Comparisons between protocols and publications of case-control studies: analysis of potential causes of non-reproducibility and recommendations for enhancing the quality of personalization in healthcare

Haifeng Hou^{1,2} · Guoyong Ding¹ · Xuan Zhao¹ · Zixiu Meng¹ · Jiangmin Xu³ · Zheng Guo² · Yulu Zheng² · Dong Li¹ · Wei Wang^{1,2}

Received: 18 December 2018 / Accepted: 8 March 2019 / Published online: 19 March 2019
© European Association for Predictive, Preventive and Personalised Medicine (EPMA) 2019

Abstract

Background Selective reporting of results in published case-control studies has been widely suspected, but little comprehensive information on selective reporting is available with regard to case-control studies. We aimed to evaluate the concordance of findings between publications and the protocols of case-control studies and to assess the level of selective reporting of results in case-control studies.

Methods The databases of Embase, Medline, Scopus, and Web of Science were searched to identify case-control study protocols published between January 1, 1990 and December 31, 2017. The numbers and characteristics of predefined exposures (or factors) were extracted from the protocols. The reported and unreported factors were both collected from the published studies and protocols. The frequency of selective reporting of results were estimated by identifying the discrepancies of factors between the protocols and the published studies. Study sample size and the extent of selective reporting of factors were measured by a Spearman correlation analysis.

Results Fourteen protocols with 24 published studies and 159 factors were identified, of which eight protocols (57.1%) had discrepancies between the publications and protocols. The prevalence of incomplete reporting in published case-control studies was 42.9% (6/14), with participant characteristics, anthropometric and laboratory measurement variables more likely to be unreported. A total of 16,835 cases and 56,049 controls were recruited in the 14 protocols of case-control studies (sample size ranges from 428 to 52,596 per study). Sample size had no statistical significance with selective reporting of results ($P > 0.05$).

Haifeng Hou, Guoyong Ding and Xuan Zhao contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13167-019-00165-2>) contains supplementary material, which is available to authorized users.

✉ Wei Wang
wei.wang@ecu.edu.au

Haifeng Hou
hfhhou@163.com

Guoyong Ding
gyding@tsmc.edu.cn

Xuan Zhao
xzhao1995@163.com

Zixiu Meng
zixiumeng@163.com

Jiangmin Xu
xuji@wssu.edu

Zheng Guo
zgao0@our.ecu.edu.au

Yulu Zheng
yuluz@our.ecu.edu.au

Dong Li
tsmdongli@163.com

¹ School of Public Health, Taishan Medical University, Taian, China

² School of Medical and Health Sciences, Edith Cowan University, Perth, WA 6027, Australia

³ School of Health Sciences, Winston-Salem State University, Winston-Salem, USA

Conclusion The study protocols should be publicly available prior to the completion of case-control studies so that the potential bias can be assessed by the readers. Our findings highlight the need for investigators, peer reviewers, and readers to exercise increased awareness and scrutiny due to the undesirable practice of selective reporting of results in medical sciences causing the loss of potentially important information, thus impacting quality of personalized attitude in healthcare in the context of the predictive, preventive, and personalized medicine.

Keywords Case-control study · Protocol · Selective reporting of results · Publication bias · Predictive preventive personalized medicine · Quality of healthcare

Abbreviations

CCS	Case-control study
CONSORT	The Consolidated Standards of Reporting Trials
FDA	Food and Drug Administration
ICMJE	The International Committee of Medical Journal Editors
NIH	National Institutes of Health
NOS	Newcastle-Ottawa Scale
PPPM	Predictive, preventive, and personalized medicine

Introduction

The selective reporting of statistically significant results and favorable outcomes have raised widespread concerns about the validity of clinical studies [1–3]. Selective reporting of results distorts the reproducibility of research, produces publication bias, and results in the dissemination of inaccurate results, particularly if these selections are made subjectively to display apparently positive results [4, 5]. The prevalence of selective result reporting was observed to be up to 62% in previously reported clinical trials [1]. As a main source of publication bias, selective result reporting can lead to falsely high effect estimates of factors when a meta-analysis is conducted referring to these biased studies.

Many efforts were made to establish some standards to reduce selective reporting of results. Prompted by the Food and Drug Administration (FDA) and the National Institutes of Health (NIH), it was suggested that trial registration should be implemented prior to the onset of a clinical trial [6, 7]. The Consolidated Standards of Reporting Trials (CONSORT) statement, developed by the International Committee of Medical Journal Editors (ICMJE) in 1996, stated the requirement that trial registration be a precondition for the publication of each clinical trial [4, 8]. The disclosure of protocol prior to commencing a study was required to reduce subjective reporting of statistically positive findings or favorite results and increase awareness of potential publication bias, by facilitating readers to compare published articles to the details interpreted in the published protocols [9].

Although trial registration effectively reduces selective outcome reporting [10, 11], this registration strategy has not been established in either observational studies or non-interventional studies of human beings. Recently, several protocols were published before the enrollment of participants in observational studies (i.e., case-control study [CCS] and cohort study) [12–14].

CCS is one of the most important methodologies in health and medical sciences, and has been widely used in exploration of etiology, comparison of diagnostic approaches, evaluation of clinical treatments, and other healthcare investigations [15–17]. To resolve the co-morbidities and complex clinical situations worldwide, new scientific findings of CCSs are required for the development of predictive, preventive, and personalized medicine (PPPM)-related approaches [18, 19]. However, selective reporting of results might exist in CCSs, which limits the establishment of innovative strategies based on personalized prediction, prevention, and treatment. To date, there has been no direct evidence of such selective reporting within CCSs, nor have articles on such topics been reported, even though they may be subject to publication bias. We conducted this study to estimate the prevalence of selective result reporting in published CCSs and to identify discrepancies between the factors predefined in the protocols and those reported in the published CCSs.

Methods

Search strategy

Embase, Medline, Scopus, and Web of Science databases were used to search for protocols and also protocol amendments for CCSs published from January 1, 1990 to December 31, 2017. The search strategy was “protocol [Title/Abstract]” AND “case control study [Title/Abstract].” For the included protocols, the relevant CCSs were searched according to the following information reported in the protocols: “authors,” “key words,” and “affiliates.”

Selection criteria and data extraction

The literature was included in this study if it meets the following criteria: (1) studies conducted on human populations; (2)

studies conducted to investigate a specific disease, injury, or psychological disorder; (3) CCSs designed dependently; and (4) studies published in English. Protocols with at least one publication were included in this study. For each protocol, all published studies reporting full or part results were included.

For each CCS, two authors independently reviewed the study protocol and all published studies, and extracted the following data: (1) characteristics of studies, (2) number and nature of reported factors (e.g., statistical significance, completeness of reporting, and specification of stratification analysis), and (3) number and specification of unreported factors. Any inconsistencies between the two authors were resolved by discussion with a third author. Information from amended protocols took precedence over information from earlier versions [1]. An exposure was a risk factor that was used to assess the association with specific disease or event between case and control groups. Unreported factors were those factors that were designated in the protocol but not reported in the published studies. All factors in a specific protocol were identified whether reported or not. Because statistical significance is the major agent of impact on selective result publication, we reviewed all the included studies to retrieve any reported exposure that did not reach statistical significance but was prespecified in the protocols. The nature and amount of the abovementioned factors in each protocol and relevant published studies were summarized and then entered into a spreadsheet.

Assessment of methodological quality

The methodological quality of each study was assessed in accordance with the quality assessment scale by reference to Newcastle-Ottawa Scale (NOS) guideline (Table S1) [20]. No study was excluded based on the assessment of methodological quality since each of the studies was of sufficient quality.

Statistical analyses

Our goal was to evaluate the consistency of reported factors between the protocol and corresponding published studies by determining the number and proportion of discrepant factors in these studies. Frequencies or percentages were used to describe qualitative data. If quantitative data were not normal distributions, they were expressed as medians and quartiles (P_{25} – P_{75}) or range (P_0 – P_{100}). Comparisons between groups were performed by Mann-Whitney tests. The correlation between study sample size and the extent of selective reporting of factors was measured by a Spearman correlation analysis. Statistical analyses were performed using SPSS version 24.0 (IBM, New York, USA). A *P* value less than 0.05 was considered statistically significant.

Results

Search results

The protocol search retrieved 2981 records (556 from Medline, 544 from Web of Science, 982 from Embase, and 899 from Scopus). After the removal of 1443 duplicates, 1538 records were screened with 1507 records excluded due to the following reasons: cohort studies and protocols ($n = 95$), no human studies ($n = 185$), clinical trials and protocols ($n = 198$), CCS articles without corresponding protocols ($n = 826$), reviews and other reasons ($n = 203$). Thirty-one protocols of CCSs were thus identified. The studies conducted by following these selected CCS protocols were further searched and 17 protocols that had not published any corresponding studies were excluded. Finally, 14 protocols and the related 24 published CCSs were included in this current study [21–34]. The flow chart of literature search is shown in Fig. 1.

Characteristics of included studies

Characteristics of these protocols of CCSs are shown in Table 1. A total of 16,835 cases and 56,049 controls were recruited in the 14 CCSs. The median sample size was 1227.0 (range 428–52,596). Cardiovascular (3/14; 21.4%), injury (2/14; 14.3%), and respiratory (2/14; 14.3%) were the leading medical domains having published research protocols. The majority of the protocols were published in Europe (9/14; 64.3%). Seven protocols (50.0%) were published from 2010 to 2017, four protocols (28.6%) between 1990 and 2000, and three (21.4%) between 2001 and 2010. Four protocols (28.6%) were identified with more than two CCSs published subsequently, of which one (7.1%) was identified with eight publications of the relevant CCSs. With regard to the number of predefined factors in protocols, 3 to 29 factors were identified (median value = 12.0).

Analysis between protocols and publications

Further analyses examined whether the factors predefined in a CCS protocol were reported in the subsequent published studies. As shown in Tables S2–S3 and Fig. 2, there were six CCSs (42.9%), in which the factors reported in publications were similar to the defined factors in the protocols. Two studies (14.3%) reported more factors whereas six studies (42.9%) reported fewer factors than the protocols. As shown in Table 2 and Fig. 2, the median of the numbers of factors in the 14 protocols was 12.0 (quartile 7.5–17.5, range 3–29), which was not significantly different from that of factors reported in the published study results (median 10.0, quartile 3.0–17.5, range 1–29). Moreover, the number of unreported factors in published studies was with a median of 0.0 (quartile 0.0–3.3, range 0–12), whereas the reported factors with

Fig. 1 Flow chart of literature search

statistical non-significance in the published studies was 3.5 (quartile 1.0–7.0, range 0–12).

In this study, the factors in the CCSs were classified into six domains: (1) participant characteristics, (2) anthropometric and laboratory measurements, (3) history and family history of disease, (4) environmental factors, (5) lifestyles and work-related agents, and (6) genetic and other factors. The details of

reported or unreported variables in these classifications are shown in Tables S1 and S2. Variables concerning participant characteristics, anthropometric variables, and laboratory measurement were more likely to be unreported when they were found statistically non-significant.

Furthermore, we tested the hypothesis that larger sample sizes effectively reduced the selective reporting of results. We examined the correlation between sample size and the number of protocol-defined factors published in subsequent CCSs, as well as the correlation of sample size with the ratio of number of published variables to number of protocol-defined variables. As shown in Figs. 3 and 4, no statistically significant correlations were detected, indicating that the sample size was not an underlying indication for predicting the selective publication of results.

Discussion

To evaluate the prevalence of selective result reporting for CCSs, we searched all published protocols in Embase, Medline, Scopus, and Web of Science databases from January 1, 1990 to December 31, 2017. Our study has some interesting findings. Only six CCSs (42.9%) reported similar factors in published studies with those predefined in the protocols, while two studies (14.3%) reported more factors and six studies (42.9%) reported fewer factors than the correspondent protocols. Statistical non-significance was the major reason leading to incomplete result reporting. Sample size was not correlated with selective reporting. Although the current study was based on only 14 protocols with 24 published journal articles, we suspect that selective reporting of results of CCSs is a common practice, and much more serious than the

Table 1 Characteristics of the included studies

Characteristics	No. of protocols	Frequencies (%)
Medical domains		
Cancer	1	7.1
Cardiovascular	3	21.4
Endocrine and metabolic	1	7.1
Infectious disease	1	7.1
Injury	2	14.3
Nervous	1	7.1
Respiratory	2	14.3
Others	3	21.4
Areas		
America	3	21.4
Asia	2	14.3
Europe	9	64.3
Year of publication		
1990–2000	4	28.6
2001–2010	3	21.4
2010–2017	7	50.0
No. of case-control studies published		
1	10	71.4
2	3	21.4
8	1	7.1
Type of participants		
Hospital-based	13	92.9
Community-based	1	7.1
Sample size of participants		
Median	1226.5	–
Range	428.0–52,596.0	–
No. of predefined factors in protocols		
Median	12	–
Range	3–29	–

Fig. 2 Classifications of factors in protocols and their published case-control studies

estimates of this study indicated, since there are numerous CCSs without an available protocol.

The incomplete reporting of results precludes both precise interpretation of corresponding results and the entire inclusion of these findings in further systematic review [1]. Registries of clinical trials have been widely utilized to prevent selective outcome reporting [4]. The disclosure of protocols to the public obviously improves the reliability and consistency of the full reporting of results [35, 36]. At present, registration of an available protocol is not a requirement of journal publication policy for CCSs, which not only limits access to full study protocols and expected results, but also restricts research credibility and transparency.

Facing co-morbidities and complex clinical challenges, physicians are requested to develop precise prediction, differential diagnosis, and personalized treatment, and to improve quality of healthcare [37]. However, selective reporting of results lessens the accuracy of evidences from clinical researches. One study showed that around 40% of trials funded by the Canadian Institutes of Health Research were reported with discrepant outcomes with protocols [38]. Incomplete outcome reporting was also found in 62% trials approved by an ethics committee in Denmark [1]. After the policy of registration was embraced by

ICMJE, a lower level of discrepancies (31%) were revealed between the outcomes registered in protocols and those in published clinical trials [39]. Our findings showed that discrepancies of factors between the published CCSs with accessible protocols (42.9%) were not significantly higher than those in the above registered clinical trials, indicating that protocol disclosure may also improve full reporting of results within CCSs. Even though incomplete reporting was found in nearly 50% of the CCSs, the statistical analysis showed no significant discrepancies between the protocol-defined variables and the reported variables in the published studies. This finding might be due to relatively small number of included studies, resulting from a lack of sufficient protocols publicly available. Besides well-established register strategy of the protocols of clinical trials, hundreds of researchers published protocols of cohort studies. However, only 31 protocols were disclosed before the start of the studies among the thousands of CCSs published online. Since protocol disclosure and register have been ascertained to limit selective reporting of results [40, 41], we strongly recommend that the researchers publish a protocol before the commencement of a CCS. Increased transparency in CCSs will contribute not only to the improvement of study credibility, but also to the reduction of the potential publication bias [42], by which more accurate results

Table 2 Comparisons of factors published in studies with those predefined in protocols

Categories of factors	Median	Quartile	Range
Factors in protocols	12.0	7.5–17.5	3–29
Factors in publications	10.0	3.0–17.5	1–29
Factors not reported in publications	0.0	0.0–3.3	0–12
Factors reported non-significantly in publications	3.5	1.0–7.0	0–12

Factors in protocol vs. factors in publication by a non-parametric analysis: $Z = 0.784$, $P = 0.433$; factors not reported in publication vs. factors reported non-significantly in publication by a non-parametric analysis: $Z = 2.098$, $P = 0.036$

Fig. 3 Bubble chart of correlation between sample size and number of variables reported in case-control studies. Spearman correlation (r_s) = 0.113, P = 0.700

based on personalization will be released according to the optimal approaches in medical services from the perspectives of PPPM [43].

We postulated that sample size of participants would be associated with incomplete reporting of results, and hypothesized that studies with larger sample size would involve less selective reporting. However, the current study disproved this hypothesis by illustrating that sample size of participants is not correlated with selective result reporting. Thus, readers cannot speculate on the extent of potential selective reporting of results simply on the basis of the sample size of a CCS.

Here, we want to emphasize the importance of proper reporting of findings from the perspectives of PPPM. The statistical analyses, in general, are essentially done on particular groups of patients/controls of a particular size. From the point of view of personalized attitude, however, an individual

deviation from the observed trend in the group remains undealt with, neglected or rejected as erroneous, or simply taken as being statistically insignificant, despite the great predictive/diagnostic/prognostic potentials which may be hidden right in these individual/personal “deviations” if the finding is properly reported. Therefore, scientific community may proceed further to search for explanations and available solutions at an individual level. Thus, the practice of selective reporting, as documented, according to our opinion, may have strong impact on the medical attitudes from the PPPM point of view. The medical professionals should take this perspective and undesirable practice of selective reporting into consideration when they author publications, and develop new strategies for PPPM. At the same time, the caution should be applied across all medical journals in order to eliminate selective reporting as much as possible.

Fig. 4 Bubble chart of correlation between sample size of participants and ratio of published factors to factors in protocols. Spearman correlation (r_s) = 0.318, P = 0.268

Limitations

Potential limitations should be acknowledged in this study. Our results might underestimate the extent of selective result reporting since we only evaluated CCSs with an accessible protocol. With regard to protocols published in recent years, the studies might not yet be completed, which results in overestimation of the discrepancies between protocols and published studies. Moreover, there are currently insufficient published protocols of CCSs available publicly, which limited the sample size of this study, resulting in a low power of statistical analyses.

Conclusion

Our findings highlight the need for investigators, peer reviewers, and readers to exercise increased awareness and scrutiny of selective reporting of results, as well as the requirement of explanation of the findings from the perspectives of PPPM. In this way, precise evidences at individual patient level might be released by case-control studies to promote PPPM-related approaches for enhancing quality of personalization in healthcare. As a substantial effort made toward preventing the occurrence of selective result reporting, disclosure of protocol is required for publication of CCSs in all academic journals. Protocols should be publicly available prior to the completion of case-control studies so that the potential bias can be assessed by readers. Moreover, deviations from protocols must be explained in the published CCSs so that the potential bias can be readable.

Author contributions HH and WW designed this study. GD, XZ, ZM, YZ, and ZG collected the data. HH, and DL performed the statistical analysis and interpretation of the data. JX provided statistical expertise. HH, WW, and JX wrote the manuscript. All authors have read and approved the manuscript.

Funding information This work was supported by the Natural Science Foundation of Shandong Province, China (No. ZR2017MH100). The sponsor did not participate in the study design, data collection, data analysis, or manuscript preparation.

Compliance with ethical standards

Competing interests The authors declare that they have no competing interests.

Consent for publication Not applicable.

Ethics approval and consent to participate Not applicable.

References

1. Chan AW, 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(20):2457–65.
2. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med*. 2008;358(3):252–60.
3. 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(9634):201.
4. Zhang S, Liang F, Li W. Comparison between publicly accessible publications, registries, and protocols of phase III trials indicated persistence of selective outcome reporting. *J Clin Epidemiol*. 2017;91:87–94.
5. 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;361:1963–71.
6. You B, Gan HK, Pond G, Chen EX. Consistency in the analysis and reporting of primary end points in oncology randomized controlled trials from registration to publication: a systematic review. *J Clin Oncol*. 2012;30(2):210–6.
7. DeAngelis CD, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, et al. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *JAMA*. 2004;292(11):1363–4.
8. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA*. 1996;276(8):637–9.
9. Jones CW, Handler L, Crowell KE, Keil LG, Weaver MA, Platts-Mills TF. Non-publication of large randomized clinical trials: cross sectional analysis. *BMJ*. 2013;347:f6104.
10. Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC. The Clinical Trials.gov results database update and key issues. *N Engl J Med*. 2011;364:852–60.
11. Zarin DA, Tse T, Sheehan J. The proposed rule for U.S. clinical trial registration and results submission. *N Engl J Med*. 2015;372:174–80.
12. Weerasinghe M, Konradsen F, Eddleston M, Pearson M, Gunnell D, Hawton K, et al. Risk factors associated with purchasing pesticide from shops for self-poisoning: a protocol for a population-based case-control study. *BMJ Open*. 2015;5(5):e007822.
13. Estcourt LJ, Stanworth SJ, Collett D, Murphy MF. Intracranial haemorrhage in thrombocytopenic haematology patients—a nested case-control study: the InCiTe study protocol. *BMJ Open*. 2014;4(2):e004199.
14. Cramond F, Irvine C, Liao J, Howells D, Sena E, Currie G, et al. Protocol for a retrospective, controlled cohort study of the impact of a change in Nature journals' editorial policy for life sciences research on the completeness of reporting study design and execution. *Scientometrics*. 2016;108:315–28.
15. Stosic M, Vukovic D, Babic D, Antonijevic G, Foley KL, Vujcic I, et al. Risk factors for multidrug-resistant tuberculosis among tuberculosis patients in Serbia: a case-control study. *BMC Public Health*. 2018;18(1):1114.
16. Soares AM, Szejnfeld VL, Enokihara MY, Michalany N, Castro CH. High serum 25-hydroxyvitamin D concentration in patients with a recent diagnosis of non-melanoma skin cancer: a case-control study. *Eur J Dermatol*. 2018;28(5):649–53.
17. Canova C, Anello P, Barbiellini Amidei C, Parolin V, Zanier L, Simonato L. Use of healthcare services at the end of life in decedents compared to their surviving counterparts: a case-control study among adults born before 1946 in Friuli Venezia Giulia. *PLoS One*. 2019;14(2):e0212086.
18. Golubnitschaja O, Costigliola V, EPMA. General report & recommendations in predictive, preventive and personalised medicine 2012: white paper of the European Association for Predictive, Preventive and Personalised Medicine. *EPMA J*. 2012;3(1):14.
19. Lu M, Zhan X. The crucial role of multiomic approach in cancer research and clinically relevant outcomes. *EPMA J*. 2018;9(1):77–102.

20. Hou H, Zhao Y, Yu W, Dong H, Xue X, Ding J, et al. Association of obstructive sleep apnea with hypertension: a systematic review and meta-analysis. *J Glob Health*. 2018;8(1):010405.
21. Leone M, Bottacchi E, Beghi E, Morgando E, Mutani R, Amedeo G, et al. A case-control study on alcohol and seizures: study design, protocol, and data collection. *Ital J Neurol Sci*. 1997;18(2):119–24.
22. Heinemann LA, Garbe E, Lewis M, van der Woude F, Graf H. Case-control study on analgesics and nephropathy (SAN): protocol. *BMC Nephrol*. 2005;6:9.
23. Boneberger A, Radon K, Baer J, Kausel L, Kabesch M, Haider D, et al. Asthma in changing environments—chances and challenges of international research collaborations between South America and Europe—study protocol and description of the data acquisition of a case-control-study. *BMC Pulm Med*. 2010;10:43.
24. Shariful Islam SM, Lechner A, Ferrari U, Froeschl G, Niessen LW, Seissler J, et al. Social and economic impact of diabetics in Bangladesh: protocol for a case-control study. *BMC Public Health*. 2013;13:1217.
25. Grigg MJ, William T, Drakeley CJ, Jelip J, von Seidlein L, Barber BE, et al. Factors that are associated with the risk of acquiring *Plasmodium knowlesi* malaria in Sabah, Malaysia: a case-control study protocol. *BMJ Open*. 2014;4(8):e006004.
26. Tovilla-Zárate CA, González-Castro TB, Juárez-Rojop I, Pool García S, Velázquez-Sánchez MP, Villar-Soto M. Study on genes of the serotonergic system and suicidal behavior: protocol for a case-control study in Mexican population. *BMC Psychiatry*. 2014;14:29.
27. Lewis DM, Assmann A, Heinemann L, Spitzer WO. Interim review of the transnational case-control study of oral contraceptives and health: approved protocol revisions through September 1995. *Pharmacoepidemiol Drug Saf*. 1996;5(1):43–51.
28. Wynn P, Stewart J, Kumar A, Clacy R, Coffey F, Cooper N, et al. Keeping children safe at home: protocol for a case-control study of modifiable risk factors for scalds. *Inj Prev*. 2014;20(5):e11.
29. Heinemann LA, Thomas DB, Möhner M. Multicentre international liver tumour study protocol of the case-control study on hepatocellular cancer. *Pharmacoepidemiol Drug Saf*. 1996;5(3):173–86.
30. Klusmann A, Gebhardt H, Liebers F, von Engelhardt LV, Dávid A, Bouillon B, et al. Individual and occupational risk factors for knee osteoarthritis - study protocol of a case control study. *BMC Musculoskelet Disord*. 2008;9:26.
31. Picot VS, Bénet T, Messaoudi M, Telles JN, Chou M, Eap T, et al. Multicenter case-control study protocol of pneumonia etiology in children: global approach to biological research, infectious diseases and epidemics in low-income countries (GABRIEL network). *BMC Infect Dis*. 2014;14:635.
32. Gomez-Marcos MA, Gonzalez-Sarmiento R, Recio-Rodríguez JJ, Agudo-Conde C, Gamella-Pozuelo L, Perretta-Tejedor N, et al. Relationship between target organ damage and blood pressure, retinal vessel calibre, oxidative stress and polymorphisms in VAV-2 and VAV-3 genes in patients with hypertension: a case-control study protocol (LOD-Hipertension). *BMJ Open*. 2014;4(4):e005112.
33. Vinogradova Y, Coupland C, Hippisley-Cox J. Exposure to combined oral contraceptives and risk of venous thromboembolism: a protocol for nested case-control studies using the QResearch and the CPRD databases. *BMJ Open*. 2014;4(4):e004499.
34. Buckley CM, Ali F, Roberts G, Kearney PM, Perry IJ, Bradley CP. Timing of access to secondary healthcare services for diabetes management and lower extremity amputation in people with diabetes: a protocol of a case-control study. *BMJ Open*. 2013;3(10):e003871.
35. Summerskill W, Collingridge D, Frankish H. Protocols, probity, and publication. *Lancet*. 2009;373(9668):992.
36. Schulz KF, Altman DG, Moher D. Protocols, probity, and publication. *Lancet*. 2009;373(9674):1524.
37. Cheng T, Zhan X. Pattern recognition for predictive, preventive, and personalized medicine in cancer. *EPMA J*. 2017;8(1):51–60.
38. Chan AW, Krljeza-Jerić K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *CMAJ*. 2004;171(7):735–40.
39. Mathieu S, Boutron I, Moher D, Altman DG, Ravaud P. Comparison of registered and published primary outcomes in randomized controlled trials. *JAMA*. 2009;302(9):977–84.
40. Wayant C, Scheckel C, Hicks C, Nissen T, Leduc L, Som M, et al. Evidence of selective reporting bias in hematology journals: a systematic review. *PLoS One*. 2017;12(6):e0178379.
41. Calmèjane L, Dechartres A, Tran VT, Ravaud P. Making protocols available with the article improved evaluation of selective outcome reporting. *J Clin Epidemiol*. 2018;104:95–102.
42. Swaen GMH, Urlings MJE, Zeegers MP. Outcome reporting bias in observational epidemiology studies on phthalates. *Ann Epidemiol*. 2016;26(8):597–9.
43. Golubnitschaja O, Baban B, Boniolo G, Wang W, Bubnov R, Kapalla M, et al. Medicine in the early twenty-first century: paradigm and anticipation - EPMA position paper 2016. *EPMA J*. 2016;7:23.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.