Are the Subgroup Analyses in Stroke Clinical Trials Credible? – Protocol for a Systematic Review

Ayoola Ademola (✉ ayoola.ademola@ucalgary.ca)

University of Calgary   https://orcid.org/0000-0002-8412-6246

Kevin A Hildebrand

University of Calgary

Samuel Babatunde

University of Calgary

Darren Mazzei

University of Calgary

Michael D Hill

university of calgary

Lehana Thabane

McMaster University

Tolulope T Sajobi

University of Calgary

Protocol

Abstract

Background: Clinicians and decision-makers are mostly interested in how the overall effect in a randomized controlled trial impact patient differently due to differential treatment effects. Subgroup analyses of randomized trials examine if the overall treatment effect is consistent among the subgroup or to identify groups that modify the intervention effects. Well defined and performed subgroup analyses led to critical policy decisions; however, inappropriately performed subgroup analyses had resulted in incorrect decisions with negative consequences. This systematic review examines the reporting quality and the subgroup analyses' credibility in stroke clinical trials.

Methods and analysis: We will extract relevant studies from PubMed, Embase, Cochrane Central Register of Controlled Trials (CENTRAL), and Web of Science using three concepts in the Medical Subject Headings (MESH) heading. Two reviewers will independently screen the title, abstract, and full text of relevant studies. We will examine the risk of bias and the credibility of reported subgroup analyses using two validated instruments, i.e., Cochrane Risk-of-Bias tool for randomized trials version 2 (RoB 2) and the Instrument for assessing the Credibility of Effect Modification Analyses (ICEMAN), respectively. Random effects regression will be used to evaluate study characteristics associated with the credibility of subgroup analyses in the included studies.

Expected outcomes: This research aims to review the quality of subgroup analyses' results and reporting. The research's results will also provide critical methodological contributions to the statistical literature of clinical trials.

Ethics and dissemination: This systematic review does not require primary patient data and does not require ethical approval. The review's results will be published in a peer-review and scientific conferences.

Trial registration number: CRD42020223133

Background

Randomized controlled trial (RCT) is generally considered the gold standard for obtaining the highest level of evidence about the efficacy or effectiveness of medical interventions [1-3]. Despite being foundational to numerous medical breakthroughs and innovations in medical research, the generalizability of RCT's results and its external validity had been a critical discussion over the years [4, 5]. Sometimes, the overall treatment effect in a trial is less appealing to clinicians since they are more interested in how a new treatment affects their patients. Although patients in RCTs may seem homogenous, the intervention's impact might vary across different patient demographic or clinical characteristics. Hence, subgroup analysis is essential if there are potentially large differences between groups, heterogeneity of treatment effects, pragmatic questions about when to treat, and confusion about benefit or harm in specific groups[6, 7]. Subgroup analysis aims to determine the existence of variability in treatment effects among patients with different baseline characteristics. When adequately performed, subgroup analyses produce reliable information for decision-making and generated future hypotheses. It

also provides an opportunity to broaden treatment to groups of patients previously omitted in clinical practice and prevent harming patients with ineffective treatment.

Nevertheless, poorly conducted subgroup analyses had been shown to generate incorrect outcomes due to the use of inappropriate statistical analysis, false-positive results from multiple testing, and/or false-negative results due to low statistical power [8, 9]. The Second International Study of Infarct Survival (ISIS-2) trial reported that patients with myocardial infarction born under Gemini or Libra's zodiac signs had reduced vascular mortality compared to patients born under other zodiac signs [10]. The subgroup analysis result from the randomized trial of aspirin and sulfinpyrazone in stroke patients stroke had aspirin as an ineffective treatment for secondary stroke prevention in women until this result was proven to be inaccurate in a later trial [11, 12]. Therefore, researchers have an ethical obligation to conduct subgroup analysis to the highest standards.

More importantly, few systematic reviews have been conducted to investigate the quality and reporting of subgroup analyses in clinical trials. However, the included studies in these reviews were obtained from only one electronic biomedical journal, sampled for available studies, or randomly selected by stratifying in ratio 1:1 between high/low impact journals [13-16]. Likewise, there had been proposed guidelines and checklists for evaluating the plausibility of subgroup analyses' results to address inconsistencies in reporting subgroup analyses' results [7, 8, 17-20]. However, these proposed guidelines and checklists consisted of long lists of items and were not systematically developed or validated. Recently, Schandelmaier et al. developed a brief and validated the Instrument for assessing the Credibility of Effect Modification Analyses (ICEMAN) checklist from the existing guidelines reported in the literature [21]. ICEMAN, which was systematically developed through a consensus of fifteen experts and tested by seventeen potential users, consists of five-core questions for RCTs and eight-core questions for meta-analysis with four-level Likert scale response options. Thus, these checklists' recommended subgroup analyses to be prespecified and justified with few clinically important questions tested. Also, every subgroup analysis performed should be reported with reproducible results and supported by other studies.

Research objective

The primary objective of this systematic review is to evaluate the credibility of subgroup analyses in stroke trials. Also, we will examine the reporting quality of stroke trials. The review's results will provide suggestions on how to improve stroke trials the reporting standard.

Methods And Analysis

Protocol design and registration

The design, conduct, and reporting of this systematic review will follow the Cochrane Handbook of Systematic Reviews of Interventions [22]. The protocol has been registered with PROSPERO, an international database of prospectively registered systematic reviews to prevent duplication (Trial

registration number: CRD42020223133). Considering this is a methodological review, we will adopt the Studies, Data, Methods, Outcomes (SDMO) framework to address critical questions in this review [23]. The study design (S) focuses on RCTs due to its level of evidence [24]. The data (D) are the results from original published studies in biomedical journals (PubMed, Embase, Cochrane Central Register of Controlled Trials (CENTRAL), and Web of Science). The method (M) compares the quality of reporting subgroup analyses between pre- and post-publication of the updated the Consolidated Standards of Reporting Trials (CONSORT) guidelines [25-27]. Lastly, the outcome (O) is to determine if there is an improvement in the quality subgroup analysis reporting in stroke clinical trials after the publication of the CONSORT statement.

Eligibility criteria and search strategy

We will include studies with subgroup analyses in stroke randomized controlled trials by searching the electronic biomedical databases up to November 30th, 2020. Studies in this systematic review will include (1) randomized controlled trials, (2) stroke trials, (3) trials with subgroup analysis, effect modification, or interaction term, and (4) trials with published protocols or trial registration. Hence, we will exclude (1) subgroup analysis defined by intervention characteristics, (2) studies in not in human, (3) systematic review, literature review, or meta-analysis, (4) grey literature or conference abstracts, and (5) non-English language publications. The flowchart of the inclusion and exclusion is in Figure 1. The search strategy will include three main concepts, i.e., randomized controlled trial, stroke, and subgroup analysis/effect modifications. Below in Table 1 is the detailed Medical Subject Headings (MeSH) for Medline electronic database and Figure 1A in the appendix for the Medline's search result.

Two reviewers (AA and SB) will independently execute the search strategy in each of the preselected electronic databases and compare the number of articles obtained from each database. The reviewers will upload these identified studies into Covidence, a web-based application for undertaking systematic reviews [28]. Covidence will automatically remove duplicated studies and provide a standard template where the reviewers can independently review them.

Selection of studies

AA and SB will separately screen titles and abstracts of identified articles using both inclusion and exclusion criteria. Likewise, AA and SB will individually perform a full-text review of all eligible articles and extract relevant information. Unresolved disagreement on any article between the two reviewers will be marked "Further review required." They will forward such articles to LT, MDH, and TTS, who serves as an adjudicator making the final decision. The results will summarize in the PRISMA flow diagram [29]. Also, we will compute and report the interrater reliability score.

Data Items and Outcomes

The primary outcome is the credibility of subgroup analyses results of published stroke randomized controlled trials. The outcome will be obtained using ICEMAN to evaluate each subgroup analyses. The

main measure of effect is the proportion of subgroup analysis with credible claim. AA and SB will extract relevant characteristics reported for each study' subgroup analyses after a thorough full-text review. Data items to be extracted are the first author name, country of the first author, year of publication, journal of publication, study design, trial intervention, prespecified analysis in the original protocol, and the number of subgroups analyzed. Other information includes sample size, type of endpoint for subgroup analysis (primary or secondary), type of endpoints (binary, continuous, and ordinal), analysis type (linear regression, generalized linear model, or proportional regression). Also, AA and SB will extract the statistical significance of effect modification (subgroup analysis), and the magnitude of overall treatment effect size reported (small, median, large). More detailed relevant characteristics to be extracted from each study are in Appendix Table 1A.

Risk of bias and assessment of subgroup analyses' credibility.

We will use two different validated assessment tools to appraise the study design's quality and the subgroup analyses' trustworthiness of each study. A poorly conducted trial invalidates such trial's results; therefore, we will assess each study design through the Cochrane Risk-of-Bias tool for randomized trials version 2 (RoB 2) [30]. This instrument will evaluate pertinent study qualities such as participants' selection, randomization, concealment, and allocation. Next, we will use the ICEMAN checklists to evaluate each study subgroup analyses' credibility. The ICEMAN has five-core questions for RCTs. Each core question has four response options that ranged from "probably reduced credibility" to "probably increased credibility. Reviewers (AA and SB) will independently use these checklists to grade each study's quality and credibility of reported subgroup analyses for these studies. LT, MDH, and TTS will adjudicate any disagreement between the two reviewers. Based on RoB 2, we will classify each study ranging from low to high quality. AA and SB will use the ICEMAN checklists to assess each subgroup analysis credibility through the four-rating score, i.e., from very low credibility to high credibility.

Data analysis (include analysis based on ICEMAN)

Descriptive analysis of all the extracted data from all the included studies will be conducted. We will analyze the extracted data using both ordinal, binary and cluster analysis. These results will be stratified by the reported studies' credibility, type of intervention, whether the subgroup analysis result is significant, population group (acute versus non-acute stroke), and publication journal. Mixed-effect regression will be used to evaluate the impact of study characteristics (publication journal, sample size, type of analysis, type of endpoint [primary or secondary], disease population, intervention, and risk of bias) on the credibility of subgroup analysis reported in the studies. Trend analysis of the reporting will be performed to appraise any improvement in the quality of reporting over the last few years. Lastly, we will analyze the impact of publication journals on the reporting quality trials' results [31, 32].

Discussion

The goal of this systematic review is to evaluate the reporting quality of the results of stroke trials. The review will also examine the subgroup analyses' credibility of those with or without a claim of statistical

significance. Adequately performed subgroup analyses could lead to changes in clinical policy; however, it is essential that researchers appropriately perform these subgroup analyses. Findings from this study will provide an up-to-date picture of the quality of reporting of subgroup analysis in RCTs and recommendations on how to improve the credibility of subgroup analyses in stroke trials.

The strength of this review is independence, transparency, and reproducibility of results. Two reviewers will independently obtain eligible studies and data used in this systematic review. We will publish the methods, the MESH terms, and the different analyses to ensure the reproducibility of results at any time. A possible limitation might have been the inability to perform meta-analysis; however, this review is about answering a methodological question rather than a particular treatment effect. Regardless, we believe this systematic review will provide information on the quality of reporting subgroup analyses in stroke clinical trials, enabling us to determine if there are improvements in these results. Also, we believe the results will evaluate the impact of publication journals on the quality of reporting of clinical studies. Lastly, we will be able to provide appropriate suggestions on the areas of reporting requiring enhancement.

List Of Abbreviations

S/N	Abbreviation	Meaning
1	Embase	Excerpta Medica Database
2	CENTRAL	Cochrane Central Register of Controlled Trials
3	MESH	Medical Subject Headings
4	RoB 2	Cochrane Risk-of-Bias tool for randomized trials version 2
5	ICEMAN	The Credibility of Effect Modification Analyses
6	PROSPERO	The International Prospective Register for Systematic Reviews
7	RCT	Randomized controlled trial
8	ISIS-2	The Second International Study of Infarct Survival
9	SDMO	The Studies, Data, Methods, Outcomes
10	AA	Ayoola Ademola
11	CONSORT	The Consolidated Standards of Reporting Trials
12	SB	Samuel Babatunde
13	LT	Lehana Thabane
14	MDH	Michael D. Hill
15	TTS	Tolulope T. Sajobi
16	US	United States

Declaration

Acknowledgments

Not applicable.

Funding

Availability of data and materials

The data obtained or analyzed during this study will be included in the published article (and its supplementary information files.)

Competing interests

Not applicable.

Consent to publish

Not applicable (no patient information or details is included in the manuscript).

Ethical considerations

Not applicable. The review does not involve a human study.

Authors Contribution

AA, TTS, and KH:- developed the concept; AA:- drafted, extracted data, reviewed and submitted the manuscript; SB:- extracted data, reviewed the manuscript; DM:- reviewed the manuscript; MDH:- reviewed the manuscript; LT:- reviewed of the manuscript; KAH:- reviewed the manuscript & TTS:- reviewed the manuscript. The final manuscript was read and approved by all authors.

Reference

1. Akobeng, A.K., *Understanding randomised controlled trials.* Arch Dis Child, 2005. **90**(8): p. 840-4.

2. Hariton, E. and J.J. Locascio, *Randomised controlled trials - the gold standard for effectiveness research: Study design: randomised controlled trials.* BJOG, 2018. **125**(13): p. 1716.

3. Stang, A., C. Poole, and O. Kuss, *The ongoing tyranny of statistical significance testing in biomedical research.* European Journal of Epidemiology, 2010. **25**(4): p. 225-230.

4. Ahn, C. and D. Ahn, *Randomized clinical trials in stroke research.* Journal of investigative medicine : the official publication of the American Federation for Clinical Research, 2010. **58**(2): p. 277-281.

5. Susukida, R., et al., *Generalizability of findings from randomized controlled trials: application to the National Institute of Drug Abuse Clinical Trials Network.* Addiction (Abingdon, England), 2017. **112**(7): p. 1210-1219.

6. Rothwell, P.M., *Can overall results of clinical trials be applied to all patients?* Lancet, 1995. **345**(8965): p. 1616-1619.

7. Rothwell, P.M., *Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation.* Lancet, 2005. **365**(9454): p. 176-86.

8. Brookes, S.T., et al., *Subgroup analyses in randomised controlled trials: quantifying the risks of false-positives and false-negatives.* Health Technol Assess, 2001. **5**(33): p. 1-56.

9. Schulz, K.F. and D.A. Grimes, *Multiplicity in randomised trials II: subgroup and interim analyses.* Lancet, 2005. **365**(9471): p. 1657-1661.

10. Group, I.-C., *Randomized trial of intravenous streptokinase, oral aspirin, both, or neither among 17,187 cases of suspected acute myocardial infarction: ISIS-2.* Journal of the American College of Cardiology, 1988. **12**(6): p. A3-A13.

11. Ridker, P.M., et al., *A Randomized Trial of Low-Dose Aspirin in the Primary Prevention of Cardiovascular Disease in Women.* New England Journal of Medicine, 2005. **352**(13): p. 1293-1304.

12. The Canadian Cooperative Study, G. and G. Canadian Cooperative Study, *A Randomized Trial of Aspirin and Sulfinpyrazone in Threatened Stroke.* N Engl J Med, 1978. **299**(2): p. 53-59.

13. Fan, J., F. Song, and M.O. Bachmann, *Justification and reporting of subgroup analyses were lacking or inadequate in randomized controlled trials.* J Clin Epidemiol, 2019. **108**: p. 17-25.

14. Sun, X., et al., *Credibility of claims of subgroup effects in randomised controlled trials: systematic review.* BMJ, 2012. **344**: p. e1553.

15. Wang, R., et al., *Statistics in Medicine — Reporting of Subgroup Analyses in Clinical Trials.* N Engl J Med, 2007. **357**(21): p. 2189-2194.

16. Zhang, S., et al., *Subgroup Analyses in Reporting of Phase III Clinical Trials in Solid Tumors.* J Clin Oncol, 2015. **33**(15): p. 1697-702.

17. Dijkman, B., B. Kooistra, and M. Bhandari, *How to work with a subgroup analysis.* Can J Surg, 2009. **52**(6): p. 515-522.

18. Guillemin, F., *Primer: the fallacy of subgroup analysis.* Nat Clin Pract Rheumatol, 2007. **3**(7): p. 407-413.

19. Oxman, A.D. and G.H. Guyatt, *A consumer's guide to subgroup analyses.* Ann Intern Med, 1992. **116**(1): p. 78-84.

20. Sun, X., et al., *Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses.* BMJ, 2010. **340**: p. c117.

21. Schandelmaier, S., et al., *Development of the Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in randomized controlled trials and meta-analyses.* Canadian Medical Association Journal, 2020. **192**(32): p. E901-E906.

22. Cumpston, M., et al., *Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions.* Cochrane Database Syst Rev, 2019. **10**: p. ED000142-ED000142.

23. Munn, Z., et al., *What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences.* BMC medical research methodology, 2018. **18**(1): p. 5-5.

24. Spieth, P.M., et al., *Randomized controlled trials - a matter of design.* Neuropsychiatr Dis Treat, 2016. **12**: p. 1341-9.

25. Eldridge, S.M., et al., *CONSORT 2010 statement: extension to randomised pilot and feasibility trials.* Pilot Feasibility Stud, 2016. **2**: p. 64.

26. Moher, D., et al., *CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials.* BMJ, 2010. **340**: p. c869.

27. Schulz, K.F., et al., *CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials.* BMC Med, 2010. **8**: p. 18.

28. Jessica, B., *Product Review: Covidence (Systematic Review Software).* Journal of the Canadian Health Libraries Association, 2014. **35**(2).

29. Stovold, E., et al., *Study flow diagrams in Cochrane systematic review updates: an adapted PRISMA flow diagram.* Systematic Reviews, 2014. **3**.

30. Sterne, J.A.C., et al., *RoB 2: a revised tool for assessing risk of bias in randomised trials.* BMJ, 2019. **366**.

31. Shamseer, L., et al., *Update on the endorsement of CONSORT by high impact factor journals: a survey of journal "Instructions to Authors" in 2014.* Trials, 2016. **17**(1): p. 301.

32. Zwarenstein, M., et al., *Improving the reporting of pragmatic trials: an extension of the CONSORT statement.* BMJ, 2008. **337**: p. a2390.

Table

Table 1:

Medline search strategy

Database	Randomized controlled trials [1]	Stroke	subgroup analysis/effect modifications

Medline	#1 randomized controlled
	
	        trial.pt.
	
	#2 controlled clinical trial.pt.
	
	#3 randomized.ab.
	
	#4 placebo.ab.
	
	#5 drug therapy.fs.
	
	#6 randomly.ab.
	
	#7 trial.ab.
	
	#8 groups.ab.
	
	#9 #1 or #2 or #3 or #4 or
	
	        #5 or #6 or #7 or #8
	
	#10 animals/ not humans/
	
	#11 #9 not #10
		#1 stroke.tw,kf.
		
		#2 brain Infarction*.tw, kf.
		
		#3 brain stem
		
		        infarction*.tw, kf.
		
		#4 lateral medullary
		
		        syndrome*.tw, kf.
		
		#5 cerebral
		
		        infarction*.tw, kf.
		
		#6 cadasil.tw,kf.
		
		#7 dementia, multi-
		
		        infarct.tw, kf.
		
		#8 anterior cerebral artery
		
		        infarction*.tw, kf.
		
		#9 middle cerebral artery
		
		        infarction*.tw, kf.
		
		#10 posterior cerebral
		
		        artery infarction.tw,kf.
		
		#11 intracerebral
		
		hemorrhage.tw,kf.
		
		#12 subarachnoid
		
		hemorrhage.tw,kf.
		
		#13 exp Infarction, Anterior
		
		        Cerebral Artery/ or
			#1 ((sub-group* or
			
			        subgroup*) adj2
			
			        (analys* or effect*)).tw,kf.
			
			#2 interaction term*.tw, kf.
			
			#3 effect* modification*.tw, kf.
			
			#4 modification* effect*.tw,
			
			        kf.
			
			#5 Statistical interaction.tw,kf.
			
			#6 Differential effect*.tw, kf.
			
			#7 Heterogeneity of
			
			        treatment Effect*.tw, kf.
			
			#5 #1 or #2 or #3 or #4 or #5
			
			        or #6 or #7

Infarction, Middle
Cerebral Artery/
#14 exp Infarction,
Posterior Cerebral
Artery/ or stroke/ or
brain infarction/
#15 #1 or #2 or #3 or #4 or
#5 or #6 or #7 or #8 or
#9 or #10 or #11 or #12
or #13 or #14

Reference List

1. Cumpston, M., et al., *Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions.* Cochrane Database Syst Rev, 2019. **10**: p. ED000142-ED000142.

Appendix

Table 1A . Study characteristics

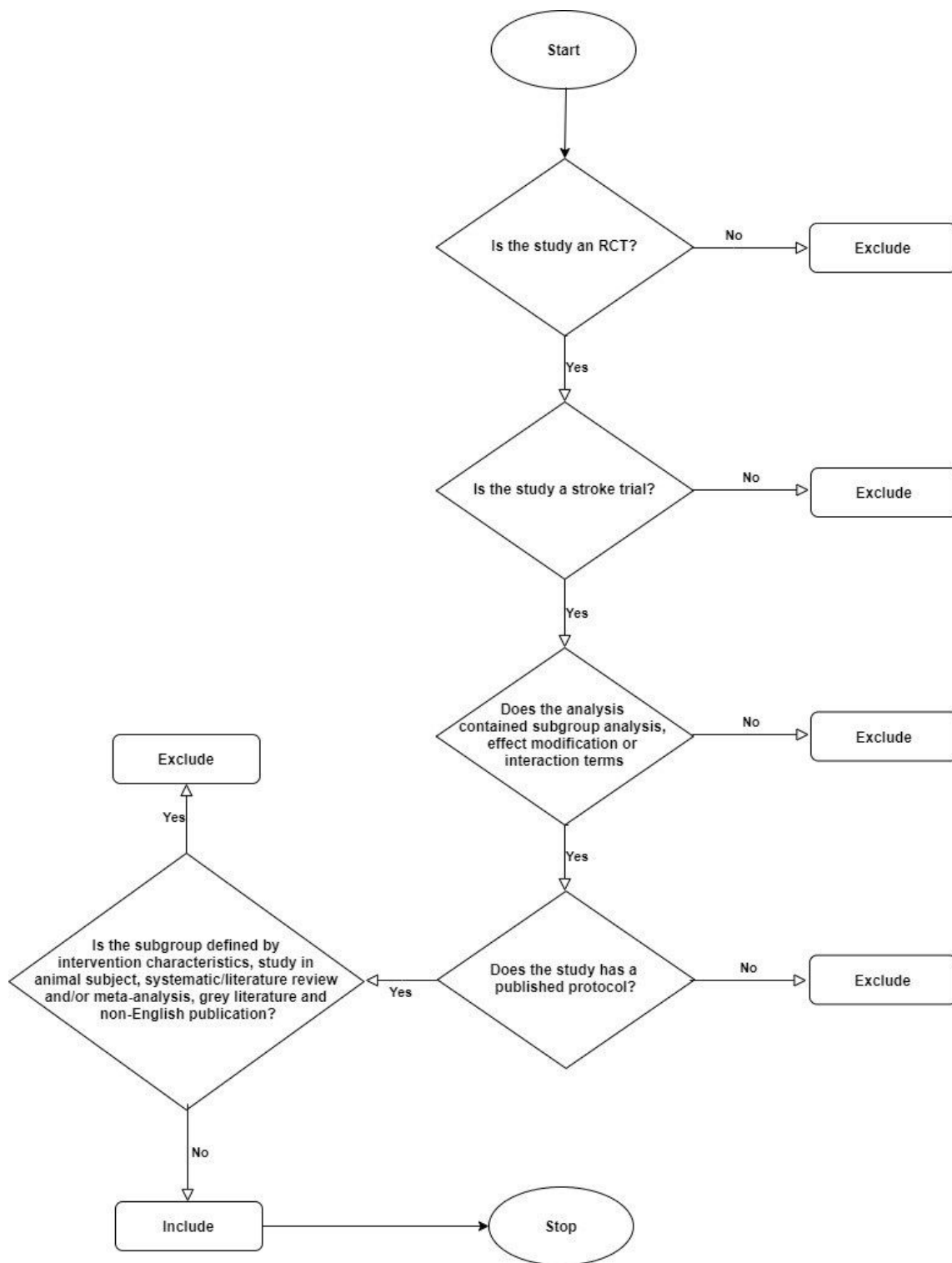| Publication characteristics	Study design	Subgroup analysis/Effect medication
1) First author-name 2) Country of the first author 3) Year of publication 4) Journal of publication	1) Study design (individually randomized parallel-group trial, cluster-randomized parallel-group trial, or individually randomized cross-over (or other matched) trial) 2) Population group (acute versus non-acute stroke) 3) Sample size 4) Number of treatment groups 5) Single or multicenter (i.e., number of centres) 6) Number of baseline characteristics 7) Type of intervention 8) Type of endpoints (binary, continuous, and ordinal) 9) Pragmatic or explanatory trials 10) Type of outcome measure 11) Randomization process (randomization sequence random, concealment, balanced) 12) Effect of assignment to intervention (patients; consent, deviation from treatment, blinding) 13) Effect of adhering to an intervention 14) Selection of the reported result process 15) Measurement of the outcome 16) Missing outcome data	1) Was the analysis predefined? 2) Number of subgroup analyses 3) Was the test of interaction used? 4) What the subgroup analysis result is significant? 5) Were all subgroup analyses reported? 6) Was there evidence to support the subgroup analysis? 7) Was the continuous covariate not dichotomized, avoiding arbitrary cut points? 8) Was the significance level adjusted for multiple subgroup analyses? 9) Statistical analysis type (descriptive only, subgroup p-values, models/interaction test (linear regression, generalized linear model, or proportional regression)). 10) Was the overall primary outcome significance? 11) Were the subgroup analyses results interpreted correctly? Appropriate?

Figures

Figure 1

Inclusion and exclusion flowchart.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Appendix.docx