Protocol

Erroneous risks in pharmaceutical versus non-pharmaceutical interventions during data extraction in evidence synthesis practice: Study protocol for a randomized controlled trial

Yi Zhu[1,2,3], Pengwei Ren[4], Suhail A.R. Doi[5], Luis Furuya-Kanamori[6], Lifeng Lin[7,8], Xiaoqin Zhou[9], Fangbiao Tao[1,2,3], Chang Xu[1,2,3]✉

[1]Key Laboratory for Population Health Across-life Cycle, Ministry of Education, Anhui, China

[2]Anhui Provincial Key Laboratory of Population Health and Aristogenics, Anhui Medical University, Anhui, China

[3]School of Public Health, Anhui Medical University, Anhui, China

[4]Department of Clinical Research Center for Respiratory Diseases, West China Hospital, Sichuan University, Chengdu, Sichuan, China

[5]Department of Population Medicine, College of Medicine, QU Health, Qatar University, Doha, Qatar

[6]UQ Centre for Clinical Research, Faculty of Medicine, University of Queensland, Brisbane, Australia

[7]Department of Epidemiology and Biostatistics, University of Arizona, Tucson, AZ, USA

[8]Department of Statistics, Florida State University, Tallahassee, FL, USA

[9]Department of Clinical Research Management, West China Hospital, Sichuan University, China

Correspondence: Chang Xu, Ministry of Education Key Laboratory for Population Health Across-life Cycle, Anhui Medical University, Anhui, China; xuchang2016@runbox.com

Running title: Data reproducibility issues in evidence synthesis

Funding: This work will be supported by a seed fund for talented earlier researchers from Anhui Medical University (9021783201) and program grant #NPRP-BSRA01-0406-210030 from the Qatar National Research Fund (a member of Qatar Foundation).

Word count: 2760

Abstract

Background: Data extraction is the foundation for a trusted systematic review; it is often time- and labor- consuming and prone to errors. Whilst data extraction errors frequently occur in literatures, an interesting phenomenon was observed that such erroneous risks tend to be higher in trials of pharmaceutical interventions compared to non-pharmaceutical one. This phenomenon has not been verified by high-quality evidence; The elucidation of which would have implications for guidelines, practice and policy.

Methods and analyses: We prepare a 1:1 designed, randomized, multicenter, investigator-blinded, cross-over trial to elucidate the potential variants on the error rates of meta-analyses with pharmaceutical against non-pharmaceutical interventions. Eligible participants that at the 2nd year of their post-graduate program or above (e.g., doctoral program) will be recruited and randomly assigned to one of the two groups to complete pre-defined data extraction tasks: 1) group A contains 10 randomized controlled trials (RCTs) of pharmaceutical interventions; 2) group B contains 10 RCTs of non-pharmaceutical interventions. Participants who would finish the data extraction will be further assigned to the alternative group for another round of data extraction, after at least 30 mins break. Finally, those participants assigned to A or B group will be further 1:1 randomly matched based on a random-sequenced number, for the double-checking process on the extracted data. The primary outcome will be the error rates on data extraction of the pharmaceutical intervention group and the non-pharmaceutical group, *before* the double-checking process, in terms of the cell level, study level, and participant level. The secondary outcome will be the error rates on data extraction of the pharmaceutical intervention group and the non-pharmaceutical group, *after* the double-checking process, again, in terms of the cell level, study level, and participant level. The generalized linear mixed effects model (based on the above three levels) will be used to estimate the potential differences on the error rates, with a link function of log that refers to binomial distribution. Subgroup analyses are planned by the following factors: the experience of individuals on systematic reviews, time used for the data extraction.

Ethics and dissemination: This study has been approved by the institutional review board of Anhui Medical University. We plan to present our findings at international scientific meetings, and publish our findings in a high-quality peer-reviewed academic journal.

Trial registration: Chinese Clinical Trial Register Center (Identifier: ChiCTR2200062206).

Keywords: Data extraction error; pharmaceutical intervention; non-pharmaceutical intervention; Randomized controlled trial

Strengths and limitations of the study

- This will be the first trial to test the error rates of data extraction in the sight of intervention type that characterized by pharmaceutical intervention or non-pharmaceutical intervention.
- This will be the third randomized trial on strategy of data extraction in the world and the first in Asian-Pacific area.
- The use of a cross-over design provides a valid way to eliminate the potential impact of the heterogeneous contexts of the studies and thus is expected to provide less biased evidence to support a better evidence synthesis practice.
- The restrict the participants to 2nd year post-graduate students or above to ensure the feasibility of the trial, which is no doubt impact the representativeness of the samples.
- A group of useful strategies should be taken to minimize the impact of the possible sharing of completed extraction table among the participants.

Introduction

In an era of evidence-based medicine, evidence is the backbone in healthcare practice that it governs the guideline-developing, decision-making, as well as policy-formulating [1]. Systematic review and meta-analysis has been one of the most important sources of evidence and thus the validity of such evidence directly determines the reliability and the quality of the healthcare administration [2]. Unfortunately, in real-world, the evidence from systematic reviews and meta-analyses were far from valid or trust-worthy due to a package of reasons, where one of which would be the frequently occurred errors during data extraction — as recorded in previous literatures, as much as 85% of the systematic reviews faces serious issue in data reproducibility [3-7].

Data extraction is a crucial step for any type of evidence synthesis researches, it undertaken the very role of information transformation, from one 'node' to another. This means any error during this process would inevitably distort the original information and thus may bias the final evidence. In our recent research, in 288 meta-analyses with data extraction errors, 12.8% (n = 39) were moderately or largely impacted in terms of the magnitude of the effects [7]. On this base, the Cochrane Collaboration have highlighted the importance of qualified data extraction for informed decision-making and recommended the application of the duplicate data extraction strategy in Cochrane reviews [8]. Other well-known guidelines or checklists, such as the AMSTAR (Assessing the Methodological Quality of Systematic Reviews), also highlighted the importance of good data extraction practice [9].

While evidence from two randomized trials suggested some benefits of duplicate data extraction for improving the data reproducibility [10, 11]; Our empirical investigations suggested that the error rates may differ by type of interventions that those meta-analyses with pharmaceutical interventions had almost doubled error rate

than those with non-pharmaceutical interventions [6,7]. The above findings aroused our interests that why the error rates differ in meta-analyses with pharmaceutical over non-pharmaceutical interventions and whether duplicate data extraction could 'trade-off' such preference on the occurs of errors?

Therefore, in this protocol, we described a planned randomized, multicenter, investigator-blinded, cross-over trial which with the aim to elucidate the aforementioned questions.

Methods

Ethics and trial registration

This study has been approved by the institutional review board of Anhui Medical University (No.83220405), and was registered at the Chinese Clinical Trial Register Center (Identifier: ChiCTR2200062206). The study is designed in line with the CONSORT statement [12], and the reporting of current protocol follows the SPIRIT (Standard Protocol Items: Guidelines for Interventional Trials) 2013 checklist [13].

Study design and settings

This is a 1: 1 designed, randomized, multicenter, investigator-blinded, cross-over trial. The trial will be conducted in three centers in China as of Anhui Medical University, Taihe Hospital, and Guizhou Provincial People's Hospital. For the three centers, Anhui Medical University is the leading medical university in Anhui province, Taihe Hospital is the leading hospital in Shiyan city and the top 10 hospital in Hubei province, and Guizhou Provincial People's Hospital is one of the three leading hospitals in Guizhou province. All of the three centers are undertaking research and teaching programs of evidence-based medicine for post-graduate students.

The full trial will involve three periods according to the standard data extraction process (Figure 1). In the first period, participants will be randomly allocated into one of the two groups for data extraction that with studies of pharmaceutical interventions or with studies of non-pharmaceutical interventions that prepared in advance; Then, in the second period, those participants perform data extraction on studies of pharmaceutical interventions will be switched to the group of non-pharmaceutical interventions for another round of data extraction after they finish the first period, and vice versa (i.e., non-pharmaceutical switched to pharmaceutical group); In the third period, participants in of the two groups assigned in the first period will be randomly matched (1:1) for a further double-checking on their extracted data. Among each of the three periods, the participants will be granted for at least 1 hour's break before they enter to the next period.

The studies used for data extraction in the two groups will be identified before the start of the trial, based on our previous well-established database of meta-analyses of adverse events (binary outcomes) [7]. The database covers 201 systematic reviews of randomized controlled trials (RCTs) with 829 meta-analyses on pharmaceutical or non-pharmaceutical interventions, and all the meta-analytic data in the database had been carefully checked for its validity. The lead author will select one meta-analysis of pharmaceutical intervention and one of non-pharmaceutical intervention and remove the existing meta-analytic data to form a Excel sheet with only study lists and necessarily column titles (but without data) as the template for data extraction, separately (Table 1). By reviewing previous meta-epidemiological studies, we set the number of studies for data extraction in each group as 10 [14, 15]; And therefore, each of the selected meta-analyses should contain at least 10 studies. If the number of included studies exceed 10 in the selected meta-analysis, a simple random sampling scheme will be employed to randomly select 10 studies.

Participants, Interventions, and Control

Individuals with medical or health science backgrounds that are learning systematic

reviews, preparing an ongoing systematic review, or already have experience in conducting systematic reviews are eligible to participate. This may include clinicians, nurses, healthcare policy makers, medical scientists, and medical students. Whereas English is not the native language of Chinese that individuals without qualified English reading skills may have poor performance on data extraction, and the fact that medical students played a main role in data extraction in the majority of the published systematic reviews, we only consider students at the 2nd year of their post-graduate program and above (e.g., doctoral program). Based on a pilot training, we expect the time for data extraction in the three periods would be about 3 to 5 hours; to this end, each participant will be offered 150 RMB (about 22 USD, 4.5 to 7.5 USD/hour) as financial reward if they complete all the three periods.

The primary aim of the trial is to examine the error rates of data extraction in RCTs of pharmaceutical over non-pharmaceutical interventions and the role of duplicate extraction for reducing error rates in evidence synthesis practice. Thus, the intervention of current trial is the double-checking scheme in the third period of the trial. While for the control, based on the design of this trial, two controls will be involved. The first control is single data extraction, namely, data extraction with only one individual (first and second period), without the involvement of any other individuals. For the single data extraction, self-checking is allowed. The second control is the non-pharmaceutical RCTs, compared with the pharmaceutical RCTs, in terms of the data extraction error.

Randomization, blinding and allocation concealment

A third party, not involved in the trial, will generate the random sequence using a computer random-number generator before the enrollment. Participants will take part in the baseline evaluation (e.g., age, expertise, experience on systematic reviews). After providing written informed consent, eligibility criteria will be assessed. Then, participants will be randomized to one of the two groups with a 1:1 ratio through simple randomization. The random sequence will be sent to the participants directly

by the third party through email 15 mins before the formal trial separately, with corresponding data extraction form and PDF files of related 10 RCTs of the referred group.

Investigators will be blinded in the whole process owing to the aforementioned process (the investigators will never know the random sequence until the third party unblind the sequence). However, it is impossible to blind the participant because when participants receive the data extraction form and the affiliated materials, they would know the intervention type of the RCTs they are about to perform the data extraction. Allocation will be concealed through a unique password-protected data extraction form which will be allocated directly by the third party. In addition, outcome accessors and statistical analysists will be blinded owning to the employment of a third party. To prevent potential exchange on the extracted data among participants within or between groups, the 10 studies in the data extraction form will be ranked randomly for each participant.

Outcomes

The primary outcome will be the error rates on data extraction of the pharmaceutical intervention group and the non-pharmaceutical group, *before* the double-checking process, in terms of the cell level, study level, and participant level. The secondary outcome will be the error rates on data extraction of the pharmaceutical intervention group and the non-pharmaceutical group, *after* the double-checking process, again, in terms of the cell level, study level, and participant level.

In addition, the time-standardized error rates of pharmaceutical over non-pharmaceutical intervention before and after the double-checking process is also of interest. For each individual, the time taken to complete the data extraction in each period is defined as the sum of the time spent on each of the RCTs. This will be recorded through a self-programmed Excel micro by the third party. In brief, when the participants click the 'starting' button, the program will start to record the time; and

when they click the 'stopping' button, the program will stop to record the time. Participants may leave the room for some private reasons that could prolonged the time as long as they forgot to click 'stop'; To avoid such a overestimation, any participants will be asked to inform the investigators if they need a short leaving so that the investigators could record the time in leaving of the participants, and this will be subtracted from the total time.

Follow-up

The follow-up of the trial would be from the beginning of the data extraction to the complementation. No further follow-up is required due to the aim of the trial. As a result, we expect a low dropout rate during the follow-up.

Sample size estimation

We used the following formula [16] to estimate the sample size requirements for each of the two groups for an equivalence test:

$$\text{Sample size} = \frac{[Z_{1-\beta} + Z_{1-\frac{a}{2}}]^2[\pi_\lambda(1-\pi_\lambda) + \pi_o(1-\pi_o)]}{(\pi_\lambda - \pi_0)^2},$$

where Z is the standard score that refers to the number of standard deviations from the mean, a is the significance level, and β is the statistical power that reflects the ability to reject the null hypothesis when there is a true effect. In addition, π_λ is the error rate of data extraction in intervention arm and π_o is the error rate in control arm. Here, the sample size is estimated in terms of both study level and participant level. For the study level, the total sample size means the number of participants (n) multiple the number of studies (k), namely, $n * k$; for the participant level, the total sample size means the number of participants (n).

For both levels, we considered an a=0.05 and β=0.8, while the event rates π differ. For the study level, based on previous two trials [10,11], π_λ is expected to be between 15.41% and 19.90%, while π_0 is about 8.87% to 10.20%. For the participant level, based on our empirical investigation and other studies [6, 7], about

65% of the meta-analyses had data extraction errors — some referred to pharmaceutical interventions, and some referred to non-pharmaceutical interventions, indicating that π_λ would be greater than 65%, while π_0 would be less than 65%. Therefore, we empirically set π_λ = 20% for study level and 80% for participant level, while π_0 = 10%for study level and 50% for participant level. Based on such settings, we obtained a sample size for each group of 12 (118/10) in terms of the study level and 40 (398/10) in terms of the participant level. Under a dropout rate of 10%, we take 45 as the minimal sample size of each group, and thus at least 90 participants are needed in total.

Recruitment

Participants will be recruited on the base of advertising. Investigators will paste the advertisement poster in the main buildings (i.e., teaching building, dining hall) of the three centers to make sure the majority of the staffs and students could reach out the recruiting information. To maximize the visibility of the advertise, investigators and their colleagues will share the e-poster in their own social media (e.g., WeChat) or community group. Subjects who are willing to participant are also encouraged to invite their friends for participation. The recruitment will be started in Aug 2022.

Data collection

Data collection will be done along with the data extraction form mentioned above. The following baseline information will be collected: age, gender, background or expertise, experience on systematic reviews, and experience of publication. The name of participants will not be collected; instead, the student identifiers for each participant will be collected. Any additional information that are considered useful at any period of the trial will be collected through face-to-face interview. Time information on data extraction will be automatically recorded by the macro as aforementioned.

Statistical analyses

Baseline characteristics will be depicted in terms of the data type. For discrete variables (e.g., gender, background, experience on systematic reviews) frequency and proportion will be summarized, and for continuous variables (e.g., age) the mean and standard deviation (SD) or median and interquartile range (IQR) will be presented. Baseline demographic data will be compared using an independent-sample *t*-test or chi-squared test as a verification of the implementation of the randomization process.

For the main analysis, both intention-to-treat and per protocol principles will be used to examine the potential difference of data extraction error rates amongst the above comparisons. Considering that the estimation of participant-level error rate would be impacted by the study level and the cell level estimates, we will establish a generalized linear mixed model by treating each cell as level 1, study as level 2 and participant as level 3 to address this problem. The risk ratio (RR) will be used as an effect estimator under the binomial distribution with a log link faction [17, 18]. While time used for data extraction is expected to impact the error rate, time-standardized rate that measured by incidence risk ratio (IRR) will be estimated under the mixed Poisson model [19].

Subgroup analysis will be employed for the following factors that may impact the quality of data extraction, including the gender, experience of individuals on systematic reviews, time used for the data extraction, and experience on publication of the participants. All the analyses will be conducted using Stata/SE 16.0 (Stata Crop LCC, College Station, TX), with a significance level of 0.05. The statistical analysis will be done by a statistician who would be blinded of the allocation information.

Discussion

In this protocol, we depicted the design, implementation, and analysis plan of a

forthcoming trial which aims to establish informed evidence for qualified data extraction in evidence synthesis practice. To the best of acknowledgement, this will be the first trial that compares the error rates of data extraction in the sight of intervention type that characterized by pharmaceutical intervention or non-pharmaceutical intervention. The study will be the third randomized trial on strategy of data extraction, while the first in the Asian-Pacific area. The study will also provide evidence on the potential benefits of duplicate data extraction on data reproducibility, in reducing errors and trading-off the potential negative impacts of intervention type on errors. Through a cross-over design, the study presents a valid way to eliminate the potential impact of the heterogeneous contexts of the studies and thus is expected to provide less biased evidence to support a better evidence synthesis practice.

While there are some limitations in terms of the design and implementation. First, to ensure the feasibility of the trial, we restrict the participants to 2nd year post-graduate students or above (e.g., doctoral program); this would be no doubt impact the representativeness of the samples. Indeed, clinicians are one of the key contributors of published systematic reviews and meta-analyses, who will be not well-covered in this trial. Second, we will take the Excel sheet for participants to perform the data extraction; Although the Excel sheet will be encrypted protected, there still would be a risk that participants who finish the data extraction will copy the data and send it to other participants — this would 'contaminate' the dataset and introduce bias on the results. Fortunately, some useful strategies could be taken to prevent this case from happening, for example, conduct a pilot trial to evaluate the risk, disrupt the order of the study list in the sheet, remove duplicated data sheet if the information is identified as totally the same by different participants. The random assignment and blinding scheme would also be a valid approach to present the case.

In summary, the conduction of this trial is expected to provide useful evidence to guide the data extraction practice for further systematic reviews authors and new

evidence for methodologist to design a better data extraction strategy.

Declarations

Ethics

This RCT was registered on the Chinese Clinical Trial Registry (Identifier: ChiCTR2200062206). Ethics approval was obtained from the Anhui Medical University IRB.

Dissemination

Results of this protocol will be written up for publication in a high-quality peer-reviewed journal and disseminated at local, national and international meetings. Papers describing the key findings will be submitted within 12 months of the trial completion. Participants will not be identified in any publications or presentations resulting from this study.

Data sharing

The data will be shared with the public after the publication of the trial.

Conflict of interests

None

Funding

This work will be supported by a seed fund for talented earlier researchers from Anhui Medical University (9021783201) and program grant #NPRP-BSRA01-0406-210030 from the Qatar National Research Fund (a member of Qatar Foundation).

Acknowledge

We gratefully acknowledge the assistance provided by Taihe Hospital; Department of pharmacy, Guizhou Provincial People's Hospital; many other individuals on the project.

Reference

1. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. *Lancet* 2017; 390(10092): 415-423.
2. Ioannidis JP. The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses. *Milbank Q* 2016; 94: 485-514.
3. Jones AP, Remmington T, Williamson PR, et al. High prevalence but low impact of data extraction and reporting errors were found in Cochrane systematic reviews. *J Clin Epidemiol* 2005; 58:741-742.
4. Gøtzsche PC, Hróbjartsson A, Marić K, et al. Data extraction errors in meta-analyses that use standardized mean differences. *JAMA* 2007; 298:430-437.
5. Mathes T, Klaßen P, Pieper D. Frequency of data extraction errors and methods to increase data extraction quality: a methodological review. *BMC Med Res Methodol* 2017; 17:152.
6. Xu C, Doi SA, Zhou XQ, et al. Data reproducibility issues and their potential impact on conclusions from evidence syntheses of randomized controlled trials in sleep medicine: A critical evaluation. *Sleep Medicine Reviews* 2022; In revision.
7. Xu C, Yu T, Furuya-Kanamori L, et al. Validity of data extraction in evidence synthesis practice of adverse events: reproducibility study. *BMJ* 2022;377: e069155.
8. Chandler J, Cumpston M, Thomas J, et al. Chapter I: Introduction. In: Higgins JPT, Thomas J, Chandler J, et al (editors). *Cochrane Handbook for Systematic Reviews of Interventions version 6.3*. Cochrane, 2022. Available from www.training.cochrane.org/handbook.
9. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 2017; 358: j4008.
10. Buscemi N, Hartling L, Vandermeer B, et al. Single data extraction generated more errors than double data extraction in systematic reviews. *J Clin Epidemiol* 2006; 59(7): 697-703.
11. Li T, Saldanha IJ, Jap J, et al. A randomized trial provided new evidence on the

accuracy and efficiency of traditional vs. electronically annotated abstraction approaches in systematic reviews. *J Clin Epidemiol* 2019; 115: 77-89.

12. Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Intern Med* 2010; 152:726-732.
13. Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med* 2013; 158:200-207.
14. Rosenberger KJ, Xu C, Lin L. Methodological assessment of systematic reviews and meta-analyses on COVID-19: A meta-epidemiological study. J Eval Clin Pract. 2021; 27(5):1123-1133.
15. Furuya-Kanamori L, Lin L, Kostoulas P, et al. Limits in the search date for rapid reviews of diagnostic test accuracy studies. *Res Syn Meth* 2022; 1-7. doi:10.1002/jrsm.1598.
16. Julious SA, Campbell MJ. Tutorial in biostatistics: sample sizes for parallel group clinical trials with binary data. *Stat Med* 2012; 31(24):2904-36.
17. Austin PC. A Tutorial on Multilevel Survival Analysis: Methods, Models and Applications. *Int Stat Rev* 2017; 85(2):185-203.
18. Rabe-Hesketh S, Skondral A (editors). Multilevel and Longitudinal Modeling Using Stata (4th Edition). *Stata Press* 2022.
19. Bagos PG, Nikolopoulos GK. Mixed-effects Poisson regression models for meta-analysis of follow up studies with constant or varying durations. *Int J Biostat* 2009; 5:21: article 21.

Table 1. Data Extraction form.

Section 1: Basic Information							
Student ID			**Age**			**Gender**	
Background			**Experience in publication**			**Experience in Meta-analysis**	
Section 2: Key Components of PICO							
Population	*Any population*	**Intervention**	*Amoxicillin*	**Comparison**	*Placebo*	**Outcome**	*Diarrhea*
Section 3: Data Extraction form							
Study	**Intervention （Amoxicillin）**			**Comparison (Placebo)**			**Electronic timer**
	Cases	Total	Treatment Duration (*wk*)	Cases	Total	Treatment Duration (*wk*)	
Example	*3*	*8*	*16*	*1*	*8*	*16*	**Click to Start**
Burke et al							
Jørgensen et al							
Meltzer et al							
Merenstein et al							
Nduba et al							
Taylor et al							
Wald et al							
Baecher et al							
Esposito et al							
Trehan et al							
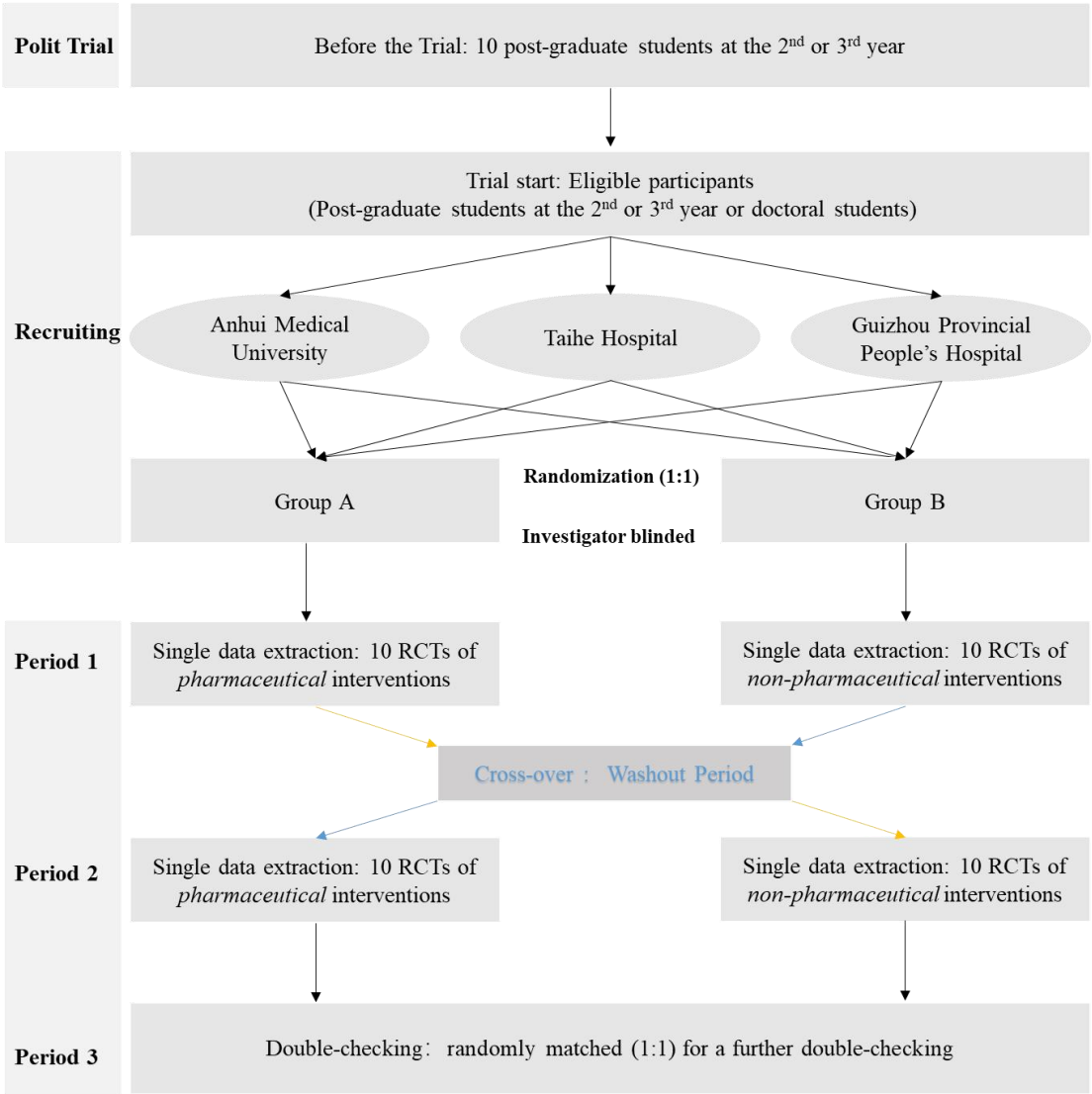
Figure 1. Study design.