The Adjustment Methods for Missing Data in Clinical Trials of Traditional Chinese Medicine

Yang XIE¹,²,* and Jia-jia WANG²

¹Department of Respiratory Diseases, The First Affiliated Hospital of Henan University of Chinese Medicine, Zhengzhou, Henan province, China
²Collaborative Innovation Center for Respiratory Disease Diagnosis and Treatment & Chinese Medicine Development of Henan Province, Henan University of Chinese Medicine, Zhengzhou, Henan Province, China

*Corresponding author

Keywords: Missing data, Adjustment methods, Traditional Chinese medicine, Clinical trials.

Abstract. Missing data occurs widely in clinical trials of traditional Chinese medicine (TCM), which seriously compromises inferences from clinical trials and undermines the precision of the results. There are four different types of adjustment methods for missing data according to the characteristics of TCM clinical trials, which are complete-case analysis, single imputation methods, estimating-equation methods, and methods based on a statistical model. The adjustment methods selected should be clarified in detail, including its rationality and objectivity for efficacy evaluation of TCM clinical trials, and the corresponding statistical analysis method should be also indicated.

Introduction

Missing data are defined as values that are not available and that would be meaningful for analysis if they were observed [1]. Missing data is a serious problem that compromises inferences from clinical trials and undermines the precision of the results, making them unreliable [2]. As for traditional Chinese medicine (TCM), an integral part of the Chinese culture and civilization, has a time-honored history of several thousand years. TCM unique characteristics and marvelous curative effects are increasingly recognized by people worldwide. Nowadays, with its gradual modernization and internationalization, a number of large sample, multi-center randomized controlled clinical trials of TCM are springing up, however, missing data occurs widely in TCM medical research, and indeed is difficult to avoid [3]. Missing data is common in TCM clinical trials, yet the topic has received little attention of researches. Researches on adjustment methods for missing data in TCM clinical trials are quite deficient. More importantly, the relevant researches that are in accordance with characteristics of TCM itself have not been found yet. So far, adjustment methods are mainly limited to complete-case analysis and simple imputation methods. These methods are relatively robust when the ratio of missing data is low. However, with the ratio ascending, the robustness of these methods decreases gradually, resulting in bias and even error [4]. Thus, how to handle the missing data in TCM clinical trials is a most urgent problem that must be addressed.

The Cause of Missing Data in TCM Clinical Trials

Although the concepts of key elements in TCM clinical trials, such as participants, intervention, outcome indicators, treatment time, methods of data collection and analysis as well as quality control, maybe the same as that in modern medicine, yet its connotation and denotation have a fundamental difference [5]. Missing data is more likely according to the characteristics of TCM clinical trials.

As for research ideas, current research mode of combination of disease and TCM syndrome dominates in TCM clinical trials, in which we grasp the occurrence and development rules of diseases dynamically by TCM syndrome differentiation on the basis of western medicine diseases
Therefore, the adjustment of the TCM treatment plan/method may increases the interview point and challenges the participants' compliance potentially, resulting in more missing data.

As for research design, clinical observation and randomized controlled trials, including large sample, multi-center randomized controlled trials, are given priority in TCM clinical trials. Generally speaking, TCM intervention is mainly aimed at diseases in stable stage. In view of the main characteristics of TCM- holism concept and syndrome differentiation and treatment, the research cycle is relatively long, which is likely to lead to more dropout, along with corresponding missing data.

As for efficacy evaluation, the standard of efficacy evaluation of modern medicine is mainly adopted in TCM clinical trials. Other than “objective indicators” of TCM collected by inspection, auscultation and olfaction, inquiry, and pulse-taking, patients’ diet and daily life, social background, psychological state, et cetera also play an important role in efficacy evaluation of TCM clinical trials. However, a scientific efficacy evaluation system in conformity with TCM’ own characteristics hasn’t been established yet, so part of the data collected can't be in statistical analysis, thus causing relative missing data.

The Adjustment Methods for Missing Data in TCM Clinical Trials

The Mechanism of Missing Data in TCM Clinical Trials

Different adjustment approaches depend on mechanism leading to missing data in clinical trials. Currently, there are three acceptable patterns of missing data: (1) Missing completely at random (MCAR), in which the probability of a particular value being missing is completely independent of both the observed data and the unobserved data. (2) Missing at random (MAR), in which the probability of a particular value being missing depends only on the observed data. (3) Missing not at random (MNAR), in which the probability of a particular value being missing depends on the unobserved data.

The Adjustment Methods for Missing Data in TCM Clinical Trials

Based on the three patterns of missing data above, there are four different types of adjustment methods: complete-case analysis, single imputation methods, estimating-equation methods, and methods based on a statistical model. We should choose corresponding processing method in terms of specific mechanism. In complete-case analysis, participants with missing data are simply excluded from the analysis. In simple imputation methods, a single value is filled in for each missing value by means of methods such as the last observation carried forward and the baseline observation carried forward. In estimating-equation methods, complete cases are weighted by the inverse of an estimate of the probability of being observed. For example, the probability of an outcome being observed might be modeled with the use of baseline data, and then the complete cases might be weighted by the inverse of their estimated probabilities of being observed.

Methods that are based on a statistical model include maximum likelihood, in which estimates and standard errors are based on the likelihood function given the observed data; bayesian methods, in which inferences are based on a statistical model that includes an assumed prior distribution for the measurements; and multiple imputation, in which multiple sets of plausible values for missing data are created from their model-based predictive distribution, and estimates and standard errors are obtained with the use of multiple-imputation combining rules.

Sensitivity analysis can assess the influence of different adjustment methods for missing data on inferences from clinical trials, which contributes to validating the robustness of the selected methods of handling missing data. Therefore, sensitivity analysis should be advocated in TCM clinical trials to assess the robustness of different adjustment methods.

Discussion

Generally, it is considerably necessary to preselect adjustment methods alternative for missing data in the design of TCM clinical trials in order to avoid to reconsider the selection of adjustment methods.
methods in statistical analysis. Notably, the adjustment methods selected should be clarified in detail, including its rationality and objectivity for efficacy evaluation of TCM clinical trials, and the corresponding statistical analysis method should be also indicated.

Missing data in TCM clinical trials are affected by many factors, such as the length of research cycle, the number of interview point, the way of treatment, the indicators of efficacy evaluation, etc. Measures should be taken correspondingly to reduce the missing data. For example, as to long research cycle and numerous interview points, health education lectures and appropriate free inspections can be adopted in some interview point; For efficacy evaluation of TCM clinical trials, relevant standards should be formulated and improved, and more clinical endpoint indicators designed as well, such as case fatality rate, recurrence rate, complication rate, survival time and quality of life assessment [7]. Regardless of subjects' compliance, missing data can be reduced through optimizing study design, strengthening data management, encouraging researchers to continue to collect data after suspension of the research, especially those related to clinical endpoints [8].

The adjustment methods for missing data are based on certain assumption of missing data mechanism, so the recognition of missing data mechanism of TCM clinical trials is the precondition of selecting optimum adjustment methods. However, part of missing data mechanism of TCM clinical trials is difficult to explicit in that there are few detailed descriptions of dropout or efficacy evaluation indicators in conformity with own characteristics in TCM clinical trials, making it difficult to select corresponding adjustment methods. In TCM clinical trials, therefore, on the one hand, steps should be taken to try to avoid missing data, and the situation of dropout also recorded in detail; on the other hand, adjustment methods for missing data suitable for the characteristics of TCM itself should be pressingly probed based on the data of TCM clinical trials.

Acknowledgments

The current work was partially supported by the Special Science Research Program of Chinese medicine of Henan province (No.2015ZY02043), Key Science and Technology Program of Henan province (No. 172102310200), and the National Natural Science Foundation Project of China (No. 81674098).

References