or different clinical fields (Supplementary Appendix 3).

Fifty-nine (21.1%) of the 280 trial publications reported subgroup analyses (Table 1). The reporting of subgroup analyses was associated with larger sample size ($P < 0.001$), objectively measured primary outcomes ($P = 0.006$), and higher journal impact factor ($P = 0.002$). The proportion of trials reporting subgroup analyses was 12.9% in trials published in journals with an impact factor <5, compared with 36.2% in those with an impact factor ≥ 10 (Table 1). There was no statistically significant association between reporting of subgroup analyses in trial publication and other study characteristics.

3.2. Subgroup analyses planned or reported

The main characteristics of the 94 protocols with planned subgroup analyses and 59 trial publications reported subgroup analyses are shown in Table 2. All the planned subgroup analyses were exploratory, including nine (9.6%) that mentioned subgroup effects in trial objectives. Of the 94 protocols, only four provided justification for the planned subgroup analyses, three considered issues concerning statistical power; and none explicitly anticipated the direction of subgroup effects (Table 2). Test for interaction was mentioned in only 31.9% of the 94 trial protocols with planned subgroup analyses. Subgroup variables were explicitly listed in 78 (83.0%) of the 94 trial protocols, including disease characteristics (78.2%), age (30.8%), gender (29.5%), socioeconomic factors (26.9%), ethnicity (9.0%), and study center (9.0%). Cut-points were explicitly defined in only 27.3% of the 77 protocols with subgroups by continuous variables.

Of the 59 trial publications that reported subgroup analyses, more than half (55.9%) had not specified any subgroup analyses in the corresponding protocols (Table 2). Nine trial reports stated that the subgroup analyses were prespecified although there was no mention of any subgroup analyses in the corresponding protocols. Seven trial reports tried to justify at least one of the subgroup analyses, only one anticipated the direction of possible subgroup effects, and none considered issues regarding statistical power of subgroup analysis. Statistical test for interaction or subgroup effects was not conducted in 23 (39.0%), including 8 of the 20 trial publications with subgroup effect claims. Results of subgroup analyses were insufficiently reported in 21 (35.6%) of the 59 trial publications.

When trial protocols and publications were categorized into two groups by year of publication (2006−2013 and 2014−2018), there were no improvements over time in the use of a statistical test for interaction in prespecification of subgroup analyses in protocols or in reporting of sufficient data on results of subgroup analyses (Table 3).

3.3. Justification for subgroup analyses

Only 4 of the 94 trial protocols with planned subgroup analyses provided some justifications or rationales for subgroup analyses (Supplementary Appendix 4). In two protocols [21,22], subgroup analyses were based on findings from previous clinical studies. The planned subgroup analyses were simply based on the suspected differences in treatment effect by subgroups, without giving more details, in the remaining two trial protocols [23,24]. Results of these four RCTs have been published, and all reported no subgroup effects.

Table 1. The main characteristics of included protocols of randomized controlled trials

	RCT protocols		Trial publications	
Characteristics	***N***	**Planned SGAs (%)**	***N***	**Reported SGAs (%)**
Total	479	94 (19.6)	280	59 (21.1)
Year of publication				
2006–2009	82	21 (25.6)	11	5 (45.5)
2010–2011	100	23 (23.0)	29	5 (17.2)
2012–2013	99	18 (18.2)	49	8 (16.3)
2014–2015	100	19 (19.0)	80	21 (26.3)
2016–2017/18	98	13 (13.3)	111	20 (18.0)
Journal impact factor				
<2.000	32	4 (12.5)	27	2 (7.4)
2.000–4.999	400	83 (20.8)	112	16 (14.3)
5.000–9.999	27	4 (14.8)	82	19 (23.2)
≥10.000	0	–	58	21 (36.2)
Geographical region				
Europe	252	53 (21.0)	150	38 (25.3)
Asia	61	13 (21.3)	25	5 (20.0)
Oceania	61	10 (16.4)	46	7 (15.2)
North America	59	10 (17.0)	36	4 (11.1)
South America	18	2 (11.1)	7	1 (14.3)
Africa	13	1 (7.7)	0	0 (0.0)
Multiple	15	5 (33.3)	8	4 (50.0)
Country income category				
Developed	409	83 (20.3)	250	54 (21.6)
Less developed	63	8 (12.7)	26	3 (11.5)
Mixed	7	3 (42.9)	4	2 (50.0)
Funding source				
Nonindustry	409	87 (21.3)	245	51 (20.8)
Industry	28	5 (17.9)	14	1 (7.1)
Mixed/other	42	2 (4.8)	13	7 (53.8)
Trial design type				
Superiority	384	81 (21.1)	239	50 (21.1)
Exploratory	87	12 (13.8)	38	9 (23.7)
Noninferiority	8	1 (12.5)	3	0 (0.0)
No. of centers				
Multicenter	290	68 (23.5)	177	43 (24.3)
Single center	189	26 (13.8)	103	16 (15.5)
Cluster trials				
Yes	91	21 (23.1)	61	12 (19.7)
No	388	73 (18.8)	219	47 (21.5)
Sample size category				
≤100	121	10 (8.3)	72	5 (6.9)
101–250	126	23 (18.3)	74	13 (17.6)
251–500	113	26 (23.0)	63	22 (34.9)
>500	114	35 (30.7)	71	19 (26.8)
Type of primary outcome				
Objective	240	51 (21.3)	134	39 (29.1)
Subjective	234	41 (17.5)	145	20 (13.8)
Mixed/unclear	5	2 (40.0)	1	0 (0.0)

Abbreviations: RCT, randomized controlled trials; SGA, subgroup analysis.

Table 2. Characteristics of subgroup analyses planned in trial protocols or reported in trial publications

Variables	Planned SGAs in trial protocols	Reported SGAs in trial publications
Total	94 (100%)	59 (100%)
Prespecification		
Prespecified	94 (100%)	26 (44.1%)
Post hoc	0	33 (55.9%)
Justification for SGAs		
Provided	4 (4.3%)	7 (11.9%)
Not provided	90 (95.7%)	52 (88.1%)
Described SGA variables		
Yes	78 (83.0%)	–
No	16 (17.0%)	–
Anticipated SGE direction		
Yes	0 (0%)	1 (1.7%)
No	94 (100%)	58 (98.3%)
Considered power for SGAs		
Yes	3 (3.2%)	0 (0%)
No	91 (96.8%)	59 (100%)
Test for interaction		
Yes	30 (31.9%)	36 (61.0%)
No	64 (68.1%)	23 (39.0%)
No. of SGAs		
1–3	45 (47.9%)	22 (40.0%)
4+ or unclear	49 (52.1%)	33 (60.0%)
Cut-points defined for continuous SGV		
Yes	21 (27.3%)	
No	56 (72.7%)	–
NA	17 (NA)	–
Sufficient SGA results		
Yes	–	38 (64.4%)
No	–	21 (35.6%)
Subgroup effects claims		
Yes	–	20 (33.9%)
No	–	39 (66.1%)

Abbreviations: SGA, subgroup analysis; SGE, subgroup effects; SGV, subgroup variables.

Justifications or rationales for subgroup analyses were provided in 7 of the 59 trial publications with subgroup analyses (Supplementary Appendix 4) although no justifications were mentioned in any of the corresponding protocols. All justified subgroup analyses according to findings from previous studies. In one case, post hoc subgroup analyses were conducted as a response to critics of previous clinical trials with negative results [25]. There were no statistically significant subgroup effects reported in the seven trial publications.

Of the 221 trial publications that did not report subgroup analyses, nine mentioned reasons for not conducting or reporting any subgroup analyses. Inadequate statistical power and inadequate recruitment of participants belonging to certain subgroups were the reasons for lack of subgroup analyses in trial reports, except for one which stated that subgroup analyses would be reported in separate publications.

4. Discussion

We found that subgroup analyses were planned in 19.6% of the 479 RCT protocols published in journals and reported in 21.1% of the 280 corresponding trial publications. The proportion of trial protocols with planned subgroup analyses was associated with sample size, multicenter design, and nonindustry funding. There was a reduced proportion of protocols with planned subgroup analyses over time. Justifications for subgroup analyses were rarely provided (4.3%), mainly based on findings from previous studies or suspected possibility of subgroup effects.

The reporting of subgroup analyses in trial publications was positively associated with sample size, objectively measured outcomes, and journal impact factors. More than half of the subgroup analyses reported in trial publications were post hoc, without specification in the corresponding

Table 3. Interaction test for subgroup effects and adequate result reporting by grouped publication year

	Trial protocols		Trial reports	
Characteristics	***N***	**Yes (%)**	***N***	**Yes (%)**
Statistical test for subgroup effect or interaction				
2006–2013	62	21 (33.9)	18	13 (72.2)
2014–2018	32	9 (28.1)	41	23 (56.1)
Sufficient data on subgroup analyses reported[a]				
2006–2013	–	–	18	12 (66.7)
2014–2018	–	–	41	26 (63.4)
Specification of subgroup analyses in protocols				
2006–2013	–	–	18	9 (50.0)
2014–2018	–	–	41	17 (41.5)

[a] The reporting of subgroup analyses in a trial publication was judged to be sufficient for cumulative subgroup analyses if the trial report provided estimated subgroup effects with standard errors, confidence intervals, or interaction *P* values [19].

protocols. Furthermore, subgroup analyses were described as prespecified in some trial publications although none were mentioned in the corresponding protocols. Compared with trial protocols, relatively more trial reports tried to retrospectively justify subgroup analyses. Seven of the 59 trial reports with subgroup analyses provided some

Table 4. Comparison with previous studies of subgroup analyses in randomized trials

Characteristics	**The present study**	**Kasenda et al. 2014 [15]**	**Chan et al. 2008 [26]**	**Al-Marzouki et al. 2008 [27]**
Source of trial protocols	479 trial protocols published in journals during 2006–2017.	894 protocols approved by six research ethics committees in Switzerland, Germany, and Canada during 2000–2003.	70 protocols approved by the scientific ethics committees for Copenhagen and Frederiksberg, Denmark in 1994–1995.	37 trial protocols with summaries published on the Lancet's Web site as of June 2007.
Trial protocols				
Protocols with planned SGAs	19.6% (94/479)	28.2% (252/894)	18.6% (13/70)	48.6% (18/37)
Justification/rationale for planned SGAs	4.3% (4/94)	–	–	2.7% (1/37)
Anticipated direction of SGE for planned SGAs	0.0% (0/94)	4.0% (10/252)	–	–
Appropriate test for interaction or subgroup effects	31.9% (30/94)	34.5% (87/252)	–	–
Full trial publications				
Reporting of SGAs	21.1% (59/280)	47.8% (246/515)	28.6% (20/70)	75.7% (28/37)
Post hoc SGAs without specification in protocol	55.8% (33/59)	19.5% (48/246)	60.0% (12/60)	39.3% (11/28)
Prespecification stated in trial reports, but not in protocols	25.7% (9/35)	34.6% (28/81)	57.1% (4/7)	–
Provided rationales for SGAs	11.9% (7/59)	–	–	–
Anticipated the direction of SGE	1.7% (1/57)	4.5% (11/246)	–	–
Considered statistical power for SGAs	0.0% (0/57)	2.4% (6/246)	–	–
Appropriate test for interaction	61.0% (36/59)	39.0% (96/246)	–	–
Sufficient reporting of SGAs	64.4% (38/59)	–	–	–
Claimed any SGEs	33.9% (20/59)	35.0% (86/246)	–	–

Abbreviations: SGAs, subgroup analyses; SGEs, subgroup effects.

rationale for subgroup analyses although none were available in the corresponding trial protocols.

Testing for subgroup effects or interaction was planned in 31.9% of the 94 protocols and conducted in 61.0% of the 59 trial publications. There was no improvement over time in appropriate testing for interaction in either trial protocols or reports or in reporting of results of subgroup analyses in trial publications.

4.1. Comparison with previous studies

Several previous studies have evaluated subgroup analyses in study protocols and their corresponding study reports [15,26−28]. Results of the present and previous relevant studies of RCT protocols [15,26,27] are shown in Table 4. The proportion of trial protocols with planned subgroup analyses in the present study was 19.6%, which was similar to one previous study [26], but lower than was reported in other studies (from 28.2% [15] to 48.6% [27]). Differing sources of trial protocols and year periods may explain some differences in planned subgroup analyses across studies. Similar to a previous study [15], we found that specification of subgroup analyses in trial protocols was positively associated with sample size and multicenter design. However, contrary to a previous study [15], we observed a significant association between subgroup analysis specification and nonindustry funding. As in previous studies, we found that trial protocols rarely anticipated the direction of subgroup effects, and there were discrepancies between subgroup analyses planned in trial protocols and presented in trial reports. Available evidence indicated that inadequate prespecification of subgroup analyses in trial protocols and reporting of subgroup analyses in trial publications have not improved over decades. A previous study (not included in Table 4 because of a mix of different study designs) found that justifications for subgroup analyses were provided in 16.3% of 76 grant applications [28]. Justifications for subgroup analyses were rarely provided in RCT protocols; which was 4.3% in the present study and only 2.7% in a small study of 37 trial protocols [27].

4.2. Strengths and limitations

This is the first study using a random sample of trial protocols published in journals and corresponding trial reports to examine planned and reported subgroup analyses in RCTs. Findings from the present study update and further strengthen existing empirical evidence regarding the appropriateness of subgroup analyses in clinical research. Particularly, the present study summarized empirical evidence about how specified subgroup analyses were justified or what rationales were used for subgroup analyses in a large number of RCTs. Furthermore, the present study is the first to examine whether results of subgroup analyses were sufficiently reported to facilitate the checking of consistency in subgroup effects reported and the conduct of cumulative subgroup analyses [18]. Similar to cumulative meta-analysis [29], cumulative subgroup analysis refers to a series of repeated meta-analyses of subgroup effects after adding data from each new trial chronologically.

The present study included protocols of trials conducted in more regions or countries compared with previous studies that considered trials restricted in high-income countries in Europe or North America. However, protocols published in journals may be systematically different from those approved by research ethics committees. For example, the proportion of trials with industry funding was 17.9% in the present study compared with 60.2% in a previous study of trial protocols approved by research ethics committees in Europe or Canada [15]. Because of the restriction of available resources and time, we included only a random sample of 479 from a total of 3,774 published protocols identified from PubMed. Furthermore, we did not check any amendments of trial protocols or statistical analysis plans of trials for possible changes in specification of subgroup analyses. However, we believe that the main findings from the present study were unlikely to be materially different by including more published protocols or by obtaining additional information from protocol amendments and statistical analysis plans. These assumptions might be worthy of further evaluation of the literature in the future.

4.3. Implications

4.3.1. For journal editors and peer reviewers

Specification of subgroup analyses in trial protocols published in journals could be much improved if journal editors and peer reviewers identified inappropriate subgroup analyses before protocols are accepted for publication. For example, investigators could be encouraged to provide justifications or rationales for planned subgroup analyses, to anticipate the direction of subgroup effects, and to use appropriate statistical test for interaction. Journal editors and peer reviewers may need to improve their own understanding of what are important issues for prespecification of subgroup analyses in trial protocols. Attention from journal editors and peer reviewers is also needed to the appropriate reporting of subgroup analyses in trial publications. For example, the discrepancies in subgroup analyses between protocols and trial reports could be avoided or explained by checking the corresponding trial protocols before accepting trial reports for publication. The journal editors and peer reviewers should also ensure sufficient reporting of results of subgroup analyses for cumulative meta-analyses of subgroup effects.

4.3.2. For researchers

Medical and health researchers need to understand how to plan and conduct appropriate subgroup analyses in clinical studies. During the preparation of trial protocols,

issues concerning subgroup analyses should be explicitly considered. Justifications and rationales for planned subgroup analyses should be provided although subgroup analyses should also be allowed purely for exploratory purpose. In trial publications, both prespecified and post hoc subgroup analyses should be reported with sufficient data for cumulative subgroup analyses [19]. The present study found that some trial publications did not report subgroup analyses because of the small sample size. For a single study, the sample size may be too small to reveal significant subgroup effects. However, if multiple small studies on the same topic conduct the same subgroup analyses, cumulative subgroup analyses may show how consistent are results of subgroup analyses across individual trials and provide a more valid overall estimate of subgroup effects [18].

4.3.3. Methodological and reporting guidance

There is much existing guidance aiming to reducing inappropriate use of subgroup analyses in clinical research [5,6,8,30−32]. The key points for assessing the credibility of subgroup analyses in clinical trials include prespecification of subgroup analyses in protocols, a small number of subgroup analyses conducted, appropriate statistical test for subgroup effects, and consistency in results of subgroup analyses across trials. As attention has been focused mainly on the problem of inflated false positives, efforts to avoid inappropriate subgroup analyses may have resulted in a recent reduction in the proportion of trial protocols with planned subgroup analyses as shown in the present study. If fewer or no subgroup analyses are conducted in clinical trials, it will be difficult or impossible to examine the consistency in results of a subgroup analysis across different trials, and valuable data from clinical research will be wasted [18]. Therefore, methodological guidance on subgroup analyses in clinical research should emphasize the importance of sufficient reporting of results of all prespecified or post hoc subgroup analyses conducted, rather than insist on only a small number of planned subgroup analyses. With sufficient reporting of subgroup analyses from multiple trials, some problems related to inappropriate subgroup analyses, such as lack of test for interaction and false positive or false negative subgroup effects in isolated individual trials, could be properly corrected with evidence accumulation.

4.3.4. Further research required

The present study found very limited empirical evidence regarding justifications or rationales for subgroup analyses planned or conducted in clinical trials. Although all existing guidance on clinical research have emphasized the importance of a priori specification of subgroup hypotheses in clinical trials, there are no guidance specifically about how to specify a priori subgroup hypotheses at the design stage of a clinical trial. Further research are required to accumulate more empirical evidence about how and why subgroup analyses are planned or conducted in clinical trials and what measures could be taken to improve the efficient use of data from clinical trials to appropriately identify clinically meaningful subgroup effects.

5. Conclusions

Inappropriate specification and reporting of subgroup analyses remain problematic in published protocols and reports of RCTs although the trial protocols published more recently are less likely to plan subgroup analyses than those published earlier. Justifications or rationales for subgroup analyses were only rarely provided in trial protocols and trial reports. Methodological and reporting guidance needs to focus more on sufficient reporting of prespecified or post hoc subgroup analyses so as to facilitate checking of consistency in results of subgroup analyses across clinical trials.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2018.12.009.

References

[1] Rothwell PM. Can overall results of clinical trials be applied to all patients? Lancet 1995;345:1616−9.
[2] Kravitz RL, Duan N, Braslow J. Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. Milbank Q 2004;82(4):661−87.
[3] NICE. Guide to the methods of technology appraisal. London: National Institute for Health and Clinical Excellence; 2013.
[4] VanderWeele TJ, Knol MJ. Interpretation of subgroup analyses in randomized trials: heterogeneity versus secondary interventions. Ann Intern Med 2011;154:680−3.
[5] Sun X, Ioannidis JP, Agoritsas T, Alba AC, Guyatt G. How to use a subgroup analysis: users' guide to the medical literature. JAMA 2014;311:405−11.
[6] Rothwell PM. Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. Lancet 2005;365:176−86.
[7] Pocock SJ, Hughes MD, Lee RJ. Statistical problems in the reporting of clinical trials. A survey of three medical journals. N Engl J Med 1987;317:426−32.
[8] Oxman AD, Guyatt GH. A consumer's guide to subgroup analyses. Ann Intern Med 1992;116:78−84.
[9] Altman DG. Within trial variation–a false trail? J Clin Epidemiol 1998;51:301−3.
[10] Feinstein AR. The problem of cogent subgroups: a clinicostatistical tragedy. J Clin Epidemiol 1998;51:297−9.
[11] Peto R. Current misconception 3: that subgroup-specific trial mortality results often provide a good basis for individualising patient care. Br J Cancer 2011;104:1057−8.
[12] Oxman AD. Subgroup analyses: the devil is in the interpretation. BMJ 2012;344:e2022.
[13] Gabler NB, Duan N, Raneses E, Suttner L, Ciarametaro M, Cooney E, et al. No improvement in the reporting of clinical trial subgroup effects in high-impact general medical journals. Trials 2016;17(1):320.

[14] Wallach JD, Sullivan PG, Trepanowski JF, Sainani KL, Steyerberg EW, Ioannidis JP. Evaluation of evidence of statistical support and corroboration of subgroup claims in randomized clinical trials. JAMA Intern Med 2017;177(4):554−60.
[15] Kasenda B, Schandelmaier S, Sun X, von Elm E, You J, Blumle A, et al. Subgroup analyses in randomised controlled trials: cohort study on trial protocols and journal publications. BMJ 2014;349:g4539.
[16] Li T, Boutron I, Al-Shahi Salman R, Cobo E, Flemyng E, Grimshaw JM, et al. Review and publication of protocol submissions to trials - what have we learned in 10 years? Trials 2016;18(1):34.
[17] Chan AW, Hrobjartsson A. Promoting public access to clinical trial protocols: challenges and recommendations. Trials 2018;19(1):116.
[18] Song F, Bachmann MO. Cumulative subgroup analysis to reduce waste in clinical research for individualised medicine. BMC Med 2016;14(1):197.
[19] Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA 2004;291:2457−65.
[20] Sun X, Briel M, Busse JW, You JJ, Akl EA, Mejza F, et al. Credibility of claims of subgroup effects in randomised controlled trials: systematic review. BMJ 2012;344:e1553.
[21] Casaer MP, Hermans G, Wilmer A, Van den Berghe G. Impact of early parenteral nutrition completing enteral nutrition in adult critically ill patients (EPaNIC trial): a study protocol and statistical analysis plan for a randomized controlled trial. Trials 2011;12:21.
[22] de Gans K, de Haan RJ, Majoie CB, Koopman MM, Brand A, Dijkgraaf MG, et al. PATCH: platelet transfusion in cerebral haemorrhage: study protocol for a multicentre, randomised, controlled trial. BMC Neurol 2010;10:19.
[23] Dinneen SF, O' Hara MC, Byrne M, Newell J, Daly L, O' Shea D, et al. The Irish DAFNE study protocol: a cluster randomised trial of group versus individual follow-up after structured education for type 1 diabetes. Trials 2009;10:88.
[24] Mujagic E, Zwimpfer T, Marti WR, Zwahlen M, Hoffmann H, Kindler C, et al. Evaluating the optimal timing of surgical antimicrobial prophylaxis: study protocol for a randomized controlled trial. Trials 2014;15:188.
[25] Muller RG, Haase N, Wetterslev J, Perner A. Effects of hydroxyethyl starch in subgroups of patients with severe sepsis: exploratory post-hoc analyses of a randomised trial. Intensive Care Med 2013; 39(11):1963−71.
[26] 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.
[27] Al-Marzouki S, Roberts I, Evans S, Marshall T. Selective reporting in clinical trials: analysis of trial protocols accepted by the Lancet. Lancet 2008;372:201.
[28] Boonacker CW, Hoes AW, van Liere-Visser K, Schilder AG, Rovers MM. A comparison of subgroup analyses in grant applications and publications. Am J Epidemiol 2011;174:219−25.
[29] Lau J, Schmid CH, Chalmers TC. Cumulative meta-analysis of clinical trials builds evidence for exemplary medical care. J Clin Epidemiol 1995;48:45−57. discussion 59−60.
[30] Burke JF, Sussman JB, Kent DM, Hayward RA. Three simple rules to ensure reasonably credible subgroup analyses. BMJ 2015;351:h5651.
[31] Wang R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. Statistics in medicine–reporting of subgroup analyses in clinical trials. N Engl J Med 2007;357:2189−94.
[32] Yusuf S, Wittes J, Probstfield J, Tyroler HA. Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. JAMA 1991;266:93−8.