Received: 6 September 2021 | Revised: 24 December 2021 | Accepted: 13 January 2022

DOI: 10.1111/jebm.12462

METHODOLOGY

WILEY

A methodological framework for tackling confounding by indication when assessing the treatment effects of Chinese herbal injections in the real world

Jing Tan[1,2,3] | Chunrong Liu[1,2,3] | Mingxi Li[1,2,3] | Hongcai Shang[4] | Wen Wang[1,2,3] | Ling Li[1,2,3] | Yiquan Xiong[1,2,3] | Shiyao Huang[1,5] | Chaolong Rao[6] | Xiaochao Luo[7] | Yana Qi[1,2,3] | Jing Wang[1,2,3] | Kang Zou[1,2,3] | Xin Sun[1,2,3]

[1] Chinese Evidence-Based Medicine Center, West China Hospital, Sichuan University, Chengdu, China

[2] NMPA Key Laboratory for Real World Data Research and Evaluation in Hainan, Chengdu, China

[3] Sichuan Center of Technology Innovation for Real World Data, Chengdu, China

[4] Key Laboratory of Chinese Internal Medicine of MOE and Beijing, Beijing University of Chinese Medicine and Pharmacology, Beijing, China

[5] Sichuan Evidence-Based Medicine Center of Traditional Chinese Medicine, Hospital of Chengdu, University of Traditional Chinese Medicine, Chengdu, China

[6] School of Public Health, Chengdu University of Traditional Chinese Medicine, Chengdu, China

[7] Acupuncture and Tuina School, Chengdu University of Traditional Chinese Medicine, Chengdu, China

Correspondence
Xin Sun, Chinese Evidence-Based Medicine Center, West China Hospital, Sichuan University, No. 37, Guoxue Lane, Wuhou District, Chengdu 610041, China.
Email: sunxin@wchscu.cn

Funding information
National Key R&D Program of China, Grant/Award Numbers: 2017YFC1700400, 2017YFC1700406; CHINA MEDICAL BOARD, Grant/Award Number: CMB19-324; National Natural Science Foundation of China, Grant/Award Numbers: 72174132, 71974138; 1·3·5 Project for Disciplines of Excellence, West China Hospital, Sichuan University, Grant/Award Number: ZYYC08003; Sichuan Youth Science and Technology Innovation Research Team, Grant/Award Number: 2020JDTD0015

Abstract

Aim: In the context of integrative medicine, whether Chinese herbal injections are effective in routine practice has become a question of broad interest. However, confounding by indication (i.e., indication bias) is a prevalent and highly challenging methodological issue when using routinely collected health care data to assess the real-world effectiveness of Chinese herbal injections.

Methods and results: We proposed a methodological approach to tackling confounding by indication in assessing the real-world effectiveness of Chinese herbal injections, incorporating empirical experiences, a literature review and interactive discussions, and a panel of external experts to finally achieve a consensus. This approach consisted of three cohesive steps, including a full understanding of treatment patterns, construction of fair comparisons by identifying appropriate combination treatments and comparators, and using statistical methods to further control for confounding. In the investigation of treatment patterns, we proposed five domains to identify treatment patterns with Chinese herbal injections, and we offered five patterns of combination treatments to characterize how Chinese herbal injections are used in conjunction with other treatments. In constructing fair comparisons, we suggested the use of both nonuse and active comparators; given the diverse combination treatments, we developed six scenarios that may form fair comparisons. In the statistical analysis, we discussed five

statistical models for controlling confounding by indication, including their pros and cons. We also included a practical example to illustrate the usefulness of the methodological approach.

Conclusion: The proposed approach may serve as an effective tool to guide researchers to reliably assess the effectiveness of Chinese herbal injections in the context of integrative medicine.

KEYWORDS
Chinese herbal injection, indication bias, real-world setting, treatment effects

1 | INTRODUCTION

Chinese herbal injections (CHIs), widely used to treat various diseases,[1] are usually administered through multiple routes and often given in combination with pharmaceutical drugs to heterogeneous patient populations.[2] As such, their safety and effectiveness in real-world practice have become important issues that warrant substantial efforts to both improve the practice of traditional Chinese medicine (TCM) and ensure patient safety.[2,3] However, there is a paucity of high-quality evidence about the treatment effects of CHIs in real-world practice.[4]

In recent years, routinely collected health care data (RCD), such as electronic medical records (EMRs), have become an irreplaceable source of information for assessing the real-world treatment effects of CHIs, especially in exploring the optimal treatment patterns and the timing of CHIs among heterogeneously treated populations.[5–7] Although RCD share apparent advantages in sample sizes, representative population, and high-dimensional variables, these data are, in their nature, observational.[8] Consequently, the resulting treatment effects of CHIs are usually threatened by a diversity of potential biases. One particular concern is confounding by indication (i.e., indication bias).[9]

Confounding by indication is a highly prevalent problem in the assessment of treatment effects of Chinese herbal medicines. This is inherent with the health care setting in the context of integrative medicine. There are two major issues that warrant careful considerations in addressing confounding by indication when assessing the treatment effects of CHIs. First, the use of CHIs is usually subject to patient conditions. Those patients with more severe conditions are more likely to receive CHIs;[10] however, such patients are destined to experience poorer health outcomes.[11] In such cases, a comparison between the use and nonuse of a CHI would be unfair given the clear imbalance in the prognostic prognosis. As a result, the apparent effect estimates are largely distorted due to the preferential use of CHIs among patients with more severe conditions.

The other issue is that CHIs are typically used in conjunction with other treatments, which may further complicate the situation. In routine clinical practice, CHIs are often used in combination with pharmaceutical drugs,[12,13] and, in most cases, as a kind of adjunctive therapy to pharmaceutical drugs for inpatients in Chinese hospitals.[14] As a result, the treatment patterns of CHIs are usually complex in real-world practice. For example, one study found that >80% of patients received combination therapy, including multiple use of CHIs and the combination of CHIs with pharmaceutical drugs, in real-world settings.[15] However, due to the significant limitations of RCD, details regarding the use of CHI and other treatments—for instance, the sequential order of use—are often unavailable, and the lack of this information can make the choice of comparison more challenging.

Confounding by indication has become a serious problem that jeopardizes the scientific assessment of CHIs, and consensus has been reached that substantial efforts are warranted to advance methodologies in assessing the treatment effects of CHIs.[16] In response to this imperative, we developed a methodological approach to tackling confounding by indication in assessing the real-world effects of CHIs.

2 | METHODS

The development of the methodological framework occurred in two cohesive steps. First, we developed an initial framework by considering key methodological components involved with the framework, based on empirical experience, existing literature,[8,10,11,17–20] and several interactive discussions within the research group. These included considerations of the special context regarding the use of CHIs, key methodological issues about confounding by indications during the use of CHIs, and potential approaches to controlling confounding by indications at both the epidemiological design and statistical analysis stages.

Then, we invited a group of external experts in clinical epidemiology, biostatistics, and TCM practice in TCM hospitals and general hospitals to independently review the existing literature and the drafted document about the methodological framework. They were asked if our considerations of the special context were complete and appropriate, if the proposed epidemiological and statistical methods were appropriate, if there were any additional methods that should be included, and if there were any leftover issues that should be included in the framework. After their reviews were complete, they offered feedback to the research group to drive updates to the proposed framework.

Finally, a panel of the external experts, together with members of our research group, was convened to undertake a consensus process to finalize the framework. During the process, they were asked to thoroughly review the updated framework; then, for each component of the

framework, the panel members were asked if there were any disagreements on the proposed methodological approaches. All discrepancies were then discussed among the panel experts, and iterative discussions were conducted until a consensus was achieved in any case of unresolved disagreement.

3 | RESULTS

3.1 | Tackling confounding by indication for CHIs: a general framework

Confounding by indication is an epidemiological phenomenon that may be essentially explained by differential prescriptions made by clinicians due to their judgment of patient prognoses. An ideal approach to eliminating confounding by indication is allocating treatments by randomization;[11] unfortunately, this is unlikely to occur when using RCD for assessing treatment effects.

To minimize this confounding, developing a conceptual framework of fair comparison is the key. A fair comparison in real-world practice means that patients in the comparison groups selected from multiple treatment modes are comparable, and the resulting difference in the effect estimates may be causally inferred from the use of CHIs. There are two assumptions for constructing a fair comparison in the real-world practice setting; one is that clinicians' prescriptions with CHIs are not dependent on patient health conditions, and the other is that patient prognosis is balanced between patient groups having different treatment modes. However, in the real-world practice setting, both of these assumptions are difficult to meet. One major issue is that researchers will not be able to parcel out, in the observational study setting, specific reasons for prescribing treatments based only on available data. The other challenge is the presence of diverse treatment modes (i.e., different combinations of CHIs with pharmaceutical drugs), which often makes the choice of treatment modes for fair comparison tricky. Things become even more complicated if researchers are unlikely to clearly identify a sequential order of the use of the treatments, which is particularly the case in an emergent setting.

To achieve a fair comparison, we have proposed a stepwise approach (Figure 1). Developing a clear understanding about how CHIs are used is the initial—but very important—step. This includes gathering an in-depth understanding of how patients are managed, when and how CHIs are used, and on what basis CHIs are prescribed. The in-depth investigation of these issues would not only facilitate the identification of treatment modes but also enhance the understanding of the potential sources of confounding by indication. Subsequently, careful selection of treatment modes to construct fair comparisons represents a critical step. Although various treatment modes may be present in real-world practice, potentially producing a number of comparisons, only a selected number of comparisons may be deemed comparable. In such cases, careful selection of the study population is the key consideration. Finally, advanced statistical methods are involved that further address the confounding by indication.

3.2 | Investigating patterns of combination use of CHIs

3.2.1 | Key components in the investigation of treatment patterns

In the careful investigation of patterns regarding the combination use of CHIs, five steps are often involved, which are as follows.

1. Specify the target disease condition for which CHIs are used. Investigators should also carefully consider the extent to which potential patients with varying comorbidities are included as this information can affect the applicability of findings. Patients with contradictions should always be excluded.
2. Understand details about the target CHI, including single dose, cumulative dose, and administration route(s) and duration. If relevant, the choice of solvents and the drip rate may also be documented.
3. Understand how other treatments are used, such as their types (e.g., pharmaceutical agents or other TCM products), doses, frequencies, and pharmacological effects. Those treatments that have pharmacological or pathophysiological effects similar to the target CHI should be clearly identified.
4. Identify the hierarchy of use in routine practice to understand the sequential order of use—that is, determine whether the target CHI is used concomitantly with other treatments or in a sequential manner. In investigating the hierarchy, gathering information regarding the timing of use is important. Notably, records about the timing of prescriptions in RCD may not indicate the timing of use, especially when multiple prescriptions were given during a single day or in an emergency or surgical setting.
5. Where needed, conduct discussions with clinicians to learn about the reasons for the prescribing behavior.

3.2.2 | Summarizing patterns of combination use

When the details about the use of target CHIs and other treatments are readily available, the subsequent step—which is critical—is to summarize the patterns of combination use. Though highly complex in the use of CHIs together with other treatments, one may conceptualize several representative scenarios of combination uses.

In the following, we denote capital letter E as the CHI of interest, and capital letter C as the other treatments. The other treatments may either be pharmaceutical drugs, CHIs other than the target one, or both, and may fall into the following five categories (i.e., C1–C5).

- E: The CHI of interest
- C1: Other treatments used in the first line
- C2: Other treatments that have pathophysiological effects similar to those of the target CHI

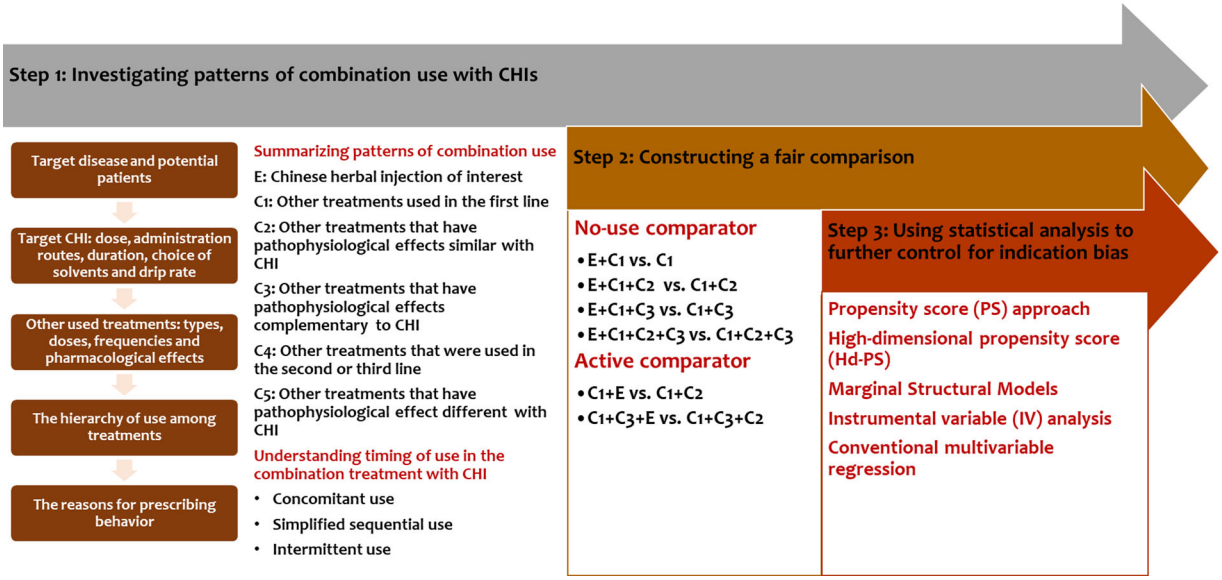

FIGURE 1 A stepwise approach for tackling confounding by indication for Chinese Herbal injections (CHIs)

- C3: Other treatments that have pathophysiological effects complementary to the target CHI
- C4: Other treatments that used in the second or third line
- C5: Other treatments that have pathophysiological effects different from those of the target CHI

Note that a treatment may fall into multiple categories, and the currently listed categories may not be inclusive of all potential scenarios of other treatments.

It is worth noting that the pattern of combination use with CHI may go beyond what we discuss here. Nevertheless, an in-depth understanding of clinicians' prescription habits and the pharmacological or pathophysiological effects of different treatments is critical. At this stage, one should be inclusive in considering all possible patterns of combinations. The key issue is to parcel out the roles of all these different treatments.

3.2.3 | Understanding the timing of use in combination treatment with CHI

Upon gaining a clear understanding of the patterns of combination use with the target CHI, the subsequent step is to identify the pattern of combination by the timing of treatment administration. Generally, there may be three scenarios.

The first scenario is concomitant use, where the target CHI and other treatments are used at the same time. In this context, the target CHI may be considered complementary in pharmacological or pathophysiological mechanisms to other treatments (e.g., pharmaceutical drugs) such that the combination may enhance the effects. The second scenario is simplified sequential use. In this case, the target CHI is given subsequent to the administration of first-line treatment (e.g., pharmaceutical drug), and no additional treatments are involved. Both of these scenarios are relatively straightforward, and the other treatments (e.g., pharmaceutical drug) may be deemed as the baseline treatment regimen or as confounders according to the classic criteria.[16,17]

The third scenario is intermittent use of CHI—that is, complex sequential use. In such a case, the target CHI may be used intermittently between two or more pharmaceutical drugs. One example is that an additional pharmaceutical drug is added if the CHI fails to achieve the expected effect. The major challenge in this context stems from the unavailability of information regarding the timing of administration, in which case one would be unable to distinguish the order of administration across the treatments. This issue unfortunately prevails in studies using RCD. If the sequential order is ignored, the unique effect of the target CHI would likely be parceled out of the complex treatment ordering. Therefore, this scenario would usually be ineligible for forming a fair comparison. However, in some cases, one would be able to identify sequential orders of multiple treatments from RCD. For instance, in exploring patterns of combination use, we found that a CHI was often used concomitantly with the first-line treatment in the first few days after diagnosis. Subsequently, other pharmaceutical drugs or alternative CHIs were added to the CHI. After that, as a result of the unsatisfied effect of current treatment regimes, CHI or add-on treatments may be replaced by other treatments. In this circumstance, time-varying exposure of CHI would be formed, in which case sophisticated statistical analyses, such as marginal structure modeling, may address this issue by involving time-varying variables. For the sake of clarity and readability, we focused on cases where the sequential order of treatments is unclear in this paper. We will discuss, in another paper, scenarios where time-varying information is readily available for assessing the effects of CHIs.

TABLE 1 Suggested comparisons for assessing treatment effects of CHIs in the context of combination use

Comparator	Scenario	Exposure	Control	Primary population	Potential exclusion or restriction	Covariates by treatments
Nonuse comparator	1	E+C1	C1	Patients receiving C1	Patients receiving C4	C2, C3, and C5, if available
	2	E+C1+C2	C1+C2	Patients receiving C1 and C2	Patients receiving C4	C3 and C5, if available
	3	E+C1+C3	C1+C3	Patients receiving C1 and C3	Patients receiving C4	C2 and C5, if available
	4	E+C1+C2+C3	C1+C2+C3	Patients receiving C1, C2 and C3	Patients receiving C4	C5, if available
Active comparator	5	C1+E	C1+C2	Patients receiving C1	Patients receiving C4	C3 and C5, if available
	6	C1+C3+E	C1+C3+C2	Patients receiving C1 and C3	Patients receiving C4	C5, if available

3.3 | Constructing fair comparisons

Following the thorough investigation of the patterns of combination use with CHI and full considerations about the timing of use, one would be ready to construct a fair comparison. The construction of fair comparison involves two types of controls, specifically nonuse and active controls. For these controls, we developed a total of six scenarios that are potentially effective in reducing the impact of confounding by indication (Table 1).

In the comparisons that involve nonuse controls, four types of comparison may be feasible. Essentially, these comparisons are constructed in a manner where users of the target CHI are compared to nonusers, while other treatments are commonly administered to both the exposure and comparison groups. In the use of an active comparator, the comparisons are mainly made between those treatments that have similar pathophysiological effects and the target CHI (i.e., C2), as this is deemed the most appropriate active comparator.

Across all these comparisons, those populations with treatments used in the second or third line are potentially considered for exclusion or as a limiting factor. These restrictions by pattern of combinations would be effective to reduce confounding by indication[17] because the information regarding the timing of use is unavailable; however, one should remain aware of the limitations by excluding C4. While confounding by indication may be reduced by excluding patients who received second-/third-line treatments, the applicability of findings may be limited because this restriction would be more likely to include patients with milder disease or better prognoses.

In these different comparisons, residual treatments—that is, treatments not included in the comparison of interest but having the potential to affect the treatment outcomes—may be used among patients. In such cases, they can be dealt by including them as covariates in the adjusted analyses.

3.4 | Using statistical analysis to further control for confounding by indication

The development of fair comparisons enables effective control of confounding by indication. Nevertheless, residual confounding by indication continues to be present. In most cases, statistical analyses would be used on top of the fair comparisons.

Several strategies are commonly used for adjusting confounding by indication. Conventional multivariable regression models and propensity score approaches have been popular.[21] In line with the development of EMR databases, however, more complex approaches have been made available to further control for confounding by indication (e.g., high-dimensional propensity score,[22,23] marginal structural models,[24] and instrumental variable (IV) analysis[25]) due to the improved linkage across databases, a wider coverage of variables, and availability of repeated measurements. However, these methods differ in their own pros and cons and have varying applicable conditions. Table 2 summarizes the potential approaches in statistical analysis.

Researchers should choose the appropriate statistical analyses based on the characteristics of data and the pattern of combination use. One should also be aware that traditional confounding-adjustment techniques, such as matching and multivariable regression models, may be inadequate to control for time-varying confounding in a real-world setting. Instrumental variables are typically used to handle unmeasured confounding, and high-dimensional propensity scores are primarily applied in large-scale databases. Details regarding these statistical models may be found elsewhere.[18,19,26]

3.5 | Special considerations

3.5.1 | Issues about the potential interaction between CHIs and Western medicines

In constructing fair comparisons, one would note that the target CHI is often used in combination with Western medicines. This raises an important question about the potential pharmacological interactions. In the proposed analytical approach, there is an assumption that the target CHI has no interaction with other medications, in which case the resulting effect estimates would be derived from the use of the target CHI. However, in some cases, this assumption may not hold. One simplified approach to testing for the postulated interaction is the comparison of resulting effect estimates for consistency. In the case of consistent estimates from different comparisons, the potential drug–drug

TABLE 2 Statistical analyses that control for confounding by indication

Methods	Advantages	Limitations	Applicable conditions
Conventional multivariable regression	• Control for all known confounders to produce a risk-adjusted treatment effect	• Inability to control for unmeasured confounders • May be over-fitted if number of events is few	No specific condition
Propensity score (PS) approach	• Balance measured confounders between comparison groups • Appropriate when number of events is fewer	• Inability to control for unmeasured confounders	Treatment measure is a dichotomous variable
High-dimensional propensity score (Hd-PS)	• Reduce both confounding by indication and the effect of unmeasured confounders. • Automatically selecting a large number of variables for calculating PS[23]	• Not all unmeasured confounders can be controlled	A vast number of variables in real-world healthcare databases
Marginal Structural Models	• Estimate the effect of exposure in the presence of time-varying confounder	• Inability to use when all subjects with that level of the covariate are certain to receive the identical treatment[24]	Used in longitudinal study designs with time-varying information regarding exposure, outcome or other covariates[24]
Instrumental variable (IV) analysis	• Estimate the effect of exposure in the presence of unmeasured confounders	Difficult to choose the IV variable, with three strong assumptions: 1) IV affects exposure, 2) IV affects the outcome only through exposure, and 3) IV does not share a common cause with outcome[33]	No specific condition, if an IV available

TABLE 3 Hypothesized scenarios for testing drug–drug interactions

Comparator	Scenario	Exposure	Control	Effect estimate
	0	C 1	No use	ES0
Nonuse comparator	1	E+C1	C1	ES1
	2	E+C1+C2	C1+C2	ES2
	3	E+C1+C3	C1+C3	ES3

interactions may not play an important role. However, this approach is sometimes confounded by the patient baseline risk given that treatment patterns may differ substantially. For example, Table 3 presents several scenarios of combination use of a target CHI with other medications. In the analysis of treatment effects of the target CHI by scenario, different effect estimates (e.g., ES1, ES2, and ES3) can be generated. In addition, an additional "placebo" comparison would be required, in which no target CHI is involved (thus, the effect estimate is ES0). By testing for pharmacological (drug–drug) interactions, one would check whether these estimates are statistically consistent.

There is, however, always a potential for pharmacological interaction—either synergistic or antagonistic—in the presence of multiple drug treatments. In the case of synergistic interaction, no additional analyses are necessary. However, in the case of an antagonistic interaction, one would have to determine whether the interaction is quantitative or qualitative in nature. In the case of a quantitative antagonistic interaction, the addition of the target CHI to Western medicine would achieve an improved—but smaller than expected—effect. In the qualitative interaction, however, the addition of a CHI would compromise the treatment effect.

3.5.2 | Issues about treatment by TCM syndrome differentiation

One special issue that warrants careful examination of the treatment effects of CHIs is the presence of differential practice patterns between Western medicine and TCM. Though CHIs are widely used both in these contexts, their practice patterns differ. Treatment by syndrome differentiation is often involved when administering CHIs in the TCM setting, but this practice is usually impracticable in the context of Western medicine. It is not impossible—and even likely—that patients with varying syndromes may present with differential treatment effects due to the nature of the interaction between the syndrome and treatment. In all possible cases, the presence of treatment by syndrome differentiation warrants careful analysis.

In the design stage, one may have a strong hypothesis about the potential interactions between CHI treatments and syndromes based on a strong biological rationale or earlier research evidence. In such cases, there are two important approaches to handling treatment by syndrome differentiation. The first is by applying restrictions to patients with a specific syndrome, for which researchers have had a strong belief that patients may benefit. This selective approach may help tailor the patient population to those who benefit most from treatments. Nevertheless, in most cases, earlier biological and epidemiological evidence may not be strong enough to warrant prespecification of a subset of the population. In particular, standardized approaches to diagnosing syndromes are often lacking, and the resulting diagnosis may vary between practitioners. Alternatively, one may not restrict the eligibility criteria to a specific subset population but instead include a broader patient population while prespecifying whether syndrome differentiation would have an interaction effect on treatment. The test for putative interactions may be operated in the statistical analysis that includes an interaction term between treatment and syndrome variables.

4 | WORKING EXAMPLE

4.1 | Background

Motherwort injection is a CHI derived from *Leonurus japonicus* Houtt. The main chemical components of motherwort injection are hydrastine, cucurbitacin, and choline.[27] Pharmacological studies have shown that motherwort injection has effects on uterine contraction and hemostasis, and it is often used for preventing or treating postpartum hemorrhage (PPH).[28] It can stimulate smooth muscle of the uterus both in vivo and in vitro and induce lasting contraction of the whole uterus.[29]

In our efforts to examine the effects (i.e., effectiveness) of motherwort injection for preventing PPH at labor in real-world clinical practice, we used the EMR database of the West China Second University Hospital of Sichuan University. The outcome of our interest was the incidence of PPH, which is a dichotomous variable defined as a bold loss of ≥500 ml.[30] We also investigated the amount of bleeding as the secondary outcome.

4.2 | Investigating patterns of combination use with motherwort injection

In our study, motherwort injection was the CHI of our interest (E), which was administered intramuscularly in the uterus during the third stage of labor for patients undergoing cesarean delivery. Oxytocin is universally recommended as the first-line uterotonic for preventing PPH (C1)[31] and presents a pathophysiological effect for uterine contraction similar to that of motherwort injection (C2). Meanwhile, oxytocin and motherwort injection may be complementary in their effects, as oxytocin has a very short duration of action (half-life with a few minutes, C3)[32] and motherwort injection may sustain the uterine-contraction effect for a relatively long period of time.[28] In addition, oxytocin may be administered through different routes. An intravenous drip is routinely used according to the guideline recommendation; however, some women may receive intramuscular injections in the uterus.

Other uterotonics, such as carbetocin, ergots, prostaglandin F_{2a} (carboprost), and prostaglandin E_1 (misoprostol), have pathophysiological effects similar to those of motherwort injections but have been primarily reserved for the prevention of PPH in high-risk groups. These treatments are often recommended as the second or third treatment options in guidelines (C4). In addition, these uterotonics may be used as an add-on when the use of motherwort injection does not have the effect expected by the practicing clinician (C4). Tranexamic acid, as a hemostatic drug, is usually used when uterotonics have failed to stop bleeding or in the case of traumatic bleeding (C4). Clearly, antibiotics have different pathophysiological effects from those of motherwort injection, thus being classified as C5. In addition, information regarding the timing of use of motherwort injection and other treatments was not available from the EMR database of West China Second University Hospital. This situation presented us with a real challenge in identifying the sequential order when using motherwort injection and second- or third-line treatments (i.e., the C4 treatments in our definition, including carbetocin, ergots, carboprost, misoprostol, and tranexamic acid).

4.3 | Constructing a fair comparison aiming to assess motherwort injection

Given the identified patterns of combination use with motherwort injection, we constructed three comparisons, including two that involved a nonuse comparator and one that included the active comparator. Oxytocin by intravenous administration was the baseline treatment in comparisons 1 and 3, and the combination of intravenous oxytocin with intramuscular oxytocin was the baseline treatment in comparison 2. Across these comparisons, we contrasted the difference between excluding the therapeutic combination with treatments failing into C4 (i.e., second- or third-line treatments) and including them so as to minimize confounding by indication in the context of comparisons. However, sensitivity analysis without excluding population received C4 could be considered.

Given the similar pathophysiological effects and consistency of administration route between motherwort injection and oxytocin, we chose comparison 3 to investigate the effectiveness of motherwort injection (Table 4).

Statistical analysis

To further control for confounding arising from the differences in population characteristics between the exposure and control groups, we applied a 1:1 propensity score–matching approach to adjust for potential confounders, including demographic characteristics, comorbidities,

TABLE 4 Comparisons developed for the assessment of treatment effects

Comparison	Exposure	Control	Covaries
Comparison 1(nonuse comparator)	Oxytocin iv. (C1) + Motherwort injection im. (E)	Oxytocin iv. (C1)	Oxytocin im. (C2, C3) and antibiotics (C5)
Comparison 2(nonuse comparator)	Oxytocin iv. (C1) + Oxytocin im. (C2, C3) + Motherwort injection im.	Oxytocin iv. (C1) + Oxytocin im. (C2, C3)	Antibiotics (C5)
Comparison 3(active comparator)	Oxytocin iv. (C1) + Motherwort injection im. (E)	Oxytocin iv. (C1) + Oxytocin im. (C2)	Antibiotics (C5)

Inclusion criteria

- Pregnant women registered at the first trimester, and underwent cesarean section between January 1, 2015 and November 30, 2019 at the West China Second University Hospital.
- Women received the oxytocin of intravenous drip at the third stage of labor.

Potential exclusion criteria

Women receiving one of following uterotonics and hemostatic drug, including carbetocin, ergots, prostaglandins F2a-carboprost, and prostaglandins E1-misoprostol and tranexamic acid.

Women receiving both Motherwort injection im. and Oxytocin im., or neither, could be excluded in active comparator.

surgical procedures, and other treatments. We then used a logistic regression model to estimate the relative risk of motherwort injection. A series of sensitivity analyses were also conducted to validate the robustness of the results. To further utilize the advantages of a large number of variables in RCD, we also used high-dimensional propensity scoring to maximize the adjustment for the confounding by the indication and unmeasured confounders.

5 | DISCUSSION

In this article, we have proposed a methodological framework that assesses the treatment effects of CHIs in the context of integrative medicine. This framework has specifically discussed several important components that would mandate the analysis and assessment of CHIs, including the investigation of treatment patterns, construction of fair comparisons, and the use of statistical analyses for controlling confounding. In each of the key components, we specifically discussed the implication of CHIs in the assessment of treatment effects. We also specifically discussed issues that are highly relevant for the CHIs, including the test for pharmacological interactions in the presence of multiple drug–drug interactions and the analysis of treatment effects with treatment decided by syndrome differentiation. We believe that the proposed methodological framework is helpful for appropriately and accurately assessing the treatment effects of CHIs in the complex setting of integrative medicine.

The proposed approach has some limitations that warrant attention. First, while this proposed framework was built upon earlier research experiences, an extensive literature review, and expert panel consensus, it has not been validated externally by different CHIs, study settings, or patient populations. However, we are planning more studies to externally address this issue. Second, our considerations of integrative medicine may be simplified. In reality, the clinical setting and practice patterns may be more complex, which would require more sophisticated statistical and epidemiological approaches to address the issue. Certainly, the most logical approach proposed continues to be based on observational study design, where bias is inherent due to the failure to control for unobserved confounding. Therefore, future studies that test for consistency between results from our framework and randomized controlled trials would be ideal to further validate the present methodological framework. Lastly, our belief that the proposed most logical approach would be able to address the question is largely based on the assumption that the data for the analysis are good enough. However, the reward data for TCM are sometimes suboptimal, in which the accuracy, completeness, and relevance of data are concerning. Substantial efforts are warranted to further improve the quality of data for supporting this maneuver.

In the context of integrative medicine, CHIs are usually used with other treatments (e.g., pharmaceutical drugs), and the patterns of combination use—not only in the number of treatments combined but also the timing of use—are often highly complex. This high level of complexity has rendered the assessment of the effects of CHI often challenging in the real-world setting. One most significant problem is the presence of confounding by indication, which is unfortunately prevalent in integrative medicine practice. To address this important methodological issue, we have proposed a stepwise methodological approach to tackling confounding by indication. While constructing fair comparisons is the core of our approach, an in-depth and clear understanding of the patterns of combination use represents the most important step that one should never ignore. Use of statistical methods may further control for the confounding by indication; however, it largely relies on successful building up of the fair comparisons. We are hopeful that our CHI-specific methodological approach would enhance the valid assessment of treatment effects in the context of integrative medicine. In addition, we would like to use multi-source real-world data to validate the rationality of this approach in the future.

ORCID

Mingxi Li https://orcid.org/0000-0002-7258-8849

Xin Sun https://orcid.org/0000-0002-6554-7088

REFERENCES

1. Li H, Wang S, Yue Z, Ren X, Xia J. Traditional Chinese herbal injection: current status and future perspectives. *Fitoterapia*. 2018;129:249-256.
2. Zhang L, Yan J, Liu X, Ye Z, Yang X, Meyboom R, et al. Pharmacovigilance practice and risk control of Traditional Chinese Medicine drugs in China: current status and future perspective. *J Ethnopharmacol*. 2012;140(3), 519-525.
3. Liang Z, Lai Y, Li M, Shi J, Lei CI, Hu H, et al. Applying regulatory science in traditional chinese medicines for improving public safety and facilitating innovation in China: a scoping review and regulatory implications. *Chin Med*. 2021;16(1), 23.
4. Xie Y, Liang X, Zhang G. Survey and analysis of 30 instructions for Chinese herbal injections. *China Pharmaceut*. 2011;20(24), 10-12.
5. Tan J, Xiong Y, Qi Y, Liu C, Huang S, Yao G, et al. Data Resource Profile: Xiamen registry of pregnant women and offspring (REPRESENT): a population-based, long-term follow-up database linking four major healthcare data platforms. *Int J Epidemiol*. 2021;50(1), 27-28.
6. Lin H, Tang X, Shen P, Zhang D, Wu J, Zhang J, et al. Using big data to improve cardiovascular care and outcomes in China: a protocol for the CHinese Electronic health Records Research in Yinzhou (CHERRY) Study. *BMJ Open*. 2018;8(2), e019698.
7. Yang Y, Zhou X, Gao S, Lin H, Xie Y, Feng Y, et al. Evaluation of electronic healthcare databases for post-marketing drug safety surveillance and pharmacoepidemiology in China. *Drug Safety*. 2018;41(1), 125-137.
8. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med*. 2015;12(10), e1001885.
9. Jorgensen NW, Sibley CT, McClelland RL. Using imputed pretreatment cholesterol in a propensity score model to reduce confounding by indication: results from the multi-ethnic study of atherosclerosis. *BMC Med Res Methodol*. 2013;13:81.
10. Cox E, Martin BC, Van Staa T, Garbe E, Siebert U, Johnson ML. Good research practices for comparative effectiveness research: approaches to mitigate bias and confounding in the design of nonrandomized studies of treatment effects using secondary data sources: the International Society for Pharmacoeconomics and Outcomes Research Good Research Practices for Retrospective Database Analysis Task Force Report–Part II. *Value Health: J Int Soc Pharmacoeconom Outcomes Res*. 2009;12(8), 1053-1061.
11. Kyriacou DN, Lewis RJ. Confounding by indication in clinical research. *JAMA*. 2016;316(17), 1818-1819.
12. Zhang Y, Xie YM, Chen C, Zhang C, Zhuang Y. Association rules analysis of Fufang Kushen injection in combination with traditional Chinese medicine or modern medications in treating malignant tumor: real-world retrospective study. *Zhongguo Zhong yao za zhi*. 2017;42(2), 378-384.
13. Wu K, Deng D, Yu B, Han Z, Huang L, He Y, et al. Evaluation of the efficacy and safety of Chinese herbal injection combined with trimetazidine for viral myocarditis: a network meta-analysis. *Front Pharmacol*. 2021;12:630896.
14. Sun LX, Xie YM, Liu H, Li YY, Zhuang Y. Analysis of clinical application characteristics and drug combination of Ciwujia injection in 12 554 cases in real world. *Zhongguo Zhong yao za zhi*. 2020;45(15), 3505-3510.
15. Jin X, Wang K, Zhai J, Liu C, Wang H, Liu Z. Analysis of injection of Xuesaitong (lyophilized) in 30 097 cases in real world. *China J Chin Materia Medica*. 2020;45(20), 5024-5028.
16. Feng S, Han M, Lai L, Wang SC, Liu JP. Research capacity at traditional chinese medicine (TCM) centers in China: a survey of clinical investigators. *Evid-Based Complement Alternat Med*. 2017;2017:4231680.
17. Psaty BM, Siscovick DS. Minimizing bias due to confounding by indication in comparative effectiveness research: the importance of restriction. *JAMA*. 2010;304(8), 897-898.
18. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivar Behav Res*. 2011;46(3), 399-424.
19. Uddin MJ, Groenwold RH, Ali MS, de Boer A, Roes KC, Chowdhury MA, et al. Methods to control for unmeasured confounding in pharmacoepidemiology: an overview. *Int J Clin Pharm*. 2016;38(3), 714-723.
20. Infante-Rivard C, Cusson A. Reflection on modern methods: selection bias—a review of recent developments. *Int J Epidemiol*. 2018;47(5), 1714-1722.
21. Xie G, Cui Z, Peng K, Zhou X, Xia Q, Xu D. Aidi Injection, a traditional chinese medicine injection, could be used as an adjuvant drug to improve quality of life of cancer patients receiving chemotherapy: a propensity score matching analysis. *Integr Cancer Therap*. 2019;18:1534735418810799.
22. Payet C, Polazzi S, Obadia JF, Armoiry X, Labarere J, Rabilloud M, et al. High-dimensional propensity scores improved the control of indication bias in surgical comparative effectiveness studies. *J Clin Epidemiol*. 2021;130:78-86.
23. Tian Y, Schuemie MJ, Suchard MA. Evaluating large-scale propensity score performance through real-world and synthetic data experiments. *Int J Epidemiol*. 2018;47(6), 2005-2014.
24. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology (Cambridge, Mass)*. 2000;11(5):550-560.
25. Koladjo BF, Escolano S, Tubert-Bitter P. Instrumental variable analysis in the context of dichotomous outcome and exposure with a numerical experiment in pharmacoepidemiology. *BMC Med Res Methodol*. 2018;18(1):61.
26. Stukel TA, Fisher ES, Wennberg DE, Alter DA, Gottlieb DJ, Vermeulen MJ. Analysis of observational studies in the presence of treatment selection bias: effects of invasive cardiac management on AMI survival using propensity score and instrumental variable methods. *JAMA*. 2007;297(3), 278-285.
27. He Cheng-jun, Peng C, Dai O, Yang L, Lliu J, Guo L, et al. Chemical constituents from *Leonurus japonicus* injection. *Chin Trad Herbal Drugs*. 2014;45(21), 3048-3051.
28. Qide Lin JH, Xinghui Liu. Consensus on application of motherwort injection in prevention and treatment of postpartum hemorrhage. *Chin J Pract Gynecol Obstetr*. 2017;33(10), 1053-1056. 2017.
29. Zhou Y. *Research of the Effects & the Safety of Leonurus Injection on the Human Uterine Contraction and Hemostatic affection in Vitro and Vivo*. Sichuan Univeristy; 2007.
30. *The Prevention and Management of Postpartum Haemorrhage: Report of Technical Working Group*, Geneva 3–6 July 1989. World Health Organization, 1990.
31. Dahlke JD, Mendez-Figueroa H, Maggio L, Hauspurg AK, Sperling JD, Chauhan SP, et al. Prevention and management of postpartum hemorrhage: a comparison of 4 national guidelines. *Am J Obstetr Gynecol*. 2015;213(1), 76.e1-.e10.
32. Prevost M, Zelkowitz P, Tulandi T, Hayton B, Feeley N, Carter CS, et al. Oxytocin in pregnancy and the postpartum: relations to labor and its management. *Front Public Health*. 2014;2:1.
33. Lousdal ML. An introduction to instrumental variable assumptions, validation and estimation. *Emerg Themes Epidemiol*. 2018;15:1.

How to cite this article: Tan J, Liu C, Li M, et al. A methodological framework for tackling confounding by indication when assessing the treatment effects of Chinese herbal injections in the real world. *J Evid Based Med*. 2022;15:64–72. https://doi.org/10.1111/jebm.12462