項目編號：**21B2_018A**

調研／研究報告

«中醫隨機對照試驗設計中不同干預模式的療效和安全性比較研究 »

項目統籌人:黃仰山

項目小組成員:梁挺雄，吳心音，黃凱琳，何子峰，何輝輝

申請機構:香港中文大學

內容

摘要

療效是中醫藥發展的根基，而隨機對照試驗(RCT)是評估中醫療效最科學的研究設計。辨證論治是中醫認識和治療疾病的基本原則，然而，對於是否有必要依照辨證論治的原則設計關於中藥療效的 RCT，目前尚無定論。本項目旨在採用薈萃-流行病學研究和網絡薈萃分析，結合專家意見咨詢的定性研究方法，回答中藥辨證論治與非辨證論治治療之間療效和安全性是否存在差異及其大小的科學問題，為中醫藥 RCT 的研究設計提供方法學指導和科學依據。

第一階段的薈萃-流行病學研究共涉及 137 個合資格的系統綜述，包括 2,064 個 RCTs 和 187,503 名研究對象，覆蓋了多種臨床疾病。所有合資格系統綜述都報告了中藥療效指標，主要以二分類指標(110，80.3%)和主觀結局(109，79.6%)為主。只有 8%的薈萃分析報告了中藥安全性指標，所有指標均為二分類指標和主觀結局。薈萃-流行病學研究結果發現，無論是二分類療效指標還是連續性療效指標，採用非辨證論治干預模式設計的 RCT 與採用辨證論治干預模式設計的 RCT 在報告中藥療效方面的差異在統計學上沒有顯著性。在中醫藥安全性方面，中醫藥 RCT 是否採用辨證論治干預模式對中藥不良反應風險大小的影響也沒有顯著性差異。二分類療效指標的亞組分析的結果顯示，對於循環系統疾病以及主觀結局，採用辨證論治的 RCT 得到的干預療效較小。在消化系統疾病方面，採用辨證論治干預模式設計的 RCT 發生不良反應的風險更高。其他亞組分析結果沒有統計學意義。

根據第一階段薈萃-流行病學研究的研究結果及團隊的研究經驗，第二階段的德爾菲研究將從消化系統疾病上選擇一個臨床問題開展第三階段研究。基於 GRADE 證據決策框架，我們開展了兩輪德爾菲研究。兩輪德爾菲專家咨詢過後，5 種臨床問題達成了 75%的陽性共識。其中，胃食管反流病(GERD)納入的 RCT 數量最多(3 個薈萃分析，48 個 RCT)，該臨床問題將進入第三階段網絡薈萃分析研究。

基於第二階段專家共識的定性研究，本項目針對 GERD，開展第三階段網絡薈萃分析。網絡薈萃分析共納入 34 個合資格的 RCT，包括 4,226 名研究對象， 評估了 28 種不同的中藥治療方案，其中有 20 項 RCT 評估了中藥與質子泵抑製劑(PPIs)的聯合使用，餘下的 14 項則對單獨使用中藥進行評估。在 34 項 RCTs 中，26 項 RCTs 在試驗設計時採用了辨證論治的原則設計(SD)，而其餘 8 項則沒有採用辨證論治的原則設計(NSD)。網絡薈萃分析結果顯示，除了胃食管反流病量表(GERDQ)癥狀評分變化外，未採用辨證論治的中藥治療在其他所有測量指標上通常最為有效。柴芍六君子湯合烏貝散(NSD)可能是治療 GERD 患者提高癥狀緩解率和反流性食管炎量表(RDQ)評分改善的最有效治療方案。和胃益气降逆汤(NSD)+PPI 在降低 GERD 患者复发率方面表现出最高的有效性。此外，半夏瀉心湯(NSD)、降逆和胃湯(NSD)和和胃益氣降逆湯(NSD)+PPI 顯示出更低的不良事件風險。儘管這些發現與我們之前的薈萃流行病學研究結論一致，但仍需要開展實證研究進一步評估。鑒於 GERD 中肝胃不和證或脾胃氣虛伴胃氣上逆證等中醫診斷的普遍性，評估它們對治療結果的預測價值，對於評估辨證論治在指導 GERD 個體化中藥方劑治療中的作用至關重要。

正文

-簡介

中醫擁有豐富的歷史，亦在中國和海外華人群體當中得到廣泛應用[1,2]。世界衛生組織第 11 版國際疾病分類(ICD-11) 亦納入了有關中醫診斷對應的疾病及證型的章節[3]。中醫的「證」為疾病的特定階段的病理變化，涉及其病因、性質和病位[4]。辨證論治為根據四診(望、聞、問、切)所收集的資料，去指導中醫師對病人進行個體化的診斷和治療[3,5]。在中醫診療中，兩個有相同西醫診斷的病人亦可以被診斷為不同的中醫證型，從而開具特定的中藥。例如，功能性消化不良患者可能被診斷為肝氣犯胃或脾胃氣虛，常用逍遙丸或香砂六君子顆粒來治療[6]。

循證醫學為中醫藥發展提供了新的契機。於療效評價方面，經過嚴謹設計的隨機對照試驗(randomised controlled trial; RCT)是評估治療有效性的黃金標準[7,8]。中藥複方 RCT 報告規範 2017 要求研究者在報告 RCT 時需要涉及辨證論治的細節[9]。就一些中醫師而言，辨證論治是中醫經典理論和臨床實踐的總合結果，也是辨別個體化中藥治療的關鍵點[10]。沒有「辨證」的 RCT 設計並不能充分體現中醫的臨床情況，局限了研究結果的有效性和普遍性，甚至低估治療效果[5]。針對辨證論治在中醫藥 RCT 的實施情況，一項系統綜述發現[11]，1999 年至 2017 年發表的 2,955 個 RCT 中，只有 25(6.3%)個中成藥 RCT 在研究設計時考慮了辨證論治的原則。在一項逍遙丸治療功能性消化不良的安慰劑對照試驗中[12]，參與研究的病人只需符合西醫對功能性消化不良的診斷，這種研究設計大大降低了模型的信度。未經辨證論治而開具中藥的情況並不少見。例如，於 Covid-19 期間[13]，儘管中國國家中醫診療指南建議使用辨證論治，但仍有很多省級或機構級的指南提倡使用指定方劑(如：肺炎一號方)。

既然辨證論治是中醫臨床實踐的基本原則之一，如果開展關於中醫療效的 RCT 研究，照搬西醫的標準治療的思維，是否會導致中醫實際療效被低估？此外，考慮治療方案的臨床應用性時，還應該考慮其安全性。沒有採用辨證論治干預模式設計的 RCT 是否會增加中醫藥治療方案的毒副作用？因此，中醫 RCT 研究是否都應該採取辨證論治的研究設計？這些是開展中醫 RCT 研究需要慎重考慮的關鍵技術問題，但是目前並沒有統一的答案。

-宗旨和目標

本項目的總體研究目標是比較中醫藥 RCTs 研究設計中，採用辨證論治干預模式設計與採用非辨證論治干預模式設計之間的療效及安全性是否存在差異及其大小。這一總體目標主要包含兩個內容:

1) 中藥辨證論治干預模式設計的 RCTs 與非辨證論治干預模式設計的 RCTs 在治療效果上，是否存在差異？如果有差異，差異的大小是多少？

2) 中藥辨證論治干預模式設計的 RCTs 與非辨證論治干預模式設計的 RCTs 在治療疾病的安全性上，是否存在差異？如果有差異，差異的大小是多少？

上述兩個研究目標主要通過三個研究從兩個層面來實現。三個研究分別為 1)薈萃-流行病學研究，2)基於專家共識的定性研究，和 3)網絡薈萃分析研究(Network meta-analysis; NMA)。兩個層面是多個疾病的總體層面和單個疾病的個體層面。

-方法

1. 第一階段:薈萃流行病學研究

1.1. *納入標準*

(1)系統綜述和薈萃分析

合資格的系統綜述需至少包含一個評價中藥 RCT 療效/安全性的薈萃分析。合資格的薈萃分析需納入:1)至少一個採取辨證論治的 RCT；和 2)至少一個採取非辨證論治的 RCT。爲確保有足夠的統計分析能力，合資格的薈萃分析需至少納入 10 個 RCT[14]。若系統綜述開展了多個薈萃分析，選擇包含中藥 RCT 數量最多的結局指標[15]。

(2) 中藥 RCT 需滿足以下標準:

- 病人類型:無限制
- 干預和對照措施: 干預需為口服中藥，且被收錄在 2020 年中國藥典中。RCT 比較組需至少滿足以下一種:1)中藥對比常規治療；2)中藥聯合常規治療對比常規治療；和 3)中藥對比安慰劑。在同一薈萃分析中，不同 RCT 需具有相同的比較組。

- 結局: 任何評估中藥 RCT 療效/安全性的指標。

1.2. 文獻檢索及篩選

本研究檢索了考科藍系統綜述數據庫、MEDLINE、EMBASE、中國生物醫學文獻數據庫、萬方、中國學術期刊全文數據庫(CNKI)和台灣期刊文獻數據庫。時間限制在 2021-2022。依據納入排除標準進行文獻篩選。

1.3. 數據提取

對合資格的薈萃分析進行數據提取，包括:發表年份，疾病類型，資金來源，RCT 偏倚風險評估工具名稱，納入 RCT 數量，療效/安全性結局數據(二分類和連續性指標)。RCT 結局數據分為主觀結局和客觀結局。

1.4. 中藥 RCT 辨證論治和偏倚風險評估

(1) 中藥 RCT 滿足以下任何一種情況時，就認爲其採取了辨證論治：

- 基綫招募屬於同一種證型的病人且給與同樣的藥方；
- 基綫招募屬於不同證型的病人且根據證型不同給與不同的藥方；
- 根據病人所屬證型，給予對應的藥方，必要時可適當進行調整。

(2) 採用考科藍偏倚風險評估工具[16]對中藥 RCT 進行質量評估。

在文獻檢索和篩查中，我們面臨著龐大的文獻數量，包括 6,140 篇系統綜述和潛在符合資格的系統綜述中的4,486篇RCTs。為了平衡項目推進和篩選質量，我們採用了以下方式：一位研究員進行文獻篩選，數據提取和偏倚風險評估，另一位研究員隨機選擇20%的文獻進行檢查，意見不統一時通過討論達成一致。對於無法達成一致的，由第三名研究者參與並做出最終決策。這種方法在實踐中被廣泛認可，既確保審查的全面性，同時也使得處理龐大的數據集更加切實可行。

1.5. 數據分析

描述性分析使用頻數和百分比描述二分類指標，中位數和範圍描述連續性指標。使用卡方檢驗評估採用辨證論治與採用非辨證論治 RCT 之間的差異。薈萃流行病學主要使用兩步法[17]來分析中藥 RCT 辨證論治與療效/安全性之間的關聯。 具體方法可見附錄 1。根據疾病類型、結局類型和資金支持進行亞組分析。在薈萃-回歸模型中調整了 RCT 樣本量、資金支持類型，進行敏感性分析。此外，本研究還調整了 RCT 的偏倚風險，以控制由於 RCT 方法學質量可能引起的混雜。

2. 第二階段:基於專家共識的定性研究

第一階段薈萃流行病學研究結果顯示，採用辨證論治干預模式與採用非辨證論治干預模式的中藥 RCT 之間中藥療效的差異無統計學意義。對於消化系統疾病，相比於採用非辨證論治干預模式的 RCT，採用辨證論治干預模式的 RCT 發生不良反應的風險更高。因此，必須深入探討辨證論治在診療消化系統疾病的價值。根據這個研究結果及團隊的研究經驗，本項目的第二階段主要從第一階段納入的薈萃分析所涉及的消化系統疾病裡選擇一個臨床問題開展第三階段研究。本階段詳細的工作流程如圖 1 所示。在數據收集之前，本項目已獲得香港中文大學調查與行為研究倫理委員會的倫理批准(參考編號：SBRE-22-0091)。在進行德爾菲研究之前，我們將第一階段的薈萃流行病學分析結果總結成一個清單(具體見附錄 2 證據概要表)，清單包括每個薈萃分析回答的臨床問題，以及兩種干預模式的療效和安全性差異。

CHM: 中藥; RCT: 隨機對照試驗; NMA，網絡薈萃分析

圖 1. 兩輪德爾菲研究流程圖

2.1. 專家的選擇及培訓

本研究的專家由 5 名中醫師，5 名經過中醫培訓的西醫師，以及 2 名臨床流行病學方法學專家組成，這些專家在各自的領域上需有十年以上工作經驗。其中十名臨床專家的挑選均參考薈萃流行病學分析涉及的臨床問題的疾病類型。所有專家的身份信息在研究過程中會保密以保證隱私及專家做出判斷的獨立性。

所有的專家均通過項目負責人和團隊成員的社交網絡透過電郵的方式招募。研究團隊根據先前所列的專家納入標準，評估專家候選人，確保他們符合選取標準。獲得專家同意並簽署知情同意書之後，我們通過電郵收集專家個人資料並發送本項目相關的培訓資料，對專家進行培訓，培訓內容包括研究目的、研究方法、薈萃流行病學分析方法以及第一階段的研究結果、GRADE 證據概要表以及德爾菲研究過程的概述。

2.2. 第一輪德爾菲研究數據收集

在專家培訓之後，我們通過電郵發送第一輪德爾菲研究的問卷。德爾菲問卷主要根據 GRADE 證據決策框架[18]設計。根據 GRADE 證據決策框架設計的德爾菲問卷主要包含考慮以下內容：

• Q1 辨證論治與非辨證論治干預模式在中藥治療該消化系統疾病療效上的差異，您認為相關證據的質量級別如何？

• Q2 您認為相比於非辨證論治干預模式，實施辨證論治對中藥治療該消化系統疾病，是否帶來很大的積極影響（如： 有較好的治療效果， 等）？

• Q3 您認為相比於非辨證論治干預模式，實施辨證論治對中藥治療該消化系統疾病，是否帶來很大的消極影響（如： 有較大的副作用， 等）？

• Q4 您認為相比於非辨證論治干預模式，實施辨證論治對中藥治療該消化系統疾病的積極影響大於消極影響嗎（如： 療效比副作用大）？

• Q5 關於比較辨證論治與非辨證論治在中藥治療該消化系統疾病療效上的差異，您認為回答這個問題在臨床上重要嗎？

• Q6 關於比較辨證論治與非辨證論治在中藥治療該消化系統疾病安全性上的差異，您認為回答這個問題在臨床上重要嗎？

綜合考慮這些因素之後，我們邀請德爾菲專家對該臨床問題開展 NMA 比較辨證論治與非辨證論治的療效和安全性差異的必要性進行打分(基於總分為 5 分的李克特量表[19]打分)。同時，我們邀請參與者提供定性意見，解釋他們給出的判斷的理由。

2.3. 第一輪德爾菲研究數據分析

對於每一個臨床問題，本研究計算了專家打分的中位數和四分位數間距，並統計同意推薦該臨床問題開展第三階段 NMA 研究的專家人數和百分比。本研究採納既往研究推薦的70%的共識標準[20]。即對於列表清單上的每一個臨床問題，如果 70%以上的專家給予必要或者非常必要的評分，則認為該臨床問題達成陽性共識；相反的，70%以上的專家給予沒必要或者非常沒必要的評分臨床問題，則認為達成陰性共識。達成陰性共識的臨床問題會移出清單列表，不再進入第二輪德爾菲研究。

2.4. 第二輪德爾菲研究

排除第一輪已達成陰性共識的臨床問題後，剩下的臨床問題進入第二輪德爾菲研究。在第二輪德爾菲研究中，我們在問卷中加入以下內容供專家參考:1)專家本人對每個臨床問題開展 NMA 必要性的第一輪評分；2)所有專家的中位評分和四分位數間距；3)每個專家不記名的定性修改意見。基於以上信息，專家再一次對於每個臨床問題開展 NMA 的必要性進行打分。

第二輪德爾菲研究的分析方法與第一輪德爾菲研究方法一致。在兩輪德爾菲研究之後，我們確定了一個具有專家共識的有必要進一步開展 NMA 的臨床問題清單。對於這些達成共識的臨床問題，我們從第一階段的薈萃流行病學研究[21]中提取了相關薈萃分析及其中 RCTs 的數量和入組參與者數量，選擇納入 RCT 最多的臨床問題進一步開展 NMA。

3. 第三階段:網絡薈萃分析研究 (NMA)

第二階段確定胃食管反流病(GERD)為第三階段 NMA 的研究重點。整個研究根據考科蘭手冊描述法開展計劃。整個研究根據考科蘭手冊中描述的方法開展。研究計劃已於 PROSPERO 資料庫註冊(註冊編號：CRD 42024516857)。

3.1. 入選標準

符合納入標準的 RCTs 需滿足以下條件：

i)招募年齡為 18 歲或以上的成年患者，經醫生確診患有 GERD，包括糜爛性、非糜爛性和非特異性 GERD。本研究對 RCT 的設置 (包括住院病人和門診病人)沒有限制。

ii) 干預需為口服中藥，且其中至少包含一種 2020 年版《中國藥典》中列出的成分。

iii)根據美國胃腸病學院(ACG)的推薦[22]，質子泵抑製劑(PPIs)，如奧美拉唑和雷貝拉唑，是 GERD 的推薦治療方法。因此，合格的 RCT 需使用 PPIs 作為對照，在 NMA 開展過程中作為橋樑，為不同中藥方之間療效和安全性的比較提供間接證據。RCT 比較組需至少滿足以下一種:1)中藥對比 PPI；2)中藥聯合 PPI 對比 PPI。

iv)根據多學會專家推薦和指南建議[22,23]，NMA 的主要結局指標是隨訪期 8 周及以上的癥狀緩解率 。次要結局指標包括通過反流性疾病問卷(RDQ)或胃食管反流病問卷(GERDQ)測量的癥狀評分變化；12 個月復發率[24]；以及隨訪期至少 8 周的不良事件發生率。RCT 報告了這些結局指標中的任何一項均被視為符合納入條件。

排除的研究類型包括：敘述性綜述、系統綜述、方案、會議摘要、評論論文或未關注 GERD 治療療效或安全性的 RCT。

3.2. 文獻檢索與篩選

本研究檢索了考科藍系統綜述數據庫、MEDLINE、EMBASE、CINAHL、聯合與補充醫學數據庫(AMED)、萬方數據庫、CBM、 CNKI、和台灣期刊文獻數據庫。檢索時間從各資料庫建立之初至 2023 年 7 月，以識別可能符合條件的 RCTs。為了提高检索的全面性，我們還搜索了臨床試驗註冊網站，包括中國臨床試驗註冊中心(http://www.chictr.org.cn)和 Clinicaltrials.gov。Drugs@FDA 及歐洲藥品管理局公開評估報告(EPAR)較少關於中藥療效

相關研究，因此在實際操作過程中不予考慮。此外，我們還評估了薈萃流行病學研究中薈萃分析所有的 RCTs 的合格性。詳細的檢索策略和檢索結果見附錄 3。

文獻檢索后，我們將檢索到的文獻導出到 EndNote 中。基於標題和摘要進行初篩，對可能符合條件的文獻進行全文評估，確定是否最終納入或排除。對於多次發表的同一個 RCTs，我們選擇最新發表的或納入病人最多的版本。

3.3. 數據提取、偏倚風險評估和證據質量評級

使用預先設計的數據提取表格來收集資料，包括 RCT 基本特徵資料、PICOS 相關的信息以及所有預先指定結局的定量結果。RCT 是否採用了辨證論治干預模式，判斷標準與第一階段薈萃流行病學研究相同。

RCTs 的方法學質量通過 Cochrane 偏倚風險評估工具 2[25]來評估， NMA 的證據質量評級通過用 GRADE 證據質量評價方法進行[26]，評級分為非常低、低、中或高。

文獻篩選、數據提取、偏倚風險評估和證據質量評級均由兩名團隊人員獨立進行。若存在分歧，則通過討論解決，必要時通過諮詢項目負責人解決分歧。

3.4. 數據分析

在這個階段的研究中，標準配對薈萃分析對於我們的研究問題沒有額外的作用，因為我們需要同時比較不同採用辨證論治干預模式的中藥治療方案、非辨證論治干預模式的中藥治療方案之間的結果。 另一方面，有些中藥治療方案是單獨使用中藥方，有些是聯合了中藥方和 PPI，這使得標準配對分析無法展示辨證論治與非辨證論治中藥干預模式的差異。因此，本項目的數據分析專注於網絡薈萃分析。

在分析不同的結局指標時，首先將評估了同一結局指標的 RCT 中所有治療方案(如辨證論治中藥方，非辨證論治中藥方，PPI)通過共同的對照措施(PPI)連接，形成網絡幾何圖，以更直觀形象的瞭解整個網絡薈萃分析所包含的治療方案。每一種治療方案將用一個點表示，

舉個例子，有一個 RCT 比較了兩種治療方案(如辨證論治中藥方與 PPI)，那麼辨證論治中藥方和 PPI 這兩個點將通過一條線連接。最終，每一個點的大小代表了 RCT 中隨機分配到該治療方案的病人比例，而線條的粗細代表進行了此線條所連接兩種治療方案療效對照的 RCT 數。NMA 的效應估計通過比值比(Risk ratios, RRs) 或者加權均數差(Weighted mean differences, WMDs)及其 95%置信區間(Confidence intervals, CIs)表示。最後，描繪每種治療方案的療效或安全性排名曲線，計算每種治療方案的累積曲線下面積(the surface under the cumulative ranking curve, SUCRA)[27]，所有的治療方案(包括辨證論治中藥方，非辨證論治中藥方，PPI)將分別按療效和安全性進行排序，找出療效最好的治療方案和安全性最高的治療方案。曲線下面積(SUCRA)用於展示不同治療方案的效應和安全性排名，SUCRA 得分越高，表示排名越靠前。

為更全面地解讀 NMA 結果，我們採用 GRADE 工作組開發的 Minimally contextualized framework[28]，綜合考慮 NMA 的效應估計、證據質量評級和 SUCRA 排名後，將每種中藥治療方案根據其療效和安全性分類到不同的組別中。首先，根據 NMA 的效應估計對每種治療方案進行分組。組 0 為與 PPIs 療效/安全性無顯著性差異的中藥治療方案。組 1 的中藥治療方案優於組 0 中的至少一種治療方案。如果存在更優的中藥治療方案，則繼續分類到更高的組別。在每個組內，再進一步根據證據質量評級對中藥治療方案進行分類：i)高證據質量評級(中或高等證據質量)和低證據質量評級(低或非常低證據質量)。最後，根據 SUCRA 排名確定所有治療方案的分類，確保一致性。

發表偏倚

發表偏倚通過 Egger' s 檢驗以及漏斗圖進行評估。顯著性水準設為 $p \leq 0.10$[29]。僅當某個結局的研究數量≥10 項時，發表偏倚的評估[30]才可行。所有數據分析均使用 R 版本 4.3.2 進行。

-結果

第一階段:薈萃流行病學研究

本研究共納入 137 個系統綜述，包括 2,064 個 RCT 和 187,503 名研究對象。所有的薈萃分析均報告了中藥干預療效結果。這些干預療效指標大多是二分類(110，80.3%)和主觀結局(109，79.6%)。只有 11 個(8%)薈萃分析報告了安全性結局，且都是二分類指標和主觀結局(表 1)。在 2,064 個 RCT 中，有 1,049(50.8%)採用了辨證論治干預模式設計(表 2)。中藥 RCT 偏倚風險評估結果見附錄 4。

對於二分類療效指標，與採用非辨證論治干預模式的中醫藥 RCT 相比，採用了辨證論治干預模式的中醫藥 RCT 得到的干預療效更小(ROR:0.93，95% CI:0.86，1.00，*P*=0.04)(圖 2)。ROR 值的 95% CI 包括無效值 1，這表明採用了辨證論治干預模式與採用非辨證論治干預模式的中藥 RCT 得到的干預療效差異無統計學意義。本研究在連續性療效指標也觀察到了類似的結果(dSMD:-1.44，95% CI:-3.95，-1.08，*P*=0.26)(圖 2)。中醫藥 RCT 是否採用辨證論治干預模式的安全性差異無統計學意義(ROR:1.13，95% CI:0.66，1.92，*P*=0.66)(圖 2)。

二分類療效指標的亞組分析中，對於循環系統疾病和主觀結局，相對於採用非辨證論治的 RCT，採用辨證論治的 RCT 得到的干預療效均更小。對於消化系統疾病，相對於採用非辨證論治的 RCT，採用辨證論治的 RCT 發生不良反應的風險更高。其餘亞組分析的結果均無統計學意義(圖 2)。二分類療效指標的敏感性分析中，當調整了對研究者和護理人員實施盲法或對結局評估者實施盲法時，相對於採用非辨證論治的 RCT，採用了辨證論治 RCT 得到的干預療效更小(附錄 5)。

表 1 本項目納入的 137 篇系統綜述的一般特徵

特徵	結果*
系統綜述的特徵	
發表年份	
2021	94 (68.6)
2022	43 (31.4)
疾病分類(ICD-11)	
循環系統疾病	35 (25.5)
泌尿生殖系統疾病	27 (19.7)
消化系統疾病	23 (16.8)
內分泌、營養或代謝疾病	10 (7.3)
其他疾病類型#	42 (30.7)
資金支持類型	
非商業資金支持	84 (61.3)
商業和非商業資金支持	1 (0.7)
未報告	52 (38.0)
採用的偏倚風險評估工具	
Cochrane 偏倚風險評估工具	105 (76.6)
Cochrane 偏倚風險評估工具和 Jadad 量表	16 (11.7)
Jadad 量表	15 (10.9)
Cochrane 偏倚風險評估工具 2.0	1 (0.7)
系統綜述中包含的薈萃分析的特徵	
薈萃分析中包含的合資格 RCT 數目的中位數, 範圍	14, 10-36
報告療效結果的薈萃分析數目	137 (100.0)
療效指標類型	
二分類指標	110 (80.3)
連續性指標	27 (19.7)
療效結局類型	
主觀結局	109 (79.6)
客觀結局	28 (20.4)
報告安全性結果的薈萃分析數目	11 (8.0)
安全性指標類型	
二分類指標	11 (100.0)
連續性指標	0 (0.0)
安全性結局類型	
主觀結局	11 (100.0)
客觀結局	0 (0.0)

*除非特殊說明，表格中的數字代表頻數(百分比)；#其他疾病類型包含 11 種分類:精神、行爲或神經發育障礙 (8, 5.8%), 睡眠-覺醒障礙 (6, 4.4%), 肌肉骨骼系統或結締組織疾病 (5, 3.6%), 神經系統疾病 (4, 2.9%), 免疫系統疾病 (4, 2.9%), 呼吸系統疾病 (4, 2.9%), 某些感染性疾病或寄生蟲病 (3, 2.2%), 腫瘤 (3, 2.2%), 癥狀、體徵或臨床所見，不可歸類在他處者 (2, 1.5%), 損傷、中毒或外因的某些其他後果 (2, 1.5%) 和視覺神經系統 (1, 0.7%)； ICD-11，國際疾病分類第 11 版

表 2 2,064 隨機對照試驗(RCT)的一般特征*

特征	總體 (N=2,064)	未採用辨證論治的 RCT (n=1,015)	採用辨證論治的 RCT (n=1,049)	*P*[#]
發表年份的中位數，範圍	2016, 1997-2021	2016, 1997-2021	2017, 1997-2021	<0.001
感興趣結局隨訪時間的中位數[a], 範圍 (單位：周)	8, 0.14-114	8, 0.14-92	8, 0.29-114	0.341
RCT 包含的參與者總人數	185,333	95,982	89,351	-
樣本量中位數，範圍	80, 28-560	82, 28-560	78, 30-480	<0.001
發表語言				
中文	2,060 (99.8)	1,013 (99.8)	1,047 (99.8)	1.000
英文	4 (0.2)	2 (0.2)	2 (0.2)	
比較組類型				
CHM ＋常規治療 vs 常規治療	1,377 (66.7)	685 (67.5)	692 (66.0)	0.464
CHM vs 常規治療	687 (33.3)	330 (32.5)	357 (34.0)	
CHM 類型				
湯劑	1,625 (78.7)	774 (76.3)	851 (81.3)	<0.001
中成藥	231 (11.2)	159 (15.7)	72 (6.9)	
CHM 顆粒	125 (6.1)	41 (4.0)	84 (8.0)	
CHM 提取物	3 (0.1)	3 (0.3)	0 (0.0)	
未報告	77 (3.7)	37 (3.6)	40 (3.8)	
RCTs 類型				
雙臂	1,993 (96.6)	981 (96.7)	1,012 (96.5)	0.825
多臂	71 (3.4)	34 (3.3)	37 (3.5)	
中心數量				
單中心	1,921 (93.1)	937 (92.3)	984 (93.8)	0.010
多中心	49 (2.4)	19 (1.9)	30 (2.9)	
未報告	94 (4.6)	59 (5.8)	35 (3.3)	
資金支持				
非商業支持	380 (18.4)	139 (13.7)	241 (23.0)	<0.001
商業支持	1 (0.1)	1 (0.1)	0 (0.0)	
未報告	1,683 (81.5)	875 (86.2)	808 (77.0)	

*除非特殊說明，數值表示頻數(百分比); CHM, 中藥; [a] 感興趣結局的中位隨訪時間表示試驗報告的從基線到結果測量的隨訪時間的中位數. [#]*P* 值與調整的 $\chi 2$ 或 t 檢驗相關.

圖 2 薈萃流行病學分析(二分類和連續性結局，療效和安全性結局)，亞組分析(基於系統綜述不同的疾病分類，結局類型和資金支持)

說明: ROR，比值比之比；RCT，隨機對照試驗；CI，置信區間；SR，系統綜述；NA，不適用，因爲該亞組裏面僅包含一個 RCT；dSMD，標準化均差之差。對於療效結局，當結局數據為二分類指標時，比值比之比小於 1 表明相對於採用辨證論治的 RCT，採用非辨證論治的 RCT 的中藥療效更大；當結局數據為連續性指標時，標準化均差之差小於 0 表明相對於採用辨證論治的 RCT，採用非辨證論治的 RCT 的中藥療效更大。對於安全性結局，比值比之比大於 1 表明相對於採用非辨證論治的 RCT，採用辨證論治的 RCT 的中藥不良反應風險更高。

第二階段:基於專家共識的定性研究

招募的德爾菲專家小組成員的人口社會學特徵詳細信息見附錄 6。參考 GRADE 證據決策框架，我們開展了兩輪德爾菲研究，第一輪德爾菲研究後， 75%的專家就功能性消化不良和腹瀉型腸易激綜合徵這 2 個臨床問題達成了陽性共識，其餘臨床問題既未達成陽性共識，也未達成陰性共識(表 3)。

第一輪德爾菲研究的結果中沒有臨床問題達成陰性共識，因此，所有消化科臨床問題均需再次進入第二輪德爾菲研究。第二輪德爾菲研究過後，75%的專家認為有必要對以下 5 個消化科臨床問題開展網絡薈萃分析，即:非糜爛性胃食管反流病 、未特指的胃食管反流病 、慢性萎縮性胃炎、功能性消化不良及腹瀉型腸易激綜合徵。胃食管反流病(含非糜爛性胃食管反流病 、未特指的胃食管反流病)納入的 RCT 數量最多(3 個薈萃分析，48 個 RCT)，該臨床問題將進入第三階段網絡薈萃分析研究(表 4)。附錄 7 呈現了專家小組基於德爾菲問卷各條目對這些達成共識的臨床問題的評價結果。項目德爾菲研究收集的數據經過一位與研究無關的專業人士，中國大陸註冊的中醫師的核查和確認。

關於決策過程德爾菲專家的定性評價分別列於附錄 8a 和附錄 8b，分別對應達成共識和未達成共識的疾病。支持針對上述臨床問題進一步開展 NMA 的觀點認為，NMA 產生的證據能夠闡明辨證論治在選擇最佳治療策略中的必要性。如果辨證論治被認為是沒有必要的，臨床醫生可以簡化中醫診斷過程，從而提高臨床診療的整體效率。在這種情況下，中醫實踐的國際化可能會得到促進，因為海外臨床醫生可以採用更簡單的中醫診斷流程。

反對針對上述臨床問題進一步開展 NMA 的觀點認為，對於具有普遍中醫證型的疾病，辨证论治通常是不必要的。在臨床實踐中，使用針對普遍中醫證型的特定中藥方劑治療該疾病已經是足夠的。而且第一階段荟萃流行病學研究結果已經表明，辨证论治對中藥療效和安全性的影響並不顯著。對於某些特定疾病，如幽門螺桿菌感染，相對於中醫而言，西醫在其治療上具有明顯優勢，因此，在這些疾病上進一步開展 NMA 顯得不那麼必要。與此同時，也有專家認為辨證論治在中醫實踐中的地位作用不應受到質疑，尤其是在複雜的消化系統疾病(如克羅恩病或潰瘍性結腸炎)中。

表 3 第一輪德爾菲研究結果

臨床問題	同意推薦的人數	百分比	中位數(四分位數間距)	薈萃分析數量 (RCT 總數量， RCT 總樣本量)
非糜爛性胃食管反流病	7	58%	4 (1.0)	1 (13, 1,243)
未特指的胃食管反流病	6	50%	4 (1.0)	2 (35, 3,552)
幽門螺桿菌相關性胃炎	5	42%	3 (3.0)	3 (49, 4,459)
慢性萎縮性胃炎	7	58%	4 (2.0)	3 (35, 3,195)
淺表性胃炎	4	33%	3 (2.0)	1 (12, 1,186)
消化性潰瘍	6	50%	4 (2.0)	2 (22, 2,282)
克羅恩病	7	58%	4 (2.0)	1 (12, 905)
潰瘍性結腸炎	6	50%	4 (2.0)	3 (46, 4,204)
功能性消化不良	9	75%	4 (1.0)	2 (24, 2,180)
腹瀉型腸易激綜合徵	9	75%	4 (1.0)	2 (31, 2,999)
老年功能性便秘	7	58%	4 (2.0)	1 (19, 1,645)
幽門螺桿菌感染(療效)	5	42%	3 (3.0)	1 (16, 2,055)
幽門螺桿菌感染(安全性)	4	33%	3 (2.0)	1 (16, 2,055)

RCT，隨機對照試驗。藍色標記的是第一輪德爾菲研究後，專家達成陽性共識的臨床問題；達成陽性共識指的是大於或等於 75%的專家同意推薦該臨床問題進一步開展網絡薈萃分析。

表 4 第二輪德爾菲研究結果

臨床問題	同意推薦的人數	百分比	中位數	四分位數間距	薈萃分析數量 (RCT 總數量， RCT 總樣本量)
非糜爛性胃食管反流病	9	75%	4	0.25	1 (13, 1,243)
未特指的胃食管反流病	9	75%	4	0.5	2 (35, 3,552)
幽門螺桿菌相關性胃炎	5	42%	3	2	3 (49, 4,459)
慢性萎縮性胃炎	9	75%	4	1.25	3 (35, 3,195)
淺表性胃炎	6	50%	3.5	2	1 (12, 1,186)
消化性潰瘍	7	58%	4	1	2 (22, 2,282)
克羅恩病	8	67%	4	2	1 (12, 905)
潰瘍性結腸炎	8	67%	4	2	3 (46, 4,204)
功能性消化不良	9	75%	4	1.25	2 (24, 2,180)
腹瀉型腸易激綜合徵	9	75%	4	0.5	2 (31, 2,999)
老年功能性便秘	8	67%	4	2	1 (19, 1,645)
幽門螺桿菌感染(療效)	4	33%	3	2	1 (16, 2,055)
幽門螺桿菌感染(安全性)	4	33%	3	1.5	1 (16, 2,055)

RCT，隨機對照試驗。橙色標記的是第二輪德爾菲研究後，專家達成陽性共識的臨床問題；達成陽性共識指的是大於或等於 75%的專家同意推薦該臨床問題進一步開展網絡薈萃分析。

第三階段: NMA

本階段研究共納入 34 項合資格的 RCTs， 涉及 4,226 名 GERD 患者。文獻檢索與篩選過程的詳細過程見圖 3，納入的研究及其基本特征見附錄 9-10。這些試驗共評估了 28 種不同的中藥治療方案，其中有 20 項評估了中藥與 PPIs 的聯合使用，餘下的 14 項則對單獨使用中藥進行評估。在 34 項 RCTs 中，26 項 RCTs 在試驗設計時考慮了辨證論治的原則，而其餘 8 項則沒有採用。研究的對照組均為 PPI。所有 RCTs 均存在高偏倚風險，詳細的偏倚風險評估結果見附錄 11-12。

NMA 結果

癥狀緩解率的網絡幾何圖包括了 14 項雙臂試驗，涉及 12 種不同的中藥方及是否根據辨證論治原則設計(圖 4)。圖 5 展示了不同的中藥方及是否採用辨證論治原則在 GERD 患者癥狀緩解率這一主要結局指標上兩兩比較的結果。可以看到，柴芍六君子湯合烏貝散(Non syndrome differentiation, NSD,沒有採用辨證論治原則設計)在與其他比較組相比時，表現出顯著更高的療效，瀉心湯(syndrome differentiation, SD, 採用辨證論治原則設計) +PPI除外。

SUCRA 結果(圖 6)顯示柴芍六君子湯合烏貝散(NSD)最有可能是提高 GERD 患者癥狀緩解率的最佳治療方案，其次是四逆降逆湯(SD)+PPI 和蝶蠣左金湯(SD)+PPI。然而，支持這些結論的證據在 30 組比較中被評為低質量證據(38.5%)，在其餘 48 組比較中被評為非常低質量證據(61.5%)，詳情見附錄 13。

根據 Minimally contextualized framework，我們對 12 種中藥方及是否採用辨證論治原則在主要結局指標 GERD 患者癥狀緩解率的療效進行了分類，結果如下(表 5):

• 組別 0 ，低質量證據支持這些治療方案與單獨使用 PPI 的療效無顯著差異 。包括柴胡加龍骨牡蠣湯(SD)、降逆護膜湯(SD)+PPI、疏肝清胃降逆湯(SD)、半夏瀉心湯(NSD)、升降並舉法(NSD)、半夏瀉心湯(SD)以及自擬方(SD)。

• 組別 1，低質量證據支持這些治療方案可能不如組別 2 中最有效的治療方案，但可能優於組別 0 的治療方案。包括四逆降逆湯(SD)+PPI、蝶蠣左金湯(SD)+PPI、瀉心湯(SD)+PPI 和降逆和胃湯(NSD)。

• 組別 2，低質量證據支持這些治療方案可能是提升 GERD 患者癥狀緩解率最有效的治療方案，本組別中僅包含柴芍六君子湯合烏貝散(NSD)。

對於次要結局指標 RDQ 評分變化，16 種中藥方及是否採用辨證論治原則被分為四組(具體見表 6)。低質量證據支持柴芍六君子湯合烏貝散(NSD)可能是對改善 GERD 患者 RDQ 評分最有效的治療方案。

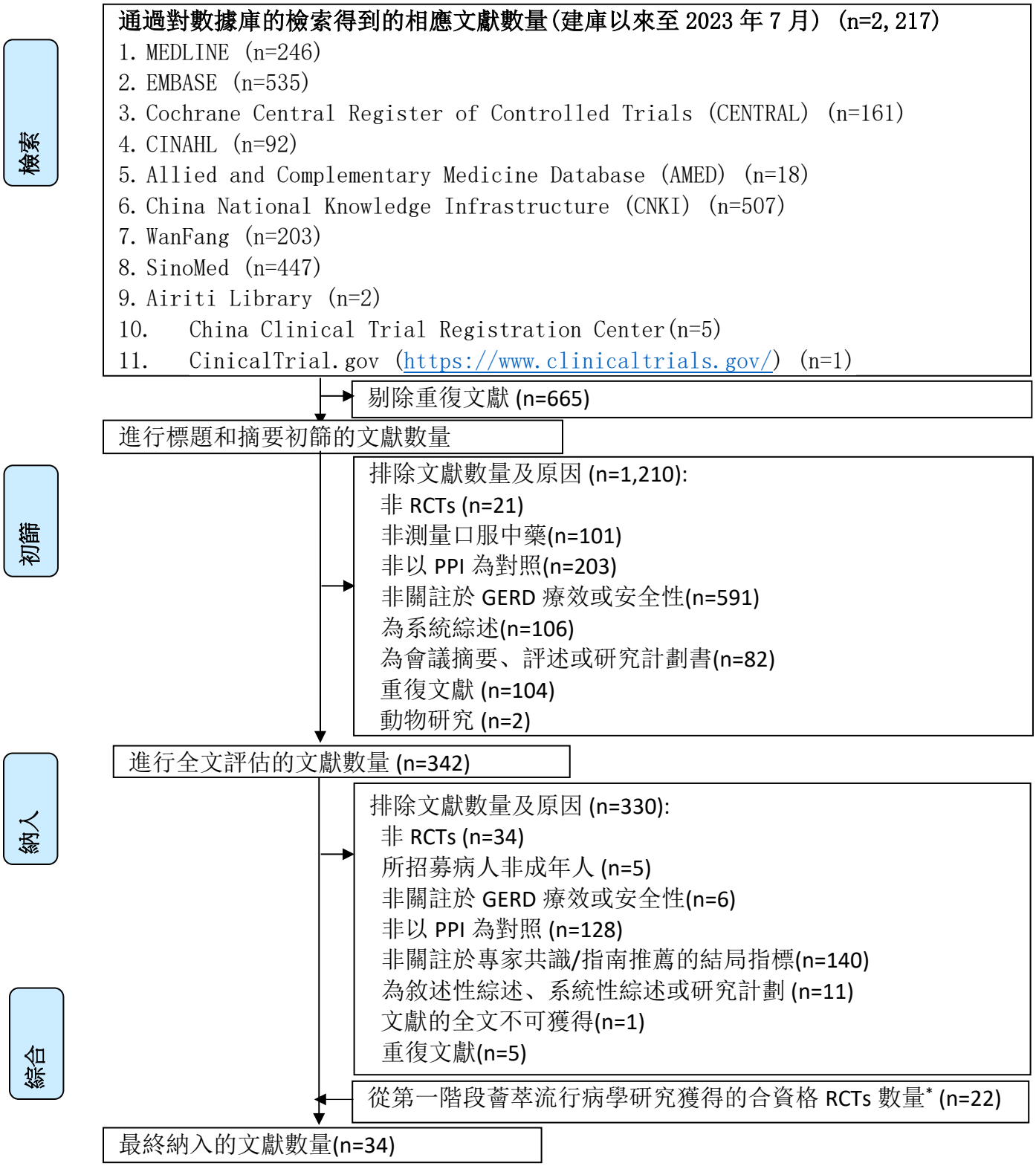
關於次要結局指標 GERDQ 評分變化，8 種中藥方及是否採用辨證論治原則被分為兩組(具體見表 7)。低質量證據表明，半夏瀉心湯(SD)、吳茱萸湯(NSD)、胃康一號方(SD)+PPI、半夏厚朴湯(SD)+PPI 以及茱苓瀉心湯(SD)+PPI 可能是對 GERDQ 評分變化最有效的治療方法。

對於復發率，3 種中藥方及是否採用辨證論治原則被分為兩組(具體見表 8)。低質量證據表明，和胃益氣降逆湯(NSD)+PPI 可能是降低 GERD 患者復發率最有效的治療方法。

在不良反應發生率方面，8 種中藥方及是否採用辨證論治原則被分為兩組(表 9)。低證據表明，半夏瀉心湯(NSD)、降逆和胃湯(NSD)以及和胃益氣降逆湯(NSD)+PPI 可能是治療 GERD 患者最安全的治療方法。

次要結局指標 RDQ 和 GERDQ 評分變化的網絡幾何圖、NMA 兩兩比較估計值、SUCRA 結果及其相關證據質量見附錄 14-17;復發率的網絡幾何圖、NMA 兩兩比較估計值、SUCRA 結果及其相關證據質量見附錄 18-21;不良反應發生率的網絡幾何圖、NMA 兩兩比較估計值、SUCRA 結果及其相關證據質量見附錄 22-25。

圖 3 文献检索和筛选的流程图



Notes: RCT: 隨機對照試驗, CHM: 中藥; GERD: 胃食管反流病, PPI: 質子泵抑製劑. * 在第一階段薈萃流行病學研究[1]中，共包含 48 個 GERD 隨機對照試驗,有 26 個不符合納入標准，原因為：i) 非以 PPI 為對照(n=15)，ii) 與已經確認合格的隨機對照試驗重複(n=11)。Reference: [1] Zhong CCW. (2023, Nov 10). The Impact of Syndrome Differentiation on Treatment Effects and Side Effects in Randomized Controlled Trials of Chinese Herbal Medicine: A Meta-Epidemiological Study [Poster presentation]. 2023 Taiwan Joint International Conference in Healthcare (TJICH), Taipei, Taiwan, China

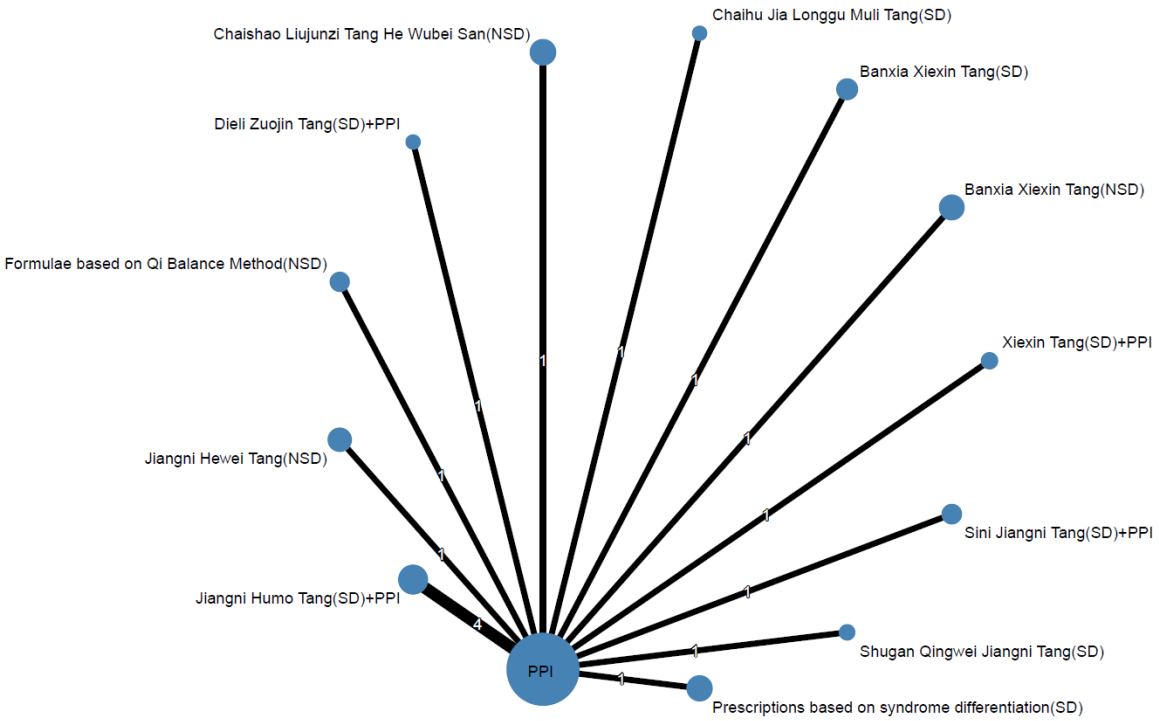


圖 4 採用或未採用辨證論治幹預模式的中藥和 PPI 比較組的網絡幾何圖，結局：癥狀緩解率

Notes: NSD, 未採用辨證論治; SD, 採用辨證論治; PPI, 質子泵抑製劑. 每一個點表示一種治療方案，隨機對照試驗中測量的治療方案通過一條線連接。每一個點的大小代表了隨機對照試驗中隨機分配到該治療方案的病人比例，而線條的粗細代表進行了此線條所連接兩種治療方案療效對照的 隨機對照試驗數。

Chaishao Liujunzi Tang He Wubei San(NSD)												
1.28 (1.03; 1.58)	**Sini Jiangni Tang(SD) +PPI**											
1.28 (1.02; 1.60)	1.00 (0.81; 1.24)	**Dieli Zuojin Tang(SD) +PPI**										
1.28 (0.99; 1.64)	1.00 (0.79; 1.27)	1.00 (0.78; 1.28)	**Xiexin Tang(SD) +PPI**									
1.31 (1.07; 1.60)	1.02 (0.85; 1.23)	1.02 (0.84; 1.25)	1.02 (0.81; 1.29)	**Jiangni Hewei Tang(NSD)**								
1.37 (1.07; 1.75)	1.07 (0.85; 1.35)	1.07 (0.84; 1.36)	1.07 (0.82; 1.40)	1.05 (0.84; 1.31)	**Chaihu Jia Longgu Muli Tang(SD)**							
1.41 (1.17; 1.70)	1.10 (0.93; 1.31)	1.10 (0.91; 1.32)	1.10 (0.89; 1.37)	1.08 (0.92; 1.26)	1.03 (0.83; 1.27)	**Jiangni Humo Tang(SD) +PPI**						
1.53 (1.13; 2.07)	1.20 (0.90; 1.60)	1.20 (0.89; 1.61)	1.20 (0.87; 1.65)	1.17 (0.88; 1.56)	1.12 (0.82; 1.54)	1.09 (0.83; 1.43)	**Shugan Qingwei Jiangni Tang(SD)**					
1.52 (1.28; 1.80)	**1.19 (1.02; 1.38)**	**1.18 (1.00; 1.40)**	1.19 (0.97; 1.45)	**1.16 (1.01; 1.33)**	1.11 (0.91; 1.35)	1.08 (0.96; 1.21)	0.99 (0.76; 1.28)	**Banxia Xiexin Tang(NSD)**				
1.53 (1.31; 1.80)	**1.20 (1.04; 1.38)**	**1.20 (1.02; 1.40)**	1.20 (0.99; 1.46)	**1.17 (1.04; 1.33)**	1.12 (0.93; 1.35)	1.09 (0.99; 1.20)	1.00 (0.77; 1.29)	1.01 (0.96; 1.07)	**PPI**			
1.55 (1.27; 1.90)	**1.22 (1.01; 1.46)**	**1.21 (1.00; 1.48)**	1.22 (0.97; 1.53)	**1.19 (1.00; 1.41)**	1.14 (0.91; 1.42)	1.10 (0.95; 1.29)	1.01 (0.77; 1.34)	1.03 (0.90; 1.17)	1.01 (0.90; 1.14)	**Formulae based on Qi Balance Method(NSD)**		
1.59 (1.30; 1.94)	**1.24 (1.03; 1.49)**	**1.24 (1.01; 1.51)**	1.24 (0.99; 1.56)	**1.21 (1.02; 1.44)**	1.16 (0.93; 1.45)	1.13 (0.96; 1.32)	1.03 (0.78; 1.37)	1.05 (0.91; 1.20)	1.03 (0.91; 1.17)	1.02 (0.86; 1.21)	**Banxia Xiexin Tang(SD)**	
1.73 (1.37; 2.19)	**1.35 (1.08; 1.69)**	**1.35 (1.07; 1.71)**	**1.35 (1.04; 1.76)**	**1.32 (1.07; 1.64)**	1.26 (0.98; 1.63)	**1.23 (1.01; 1.50)**	1.13 (0.83; 1.54)	1.14 (0.95; 1.37)	1.13 (0.95; 1.34)	1.11 (0.90; 1.37)	1.09 (0.88; 1.35)	**Prescriptions based on syndrome differentiation(SD)**

圖 5 不同的中藥方及是否採用辨證論治原則在 GERD 患者癥狀緩解率上兩兩比較的結果

Notes: NSD, 未採用辨證論治; SD, 採用辨證論治; PPI, 質子泵抑製劑. 每個格子中的數代表了縱列的治療方案相對於橫排的方案在癥狀緩解率上的相對效應值 RR 和 95%置信區間。RR 越大，說明縱列的治療方案在治療胃食管反流病的癥狀緩解率更高。有顯著性差異的數有加粗和下劃線標記。

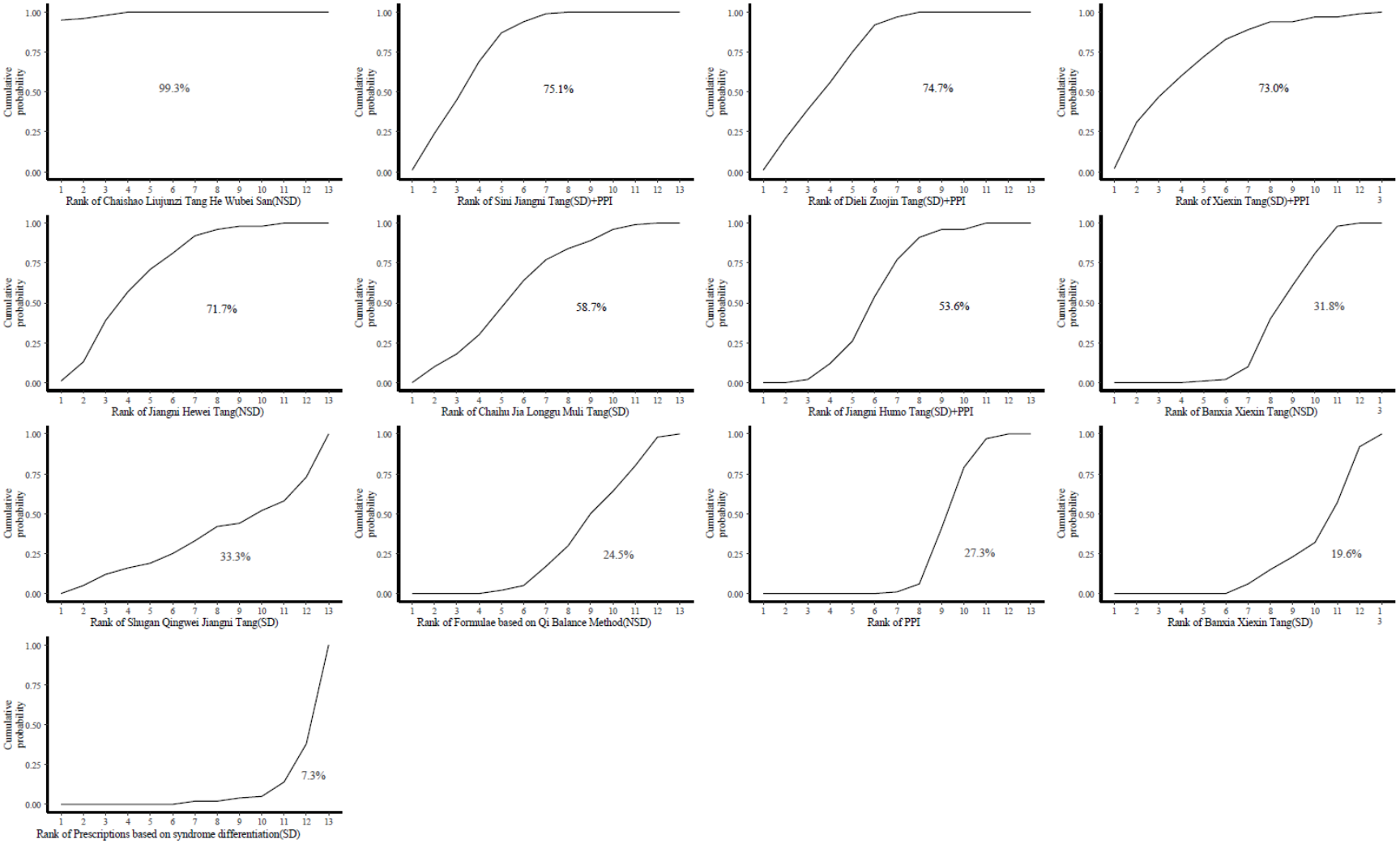


圖 6 不同的中藥方及是否採用辨證論治原則在 GERD 患者癥狀緩解率上的排序結果

Notes: NSD, 未採用辨證論治; SD, 採用辨證論治; PPI, 質子泵抑製劑. 橫軸表示每種治療方案在癥狀緩解率上可能的排名(從最佳排名到最差)。 縱軸表示每種治療方案成為最佳治療方案、第二好的治療方案、第三好的治療方案等的累積概率。 在這種圖形化方法中，通過檢查曲線下的面積來展示排名。 曲線下的面積越大，干預措施在癥狀緩解率上排名靠前或居於前幾名的可能性就越高。

表 5 根據 Minimally contextualized framework 綜合考慮療效、證據級別及 SUCRA 排名對中藥治療方案進行分類：癥狀緩解率

中藥治療方案分類	中藥治療方案	中藥治療方案 vs. PPI (RR (95% CI))	SUCRA	排序
低確定性 (低到非常低證據質量)				
組 2: 優於組 1 任一治療方案	柴芍六君子湯合烏貝散(NSD)	1.53 (1.31; 1.80)	99.3%	1
組 1：優於組 0 任一治療方案	四逆降逆湯(SD)+PPI	1.20(1.04;1.37)	75.1%	2
	蝶蠣左金湯(SD)+PPI	1.20(1.02;1.40)	74.7%	3
	瀉心湯(SD)+PPI	1.20(0.99;1.46)	73.0%	4
	降逆和胃湯(NSD)	1.17 (1.04; 1.33)	71.7%	5
組 0：與單獨使用 PPI 沒有顯著性差異的治療方案	柴胡加龍骨牡蠣湯(SD)	1.12(0.93;1.35)	58.7%	6
	降逆護膜湯(SD)+PPI	1.09(0.99;1.20)	53.6%	7
	疏肝清胃降逆湯(SD)	1.00(0.77;1.29)	33.3%	8
	半夏瀉心湯(NSD)	1.01 (0.96; 1.07)	31.8%	9
	基於升降並舉法的方劑(NSD)	0.99 (0.88; 1.11)	24.5%	11
	半夏瀉心湯(SD)	0.97(0.85;1.10)	19.6%	12
	自擬方(SD)	0.88(0.75;1.05)	7.30%	13

說明: CHM: 中藥; NSD: 未採用辨證論治; SD: 採用辨證論治; PPI: 質子泵抑製劑; RR:比值比; CI: 置信區間; SUCRA: 累積曲線下面積

表 6 根據 Minimally contextualized framework 綜合考慮療效、證據級別及 SUCRA 排名對中藥治療方案進行分類：RDQ 評分變化

治療方案分類	中藥治療方案	中藥治療方案 vs. PPI (WMD (95% CI))	SUCRA	排序
低確定性 (低到非常低證據質量)				
組 3: 優於組 2 任一治療方案	柴芍六君子湯合烏貝散(NSD)	-19.8(-21.7; -17.9)	100.0%	1
組 2: 優於組 1 任一治療方案	吳茱萸湯(NSD)	-6.95(-13.0; -0.88)	81.9%	2
	半夏厚樸湯(NSD)+PPI	-5.32(-7.33; -3.31)	80.3%	3
組 1: 優於組 0 任一治療方案	六君子湯(SD)+PPI	-5.29(-9.91; -0.67)	73.2%	4
	柴胡加龍骨牡蠣湯(SD)	-4.44(-7.25; -1.63)	68.5%	5
	調胃降逆湯(SD)+PPI	-3.77(-4.60; -2.95)	61.7%	6
	蝶蠣左金湯(SD)+PPI	-3.56(-5.64; -1.49)	56.3%	7
	旋覆代赭湯合六君子湯(SD)+PPI	-3.53(-6.66; -0.40)	55.2%	8
	降逆調胃湯(SD)	-3.36(-6.27; -0.45)	53.8%	9
	小柴胡湯(SD)	-3.06(-5.98; -0.14)	48.4%	10
	四逆降逆湯(SD)+PPI	-2.83(-4.29; -1.37)	44.4%	11
	瀉心湯(SD)+PPI	-2.61(-4.49; -0.73)	42.5%	12
	疏肝清胃降逆湯(SD)	-2.21(-4.42; 0.00)	35.8%	13
組 0：與單獨使用 PPI 沒有顯著性差異的治療方案	降逆護膜湯(SD)+PPI	-0.90(-2.89; 1.09)	19.8%	14
	柴胡梔子豉湯方(SD)	-0.90(-2.44; 0.64)	19.0%	15
	半夏瀉心湯(SD)	1.21(-0.17; 2.59)	0.70%	17

說明: CHM: 中藥; NSD: 未採用辨證論治; SD: 採用辨證論治; PPI: 質子泵抑製劑; WMD: 加權均數差; CI: 置信區間; SUCRA: 累積曲線下面積

表 7 根據 Minimally contextualized framework 綜合考慮療效、證據級別及 SUCRA 排名對中藥治療方案進行分類：GERDQ 評分變化

治療方案分類	中藥治療方案	中藥治療方案 vs. PPI (WMD (95% CI))	SUCRA	排序
低確定性 (低到非常低證據質量)				
組 1: 優於組 0 任一治療方案	半夏瀉心湯(SD)	-2.13(-3.29; -0.97)	82.0%	1
	吳茱萸湯(NSD)	-2.12(-3.92; -0.32)	78.0%	2
	胃康 1 號方(SD)+PPI	-1.86(-3.54; -0.18)	72.0%	3
	半夏厚樸湯(SD)+PPI	-1.19(-1.76; -0.62)	49.0%	5
	術苓瀉心湯(SD)+PPI	-1.18(-2.35; -0.01)	49.0%	6
組 0：與單獨使用 PPI 沒有顯著性差異的治療方案	蠲鬱降逆湯(SD)+PPI	-1.50(-3.02; 0.02)	60.0%	4
	柴半湯(SD)+PPI	-0.81(-2.36; 0.74)	37.0%	7
	小柴胡湯(NSD)	-0.43(-0.90; 0.04)	20.0%	8

說明: CHM: 中藥; NSD: 未採用辨證論治; SD: 採用辨證論治; PPI: 質子泵抑製劑; WMD: 加權均數差; CI: 置信區間; SUCRA: 累積曲線下面積

表 8 根據 Minimally contextualized framework 綜合考慮療效、證據級別及 SUCRA 排名對中藥治療方案進行分類：復發率

治療方案分類	中藥治療方案	中藥治療方案 vs. PPI (RR (95% CI))	SUCRA	排序
低確定性 (低到非常低證據質量)				
組 1: 優於組 0 任一治療方案	和胃益氣降逆湯(NSD)+PPI	0.20 (0.06; 0.66)	85.1%	1
組 0：與單獨使用 PPI 沒有顯著性差異的治療方案	吳茱萸湯聯合柴胡疏肝散(SD)+PPI	0.33(0.04; 3.03)	61.8%	2
	化肝煎合左金丸 (SD)+PPI	0.53(0.25; 1.12)	46.2%	3

說明: CHM: 中藥; NSD: 未採用辨證論治; SD: 採用辨證論治; PPI: 質子泵抑製劑; RR:比值比; CI: 置信區間; SUCRA: 累積曲線下面積

表 9 根據 Minimally contextualized framework 綜合考慮療效、證據級別及 SUCRA 排名對中藥治療方案進行分類：不良反應發生率

治療方案分類	中藥治療方案	中藥治療方案 vs. PPI (RR (95% CI))	SUCRA	排序
低確定性 (低到非常低證據質量)				
組 1: 優於組 0 任一治療方案	半夏瀉心湯(NSD)	0.02 (0.00; 0.38)	87.8%	1
	降逆和胃湯(NSD)	0.05 (0.00; 0.80)	79.8%	2
	和胃益氣降逆湯(NSD)+PPI	0.11 (0.01; 0.85)	66.0%	3
組 0：與單獨使用 PPI 沒有顯著性差異的治療方案	半夏瀉心湯(SD)	0.14(0.02; 1.09)	61.9%	4
	下氣湯(SD)+PPI	0.33(0.04; 3.13)	43.4%	5
	旋覆代赭湯合六君子湯(SD)+PPI	0.33(0.01; 7.69)	41.9%	6
	柴胡加龍骨牡蠣湯(SD)	0.50(0.05; 5.22)	33.7%	7
	吳茱萸湯(SD)+PPI	0.72(0.47; 1.10)	25.1%	8

說明: CHM: 中藥; NSD: 未採用辨證論治; SD: 採用辨證論治; PPI: 質子泵抑製劑; RR:比值比; CI: 置信區間; SUCRA: 累積曲線下面積

發表偏倚評估

對於主要指標癥狀緩解率和次要指標 RDQ 評分變化，兩者所涵蓋的 RCTs 數量均超過10個，因此具備評估發表偏倚的條件。兩者的漏斗圖均未呈現出不對稱性。Egger回歸檢測結果進一步支持了上述結果，結果顯示這兩個指標均未檢測到發表偏倚(癥狀緩解率的 Egger 回歸檢測 p 值為 0.148，RDQ 評分變化的 Egger 回歸檢測 p 值為 0.650)(附錄26-27)。

表 10 兩種重要方劑的中藥材

方劑	中藥材
柴芍六君子湯合烏貝散	Chaihu (BUPLEURI RADIX, 柴胡) 12g, Baishao (PAEONIAE RADIX ALBA, 白芍) 12g, Dangshen (CODONOPSIS RADI, 党参) 12g, Fuling (PORIA, 茯苓) 10g, Baizhu (ATRACTYLODIS MACROCEPHALAE RHIZOMA, 白术) 12g, Fabanxia (PINELLIAE RHIZOMA PRAEPARATUM, 法半夏) 7g, Chenpi (CITRI RETICULATAE PERICARPIUM, 陈皮) 10g, Haipiaoxiao (SEPIAE ENDOCONCHA, 海螵蛸) 20g, Walengzi (ARCAE CONCHA, 瓦楞子) 15g, Zhebeimu (FRITILLARIAE THUNBERGII BULBUS, 浙贝母) 12g, Huangqin (SCUTELLARIAE RADIX, 黄芩) 12g, Huanglian (COPTIDIS RHIZOMA, 黄连) 6g, Zisugeng (PERILLAE CAULIS, 紫苏梗) 10g, Gancao (GLYCYRRHIZAE RADIX ET RHIZOMA, 甘草) 6g
和胃益氣降逆湯	Xuanfuhua (INULAE FLOS, 旋覆花) 10 g, Dangshen (CODONOPSIS RADIX, 党参) 10 g, Zheshi (HAEMATITUM, 赭石) 15 g, Houpo (MAGNOLIAE OFFICINALIS CORTEX, 厚朴) 9 g, Zhishi (AURANTII FRUCTUS IMMATURUS, 枳实) 9 g, Shidi (KAKI CALYX，柿蒂) 9 g, Zhigancao (GLYCYRRHIZAE RADIX ET RHIZOMA PRAEPARATA CUM MELLE, 炙甘草) 6 g, Fabanxia (PINELLIAE RHIZOMA PRAEPARATUM, 法半夏) 9 g, Shengjiang (ZINGIBERIS RHIZOMA RECENS, 生姜) 9g, Dazao (JUJUBAE FRUCTUS, 大枣) 6 顆

-討論

關於辨證論治對中藥療效和安全性的影響，本項目整體層面薈萃流行病學研究和單個疾病層面NMA的結論相一致，即根據辨證論治進行中藥治療不一定能改善治療結果。儘管本研究表明，無論是否使用辨證論治，中藥獲得的治療結果相似，但開展 RCT 時所選定的具體中藥方劑可能會受到證型的影響。NMA 結果顯示在療效和安全性方面表現最佳的中藥方劑分別是柴芍六君子湯合烏貝散(NSD)與和胃益氣降逆湯(NSD)。這些方劑的主要功效包括疏肝解郁、健脾和胃、降逆止嘔和補氣益氣 (這兩個方劑具體包含的成分見表 10)。巧合的是，這些功效正好針對 GERD 最常見的中醫證型——肝胃不和證和脾胃氣虛伴胃氣上逆證。這種巧合可能解釋了為什麼這兩種中藥方劑的表現優於

其他方劑。雖然研究這兩種中藥有效性/安全性的 RCTs 沒有考慮患者的中醫辨證分型，但如果 RCT 招募的大多數患者被診斷為肝胃不和證或脾胃氣虛伴胃氣上逆證，那麼根據中醫理論，他们的证型与这兩種中藥方劑是适配的。在這種情況下，中藥方劑治療已經隱含地考慮了相關證型。所以我們研究的這兩種類型的 RCT 結果之間沒有差異是合理的。從這一點上看，如果疾病本身具有較高概率被診斷分類為特定的一種中醫證型，那麼對該病進行中醫治療時，辨證論治所發揮的作用可能較小。因為此時直接對患者使用針對該病該證型的固定方劑，而不進行細緻的辨證論治，也有很高的概率得到方劑證型相匹配的結果。

這可能解釋了為什麼之前的研究有中醫師認為辨证论治在臨床實踐中有時不太重要[31]，同時也是本項目第二階段德爾菲研究部分參與者所表達的觀點。實際上，在功能性消化不良[32]這一臨床問題上也能觀察到類似的情況，即功能性消化不良患者常見的中醫證型與 NMA 研究結果證實表現最佳的中藥方劑相互匹配。比如， 一項基於隱樹分析的中醫診斷研究發現，肝氣犯胃在功能性消化不良患者中的發病率較高。而在另一項針對功能性消化不良的NMA研究中，逍遙丸這一方劑表現最佳。根據中醫辨證論治，肝氣犯胃證與逍遙散，二者之間存在方證匹配的理論聯繫[33]。

-影響

與薈萃流行病學研究結果一致，NMA 結果同樣表明辨證論治在治療 GERD 方面並未提供額外的益處。然而，支援這一結論的證據質量並不高，且一些德爾菲參與者認為，基於臨床經驗，辨證論治仍然具有重要意義。因此，通過實證研究進一步評估中醫辨證論治在實踐中的重要性可能是必要的。將中醫證型如肝胃不和證或脾胃氣虛伴胃氣上逆證應用於指導中藥處方以提高療效的概念，類似於在精準腫瘤治療中使用預測性生物標誌物[34]。在我們的 NMA 中，單個 RCT(如被診斷為肝胃不和證或脾胃氣虛伴胃氣上逆證)可能根據治療結果(如癥狀減輕)將 GERD 患者分層，以便進行個性化治療(即柴芍六君子湯合烏貝散或和胃益氣降逆湯)[35]。

未來研究可以進一步評估肝胃不和證或脾胃氣虛伴胃氣上逆證在預測 GERD 患者治療效果方面的潛力[36]。研究者可以考慮採用傘式試驗設計[37]，首先對 GERD 患者進行中醫診斷並將其分到不同的亞組。其次，在各證型的亞組中，隨機分配兩種治療方案，

一種是針對其中醫診斷開具的中藥方劑，另一種是NMA研究結果所確定的最佳方劑。如果所有證型亞組中，兩組中藥方劑之間的治療結果沒有顯著差異，那麼可以表明辨證論治的價值較低。這時針對 GERD 使用 NMA 證實的最佳方劑就足夠了，同時也驗證了我們的觀察結果。在進行上述研究之前，應使用隱樹分析將患者的診斷操作標準化。

-結論

第一階段的薈萃流行病學研究結果表明，辨證論治和非辨證論治干預模式對於中藥隨機對照試驗的治療效果和安全性沒有顯著差異。從第二階段德爾菲研究的參與者反饋來看，我們觀察到了關於進一步研究辨證論治影響的不同觀點，這可能源於對中醫實踐基本信念的差異。一些參與者強調了辨證論治在中醫實踐中的關鍵作用，認為它對於實現臨床有效性至關重要，因此不認為有必要進行 NMA。另一方面，一些參與者則表達了對辨證論治在指導中藥處方中的必要性的不確定性或不同意見，主張進行額外的 NMA 分析。GERD 在第二階段德爾菲研究中達成專家共識並成為第三階段 NMA 研究的焦點。第三階段 NMA 研究結果顯示，治療 GERD 時未採用辨證論治的中藥方劑治療在所有測量指標上普遍最為有效，除了 GERDQ 評分變化之外。儘管這些發現與我們之前的薈萃流行病學研究結論一致，但仍需要開展實證研究進一步評估。鑒於 GERD 中肝胃不和證或脾胃氣虛伴胃氣上逆證等中醫診斷的普遍性，評估它們對治療結果的預測價值，對於評估辨證論治在指導 GERD 個體化中藥方劑治療中的作用至關重要。

參考文獻

[1] Griffiths S.M., Chung V.C., & Tang J.L., 2010. Integrating traditional Chinese medicine: experiences from China. Aust. J. Med. (Online). 3(7):385-396.
[2] Chung V.C., Ma P.H., Lau C.H., Wong S.Y., Yeoh E.K., & Griffiths S.M., 2014. Views on traditional Chinese medicine amongst Chinese population: a systematic review of qualitative and quantitative studies. Health Expect. 17(5):622-636.
[3] World Health Organization (WHO). International Classification of Diseases, Eleventh Revision (ICD-11), World Health Organization (WHO) 2019/2021. https://icd.who.int/browse11 (accessed 27 March 2024).
[4] World Health Organization (WHO). WHO international standard terminologies on traditional medicine in the western pacific region. 2007.
[5] Jiang, M., Lu, C., Zhang, C., Yang, J., Tan, Y., Lu, A., & Chan, K., 2012. Syndrome differentiation in modern research of traditional Chinese medicine. J. Ethnopharmacol. 140(3):634-642. https://doi.org/10.1016/j.jep.2012.01.033

[6] Ho, L., Xu, Y., Zhang, N. L., Ho, F. F., Wu, I. X. Y., Chen, S., Liu, X., Wong, C. H. L., Ching, J. Y. L., Cheong, P. K., Yeung, W. F., Wu, J. C. Y., & Chung, V. C. H., 2022. Quantification of prevalence, clinical characteristics, co-existence, and geographic variations of traditional Chinese medicine diagnostic patterns via latent tree analysis-based differentiation rules among functional dyspepsia patients. Chin. Med. 17(1):101.
[7] Higgins J.P.T., Thomas J., Chandler J., Cumpston M., Li T., Page M.J., Welch V.A. (editors), 2023. Cochrane Handbook for Systematic Reviews of Interventions version 6.4 (updated August 2023). Cochrane
[8] Kabisch M., Ruckes C., Seibert-Grafe M., Blettner M., 2011. Randomized controlled trials: part 17 of a series on evaluation of scientific publications. Dtsch Arztebl Int. 108(39):663-668.
[9] Cheng, C. W., Wu, T. X., Shang, H. C., Li, Y. P., Altman, D. G., Moher, D., Bian, Z. X., & CONSORT-CHM Formulas 2017 Group, 2017. CONSORT extension for Chinese herbal medicine formulas 2017: recommendations, explanation, and elaboration (traditional Chinese version). Ann. Intern. Med. 167(2):W7-W20.
[10] Dong J., 2013. The relationship between traditional Chinese medicine and modern medicine. eCAM. 2013, 153148.
[11] Zhang, X., Tian, R., Zhao, C., Birch, S., Lee, J.A., Alraek, T., Bovey, M., Zaslawski, C., Robinson, N., Kim, T.H., Lee, M.S., Bian, Z., 2019. The use of pattern differentiation in WHO-registered traditional Chinese medicine trials – A systematic review. Eur. J. Integr. Med. 30, 100945.
[12] Chung, V.C., Ho, R.S., Wu, X., Wu, J.C., 2016. Incorporating traditional Chinese medicine syndrome differentiation in randomized trials: methodological issues. Eur. J. Integr. Med. 8 (6), 898-904.
[13] Chung, V.C., Ho, L.T., Wu, I.X., 2020. Chinese medicine diagnosis and treatment for COVID-2019: Is China ready for implementing a national guideline? Adv Integr Med. 7 (2), 51-54.
[14] Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (editors), 2022. Cochrane Handbook for Systematic Reviews of Interventions version 6.3 Cochrane.
[15] Haring, R., Ghannad, M., Bertizzolo, L., Page, M.J., 2020. No evidence found for an association between trial characteristics and treatment effects in randomized trials of testosterone therapy in men: a meta-epidemiological study. J. Clin. Epidemiol. 122, 12-19.
[16] Higgins, J.P., Altman, D.G., Gøtzsche, P.C., Jüni, P., Moher, D., Oxman, A.D., Savovic, J., Schulz, K.F., Weeks, L., Sterne, J.A.C., 2011. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. Bmj 343, d5928.
[17] Sterne, J.A., Jüni, P., Schulz, K.F., Altman, D.G., Bartlett, C., Egger, M., 2002. Statistical methods for assessing the influence of study characteristics on treatment effects in 'meta-epidemiological' research. Stat Med. 21 (11), 1513-1524.
[18] Schünemann, H. J., Wiercioch, W., Brozek, J., Etxeandia-Ikobaltzeta, I., Mustafa, R. A., Manja, V., Brignardello-Petersen, R., Neumann, I., Falavigna, M., Alhazzani, W., Santesso, N., Zhang, Y., Meerpohl, J. J., Morgan, R. L., Rochwerg, B., Darzi, A., Rojas, M. X., Carrasco-Labra, A., Adi, Y., AlRayees, Z., … Akl, E. A, 2017. GRADE Evidence to Decision (EtD) frameworks for adoption, adaptation, and de novo development of trustworthy recommendations: GRADE-ADOLOPMENT. J. Clin. Epidemiol. 81:101-110. https://doi.org/10.1016/j.jclinepi.2016.09.009
[19] Holey E.A., Feeley J.L., Dixon J., Whittaker V.J., 2007. An exploration of the use of simple statistics to measure consensus and stability in Delphi studies. BMC Med. Res. Methodol., 7(1): 1-10. http://doi:10.1186/1471-2288-7-52

[20] Hasson F., Keeney S., McKenna H., 2000. Research guidelines for the Delphi survey technique. J Adv. Nurs. 32(4):1008-1015. https://doi.org/10.1046/j.1365-2648.2000.t01-1-01567.x
[21] Zhong C.C.W. (2023, Nov 10). The Impact of Syndrome Differentiation on Treatment Effects and Side Effects in Randomized Controlled Trials of Chinese Herbal Medicine: A Meta-Epidemiological Study [Poster presentation]. Taiwan Joint International Conference in Healthcare (TJICH), Taipei, Taiwan, China
[22] Katz PO, Gerson LB, Vela MF; null. Guidelines for the diagnosis and management of gastroesophageal reflux disease. Am J Gastroenterol. 2013 Mar;108(3):308-28; quiz 329.
[23] Slater, B. J., Collings, A., Dirks, R., Gould, J. C., Qureshi, A. P., Juza, R., Rodríguez-Luna, M. R., Wunker, C., Kohn, G. P., Kothari, S., Carslon, E., Worrell, S., Abou-Setta, A. M., Ansari, M. T., Athanasiadis, D. I., Daly, S., Dimou, F., Haskins, I. N., Hong, J., Krishnan, K., … Marks, J. M., 2023. Multi-society consensus conference and guideline on the treatment of gastroesophageal reflux disease (GERD). Surg. Endosc. 37(2), 781–806. doi:10.1007/s00464-022-09817-3
[24] Min Y.W., Shin Y.W., Cheon G.J., Park K.S., Kim H.S., Sohn C.I., Kim T.N., Moon H.C., Rhee P.L., 2016. Recurrence and Its Impact on the Health-related Quality of Life in Patients with Gastroesophageal Reflux Disease: A Prospective Follow-up Analysis. J Neurogastroenterol Motil. 22(1):86-93. doi:10.5056/jnm15124
[25] Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H. Y., Corbett, M. S., Eldridge, S. M., Emberson, J. R., Hernán, M. A., Hopewell, S., Hróbjartsson, A., Junqueira, D. R., Jüni, P., Kirkham, J. J., Lasserson, T., Li, T., McAleenan, A., … Higgins, J. P. T., 2019. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 366:l4898. doi: 10.1136/bmj.l4898
[26] Puhan, M. A., Schünemann, H. J., Murad, M. H., Li, T., Brignardello-Petersen, R., Singh, J. A., Kessels, A. G., Guyatt, G. H., & GRADE Working Group, 2014. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. BMJ 349:g5630. https://doi.org/10.1136/bmj.g5630
[27] Salanti G., Ades A.E., Ioannidis J.P., 2011. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. J Clin. Epidemiol. 64(2):163–171. https://doi.org/10.1016/j.jclinepi.2010.03.016
[28] Brignardello-Petersen, R., Izcovich, A., Rochwerg, B., Florez, I. D., Hazlewood, G., Alhazanni, W., ... & Schünemann, H. J. (2020). GRADE approach to drawing conclusions from a network meta-analysis using a partially contextualised framework. *bmj*, *371*.
[29] Sterne, J. A., & Egger, M., 2001. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. J Clin. Epidemiol. 54(10):1046-1055. https://doi.org/10.1016/s0895-4356(01)00377-8
[30] Page, M. J., Higgins, J. P., & Sterne, J. A., 2019. Assessing risk of bias due to missing results in a synthesis. Cochrane handbook for systematic reviews of interventions, 349-374.
[31] Chung, V. C. H., Ho, L. T. F., & Wu, I. X. Y., 2020. Chinese medicine diagnosis and treatment for COVID-2019: Is China ready for implementing a national guideline? Adv. Integr. Med. 7(2), 51–54. https://doi.org/10.1016/j.aimed.2020.04.001
[32] Ho, L., Xu, Y., Zhang, N. L., Ho, F. F., Wu, I. X. Y., Chen, S., Liu, X., Wong, C. H. L., Ching, J. Y. L., Cheong, P. K., Yeung, W. F., Wu, J. C. Y., & Chung, V. C. H., 2022. Quantification of prevalence, clinical characteristics, co-existence, and geographic variations of traditional Chinese medicine diagnostic patterns via latent tree analysis-based differentiation rules among functional dyspepsia patients. Chin. Med. 17(1), 101. https://doi.org/10.1186/s13020-022-00656-x
[33] Ho, L., Zhong, C. C., Wong, C. H., Wu, J. C., Chan, K. K., Wu, I. X., Leung, T. H., & Chung, V. C., 2022. Herbal medicine for functional dyspepsia: Network meta-analysis of

placebo-controlled randomised trials. J. Ethnopharmacol. 283, 114665. https://doi.org/10.1016/j.jep.2021.114665
[34] Parker, J. L., Kuzulugil, S. S., Pereverzev, K., Mac, S., Lopes, G., Shah, Z., Weerasinghe, A., Rubinger, D., Falconi, A., Bener, A., Caglayan, B., Tangri, R., & Mitsakakis, N., 2021. Does biomarker use in oncology improve clinical trial failure risk? A large-scale analysis. Cancer med. 10(6), 1955–1963. https://doi.org/10.1002/cam4.3732
[35] Menis, J., Hasan, B., & Besse, B., 2014. New clinical research strategies in thoracic oncology: clinical trial design, adaptive, basket and umbrella trials, new end-points and new evaluations of response. Eur. Respir. Rev. 23(133), 367–378. https://doi.org/10.1183/09059180.00004214
[36] Mandrekar, S. J., & Sargent, D. J., 2009. Clinical trial designs for predictive biomarker validation: theoretical considerations and practical challenges. J. Clin. Oncol. 27(24), 4027–4034. https://doi.org/10.1200/JCO.2009.22.3701
[37] Le-Rademacher, J., Dahlberg, S., Lee, J. J., Adjei, A. A., & Mandrekar, S. J., 2018. Biomarker Clinical Trials in Lung Cancer: Design, Logistics, Challenges, and Practical Considerations. J. Thorac. Oncol. 13(11), 1625–1637. https://doi.org/10.1016/j.jtho.2018.08.2019

刊物(如有)

1. 本項目第一階段研究結果已撰寫成文章，題為“Impact of syndrome differentiation on treatment effects and side effects of Chinese herbal medicine in randomized controlled trials: A meta-epidemiological study”，並已提交至國際期刊《Medicinal Research Reviews》，目前審稿中。
2. 本項目第二階段及第三階段研究結果已撰寫成文章，題為“The impact of syndrome differentiation on the effectiveness and safety of Chinese herbal medicine treatments for digestive disorders: Expert consensus and network meta-analysis ，並已提交至國際期刊《Journal of Traditional and Complementary Medicine》，目前審稿中。

專利和其他知識產權(如有)

無

致謝(如有)

感謝中醫藥發展基金對本項目的支持。

感謝所有德爾菲參與者，感謝他們參與本項目第二階段德爾菲研究，為選擇一個重要的臨床問題開展第三階段 NMA 研究提供了寶貴意見。

附錄

附錄 1 薈萃流行病學分析(兩步法)

第一步:在薈萃分析水平對療效/安全性進行評估。對於二分類指標，使用比值比(OR)測量療效大小；對於連續性指標，使用標準化均差(SMD)測量療效大小。通過薈萃-回歸分析方法估算中藥 RCT 辨證與否與療效/安全性之間的關係。對於二分類指標和連續性指標，分別採用比值比之比(ROR)和標準化均差(dSMD)來測量辨證論治與非辨證論治之間的療效/安全性差異。ROR>1 或 dSMD>0 表明相比採用非辨證論治的中藥 RCT，採用辨證論治的中藥 RCT 會得到更大的療效或更高的不良反應。

第二步:採用隨機效應模型，合併第一步中各個薈萃分析的 ROR 或 dSMD。P<0.05 表明具有統計學差異。本研究採用 I^2 和 τ^2評估薈萃分析之間的異質性。 I^2<25%，25%-50%，和>50%分別表示低，中及高度異質性。

附錄 2 基於薈萃流行病學分析結果總結辨證論治對中藥治療 12 種消化系統疾病幹預效應影響的證據概要表

疾病類型	降低證據質量級別的因素					證據質量級別*	辨證與非辨證 RCT 中藥療效差異		
	研究設計局限	不一致性	不直接性	不精確性	發表偏倚		實施辨證論治的 RCT 數目和樣本量	採用非辨證論治的 RCT 數目和樣本量	ROR (95%CI)
非糜爛性胃食管反流病	不降	不降級	不降級	降級 ΔΔ	未檢測	⊕◯◯◯非常低	11(993)	2(250)	0.87(0.34,2.19)
未特指的胃食管反流病	不降級	不降級	不降級	降級Δ	未檢測	⊕◯◯◯非常低	23(2000)	12(1552)	1.26(0.67,2.38)
幽門螺桿菌相關性胃炎	不降級	不降級	不降級	不降級^	未檢測	⊕⊕◯◯低	27(2340)	22(2119)	1.23(0.79,1.91)
慢性萎縮性胃炎	降級#	不降級	不降級	降級 ΔΔ	未檢測	⊕◯◯◯非常低	16(1669)	19(1526)	0.87(0.43,1.78)
淺表性胃炎	不降級	不降級	不降級	降級 ΔΔ	未檢測	⊕◯◯◯非常低	2(240)	10(946)	0.93(0.37,2.38)
消化性潰疡	不降級	不降級	不降級	降級 ΔΔ	未檢測	⊕◯◯◯非常低	7(740)	15(1542)	1.01(0.51,1.97)
克羅恩病	不降級	不降級	不降級	降級 ΔΔ	未檢測	⊕◯◯◯非常低	4(364)	8(541)	0.99(0.37,2.62)
潰瘍性結腸炎	不降級	不降級	不降級	降級Δ	未檢測	⊕◯◯◯非常低	21(1716)	25(2488)	1.11(0.73,1.69)
功能性消化不良	不降級	不降級	不降級	降級Δ	未檢測	⊕◯◯◯非常低	16(1226)	8(954)	0.73(0.41,1.28)
腹瀉型腸易激綜合征	不降級	降級&	不降級	降級 ΔΔ	未檢測	⊕◯◯◯非常低	22(1785)	9(1214)	1.57(0.63,3.92)
老年功能性便秘	降級#	不降級	不降級	降級 ΔΔ	未檢測	⊕◯◯◯非常低	6(478)	13(1167)	0.77(0.37,1.63)
幽門螺桿菌感染	不降級	不降級	不降級	降級 ΔΔ	未檢測	⊕◯◯◯非常低	4(435)	12(1620)	1.19(0.64,2.21)

Notes: ROR, ratio of odds ratio,比值比之比，反映的是實施辨證論治的 RCT 和採用非辨證論治的 RCT 中藥治療效果大小的差異;CI：置信區間。 ROR>1 說明實施辨證論治比採用非辨證論治能夠產生更大的治療效果。 * ROR 數據來源於薈萃流行病學分析，本質類似於二手數據分析，可視為橫斷面研究，因此證據品質評級從低質量開始，根據降低證據質量級別的因素，對存在嚴重問題的領域降低一個等級;對嚴重程度更高的降 2 個等級。

證據質量因嚴重的研究設計局限性而降級，因為所有納入的系統綜述的方法學品質都被評為 “極低” 。

& 證據品質因嚴重不一致性而被降級，表現在異質性很大，I2 值為 59%。

Δ 證據品質因較寬的 95%CI 導致嚴重不精確而降級。 ROR 點估計與 1.0 重疊，區間的上限或下限跨過了顯著益處或損害點（估計值低於 0.667，高於 1.5 被認為是臨床顯著的）。

ΔΔ 證據品質因寬的 95%CI 導致嚴重不精確而降級。 點估計與 1.0 重疊，區間的上限和下限均跨過了顯著的益處或損害點（估計值低於 0.667，高於 1.5 被認為是臨床上顯著的）。

^不降級，因為總樣本量大於 4000，足以實現基線特徵方面的預後平衡。

由於納入的文章小於 10 篇，發表偏倚未檢測。

GRADE 證據品質評級解讀

質量級別低：效果估計值的可信度有限。 真正的效果可能與效果估計值有較大差異。 質量級別非常低：效果估計值的可信度很低。 真正的效果可能與效果估計值有極大差異。

附錄 3 檢索中藥治療胃食管反流病的隨機對照試驗的策略和結果

(1) MEDLINE from 1946 to July, 2023

#	Searches	Results
1	randomized controlled trial.pt.	596399
2	randomized controlled trial.mp.	623062
3	1 or 2	623062
4	exp Drugs, Chinese Herbal/	52200
5	Chinese herb*.mp.	56888
6	exp Medicine, Chinese Traditional/	23852
7	Traditional Chinese medic*.mp.	27407
8	exp Phytotherapy/	42085
9	phytother*.mp.	42409
10	(chinese adj5 (traditional or medic*)).mp.	58029
11	(herbs or herbal).mp.	88492
12	(plant or plants).mp.	869407
13	(traditional adj5 medic*).mp.	77721
14	4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13	974325
15	exp esophagus/	54293
16	esophag$.tw.	141754
17	oesophag$.tw.	34294
18	exp gastroesophageal reflux/	29278
19	(gastroesophageal adj3 reflux).tw.	18852
20	(gastro adj3 oesophageal adj3 reflux).tw.	4020
21	(gastro adj3 esophageal adj3 reflux).tw.	1562
22	(gord or gerd).tw.	9452

23	(gastroesophageal reflux or gastro-oesophageal reflux or gastro oesophageal reflux or gastro-esophageal reflux or gastro esophageal reflux or GORD or GERD or gastric regurgitation or gastric acid reflux or esophageal reflux).mp.	35784
24	exp duodenogastric reflux/ or exp bile reflux/ or exp dyspepsia/ or exp eructation/ or exp heartburn/ or exp esophagitis/	26103
25	(duodenogastric adj2 reflux).tw.	873
26	(bile adj2 reflux).tw.	1130
27	(acid adj3 reflux).tw.	3078
28	dyspep$.tw.	13093
29	(belch$ or burp$).tw.	1336
30	eructation.tw.	228
31	(heartburn or indigestion).tw.	5893
32	esophagitis.tw.	13344
33	oesophagitis.tw.	3044
34	15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33	217694
35	3 and 14 and 34	250
36	limit 35 to humans	246

(2) Embase from 1910 to July, 2023

#	Search	Result
1	double-blind:.mp.	298481
2	placebo:.tw.	364587
3	blind:.tw.	500521
4	1 or 2 or 3	721205
5	exp Chinese medicine/	71964

6	exp oriental medicine/	2793
7	exp herbaceous agent/	60452
8	exp medicinal plant/	294125
9	exp Chinese herb/	5363
10	Chinese medic*.mp.	91845
11	oriental medic*.mp.	4284
12	herbaceous agent.mp.	60454
13	medicinal plant*.mp.	112793
14	Chinese herb*.mp.	21140
15	herb*.mp.	232790
16	5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15	533216
17	exp esophagus/ or exp gastroesophageal reflux/ or exp duodenogastric reflux/ or exp bile reflux/ or exp dyspepsia/ or exp eructation/ or exp heartburn/ or exp esophagitis/	207643
18	esophag$.tw.	226407
19	oesophag$.tw.	51789
20	(gastroesophageal adj3 reflux).tw.	32634
21	(gastro adj3 oesophageal adj3 reflux).tw.	5833
22	(gastro adj3 esophageal adj3 reflux).tw.	3600
23	gord.tw.	1686
24	gerd.tw.	21456
25	(gastroesophageal reflux or gastro-oesophageal reflux or gastro oesophageal reflux or gastro-esophageal reflux or gastro esophageal reflux or GORD or GERD or gastric regurgitation or gastric acid reflux or esophageal reflux).mp.	76699
26	(duodenogastric adj3 reflux).tw.	1053
27	(bile adj3 reflux).tw.	1910
28	(acid adj3 reflux).tw.	5989
29	dyspep$.tw.	22908
30	(belch$ or burp$).tw.	2630

31	eructation.tw.	360
32	(heartburn or indigestion).tw.	11829
33	esophagitis.tw.	26290
34	oesophagitis.tw.	4716
35	17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34	383446
36	4 and 16 and 35	553
37	limit 36 to human	535

(3) Cochrane Central Register of Controlled Trials (CENTRAL) (from database inception to July, 2023)

ID	Search	Hits
#1	MeSH descriptor: [Gastroesophageal Reflux] explode all trees	2566
#2	MeSH descriptor: [Drugs, Chinese Herbal] explode all trees	4185
#3	“Chinese herb*” or “herb*” or “Traditional Chinese medic*” or “phytother*” or “(chinese adj5 (traditional or medic*))” or “(plant or plants)” or “(traditional adj5 medic*)” or “Chinese medic*” or “oriental medic*” or “herbaceous agent” or “medicinal plant*”	3061
#4	gastroesophageal reflux or gastro-oesophageal reflux or gastro oesophageal reflux or gastro-esophageal reflux or gastro esophageal reflux or GORD or GERD or gastric acid reflux or esophageal reflux	6194
#5	duodenogastric reflux or bile reflux or dyspepsia or eructation or heartburn or esophagitis	11837
#6	(duodenogastric adj2 reflux)	6
#7	(bile adj2 reflux)	13
#8	(acid adj3 reflux)	50
#9	(dyspep$)	9
#10	(belch$ or burp$)	347
#11	(eructation)	363
#12	(heartburn or indigestion)	3217
#13	(esophagitis)	3627
#14	(oesophagitis)	3627
#15	#4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14	16181
#16	#1 or #15	16234
#17	#2 or #3	6652
#18	#16 and #17 in Trials	161

(4) CINAHL Ultimate (from database inception to July, 2023)

#	Query	Results
S1	SU randomized controlled trial OR controlled clinical trial OR randomized OR randomised OR placebo OR clinical trials OR randomly OR trial OR clinical trial	326,272
S2	SU “Chinese herb*” or “herb*” or “Traditional Chinese medic*” or “phytother*” or “(chinese adj5 (traditional or medic*))” or “(plant or plants)” or “(traditional adj5 medic*)” or “Chinese medic*” or “oriental medic*” or “herbaceous agent” or “medicinal plant*”	26,878
S3	SU gastroesophageal reflux or gastro-oesophageal reflux or gastro oesophageal reflux or gastro-esophageal reflux or gastro esophageal reflux or GORD or GERD or gastric regurgitation or gastric acid reflux or esophageal reflux	8,088
S4	TX duodenogastric adj2 reflux or bile adj2 reflux or acid adj3 reflux or dyspep$ or belch$ or burp$ or eructation or heartburn or indigestion or esophagitis or oesophagitis	19,954
S5	TX duodenogastric reflux or bile reflux or dyspepsia or eructation or heartburn or esophagitis	21,722
S6	S3 OR S4 OR S5	33,265
S7	S1 AND S2 AND S6	92

(5) AMED (Allied and Complementary Medicine) from 1985 to July 2023

Searches	Results	Result
1	double-blind:.mp.	2781
2	placebo:.tw.	4212
3	blind:.tw.	6539
4	1 or 2 or 3	8713
5	Chinese medicine.tw.	4697
6	oriental medicine.tw.	199
7	herbaceous agent.tw.	0
8	medicinal plant.tw.	1367
9	Chinese herb.tw.	210
10	Chinese medic*.tw.	5307
11	oriental medic*.tw.	239
12	herbaceous agent.tw.	0
13	medicinal plant*.tw.	3127
14	Chinese herb*.tw.	6264
15	herb*.tw.	18316
16	5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15	23332
17	(esophagus or gastroesophageal reflux or duodenogastric reflux or bile reflux or dyspepsia or eructation or heartburn or esophagitis).tw.	502
18	esophag$.tw.	563
19	oesophag$.tw.	151
20	(gastroesophageal adj3 reflux).tw.	125
21	(gastro adj3 oesophageal adj3 reflux).tw.	15
22	(gastro adj3 esophageal adj3 reflux).tw.	2
23	gord.tw.	5

24	gerd.tw.	43
25	(gastroesophageal reflux or gastro-oesophageal reflux or gastro oesophageal reflux or gastro-esophageal reflux or gastro esophageal reflux or GORD or GERD or gastric regurgitation or gastric acid reflux or esophageal reflux).tw.	138
26	(duodenogastric adj3 reflux).tw.	0
27	(bile adj3 reflux).tw.	1
28	(acid adj3 reflux).tw.	13
29	dyspep$.tw.	221
30	(belch$ or burp$).tw.	26
31	eructation.tw.	4
32	(heartburn or indigestion).tw.	105
33	esophagitis.tw.	24
34	oesophagitis.tw.	5
35	17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34	1021
36	4 and 16 and 35	18

(6) China National Knowledge Infrastructure (CNKI) from data inception to July 2023

((((((((((题名= '中医') OR (题名= '中医学')) OR (题名= '中药')) OR (题名= '草药')) OR (题名= '中草药')) OR (题名= '中成药')) OR (题名= '中医药')) OR (题名= '方剂')) OR (题名= '中西医')) OR (题名= '汤')) AND ((((((((题名= '临床试验') OR (题名= '对照试验')) OR (题名= '随机')) OR (题名= '临床观察')) OR (题名= '临床科研')) OR (题名= 'RCT')) OR (题名= '随机对照试验')) OR (题名= '临床研究')) OR (题名= '临床实验')) AND ((题名='食管反流') OR (题名= '反流性食管')) , 资源范围:总库；中英文扩展；更新时间:不限，yielded **502** citations.

(7) WanFang from data inception to July 2023

题名:(临床试验 or 对照试验 or 随机 or 临床观察 or 临床科研 or RCT or 随机对照试验 or 临床研究 or 临床实验) and 题名:(中医 or 中医学 or 中药 or 草药 or 中草药 or 中成药 or 中医药 or 方剂 or 中西医 or 汤) and 题名:(食管反流 or 反流性食管), 主题词扩展 ,yielded 203 Citations.

(8) SinoMed from data inception to July 2023

("食管反流"[标题:智能] OR "反流性食管"[标题:智能]) AND ("中医"[标题:智能] OR "中医学"[标题:智能] OR "中药"[标题:智能] OR "草药"[标题:智能] OR "中草药"[标题:智能] OR "中成药"[标题:智能] OR "中医药"[标题:智能] OR "方剂"[标题:智能] OR "中西医"[标题:智能] OR "汤"[标题:智能]) AND ("临床试验"[标题:智能] OR "对照试验"[标题:智能] OR "随机"[标题:智能] OR "临床观察"[标题:智能] OR "临床科研"[标题:智能] OR "RCT"[标题:智能] OR "随机对照试验"[标题:智能] OR "临床研究"[标题:智能] OR "临床实验"[标题:智能]), yielded **447** citations.

(9) Airiti Library from data inception to July 2023

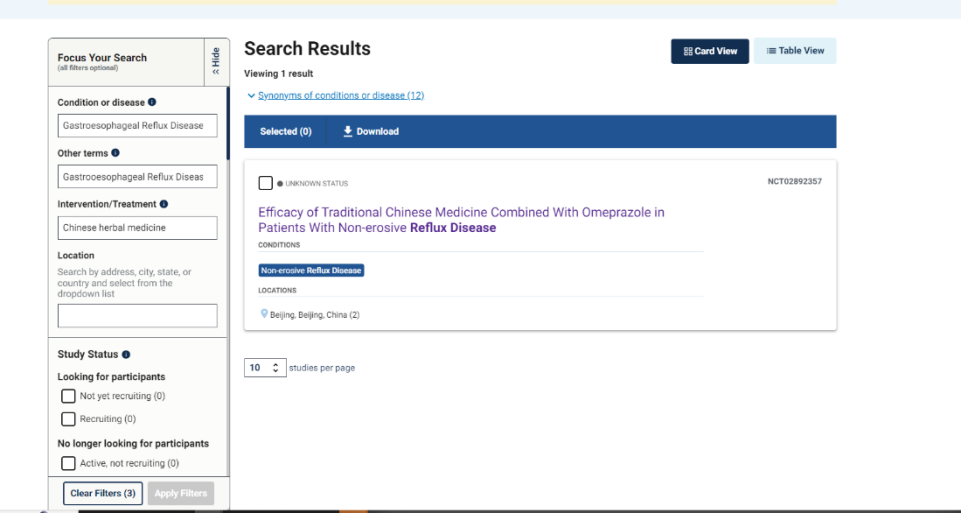
(((((((([DN]:(临床试验) OR [DN]:(对照试验)) OR [DN]:(随机)) OR [DN]:(临床观察)) OR [DN]:(临床科研)) OR [DN]:(RCT)) OR [DN]:(随机对照试验)) OR [DN]:(临床研究)) OR [DN]:(临床实验)) AND ((((((((([DN]:(中医) OR [DN]:(中医学)) OR [DN]:(中药)) OR [DN]:(草药)) OR [DN]:(中草药)) OR [DN]:(中成药)) OR [DN]:(中医药)) OR [DN]:(方剂)) OR [DN]:(中西医)) OR [DN]:(汤)) AND ([DN]:(食管反流) OR [DN]:(反流性食管)), yielded **2** citations.

(10) China Clinical Trial Registration Center (**http://www.chictr.org.cn)：** 5

(11) National center for Biotechnology Information (https://www.clinicaltrials.gov/): 1

Condition: Gastroesophageal Reflux Disease\(GERD\)

Intervention: Chinese Herbal Medicine

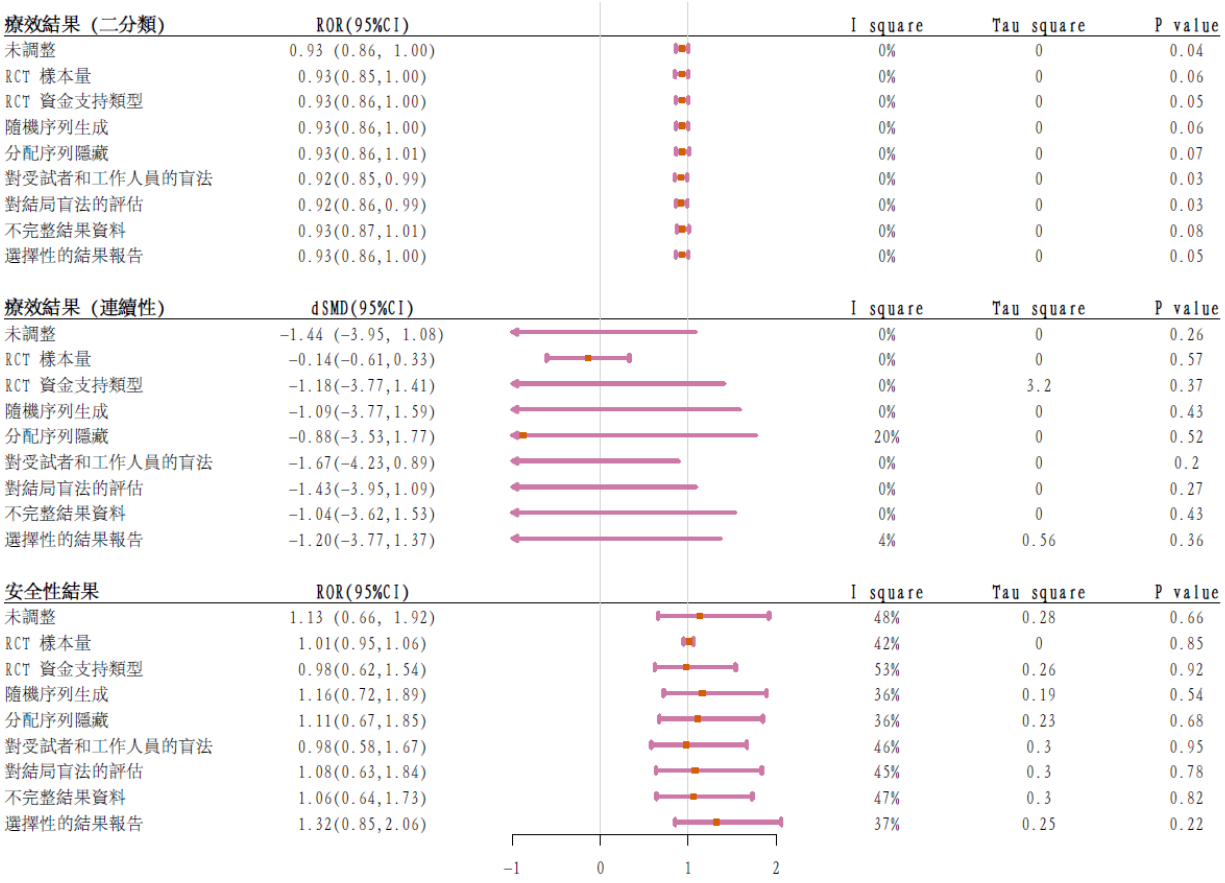


附錄 4 所納入的 2,064 項隨機對照試驗的偏倚風險*

偏倚風險維度及風險等級	RCT 匯總 (N=2,064)	採用辨證論治的 RCT (N=1,015)	採用非辨證論治的 RCT (N=1,049)	*P*#
隨機序列生成				
低偏倚風險	910 (44.1)	365 (36.0)	545 (52.0)	<0.001
偏倚風險不確定	1,049 (50.8)	597 (58.8)	452 (43.1)	
高偏倚風險	105 (5.1)	53 (5.2)	52 (5.0)	
分配序列隱藏				
低偏倚風險	34 (1.6)	12 (1.2)	22 (2.1)	0.257
偏倚風險不確定	1,925 (93.3)	950 (93.6)	975 (92.9)	
高偏倚風險	105 (5.1)	53 (5.2)	52 (5.0)	
對受試者和工作人員的盲法				
低偏倚風險	59 (2.9)	30 (3.0)	29 (2.8)	0.664
偏倚風險不確定	1,996 (96.7)	982 (96.7)	1,014 (96.7)	
高偏倚風險	9 (0.4)	3 (0.3)	6 (0.6)	
對結局盲法的評估				
低偏倚風險	455 (22.0)	191 (18.8)	264 (25.2)	<0.001
偏倚風險不確定	1,605 (77.8)	823 (81.1)	782 (74.5)	
高偏倚風險	4 (0.2)	1 (0.1)	3 (0.3)	
不完整結局數據				
低偏倚風險	1,918 (92.9)	967 (95.3)	951 (90.7)	<0.001
偏倚風險不確定	9 (0.4)	4 (0.4)	5 (0.5)	
高偏倚風險	137 (6.6)	44 (4.3)	93 (8.9)	
選擇性的結果報告				
低偏倚風險	0(0)	0(0)	0(0)	0.658
偏倚風險不確定	1,989 (96.4)	980 (96.6)	1,009 (96.2)	
高偏倚風險	75 (3.6)	35 (3.4)	40 (3.8)	

RCT，隨機對照試驗

*除非特殊說明，表格中的數字代表頻數(百分比)；#*P* 值為分組調整卡方檢驗的結果



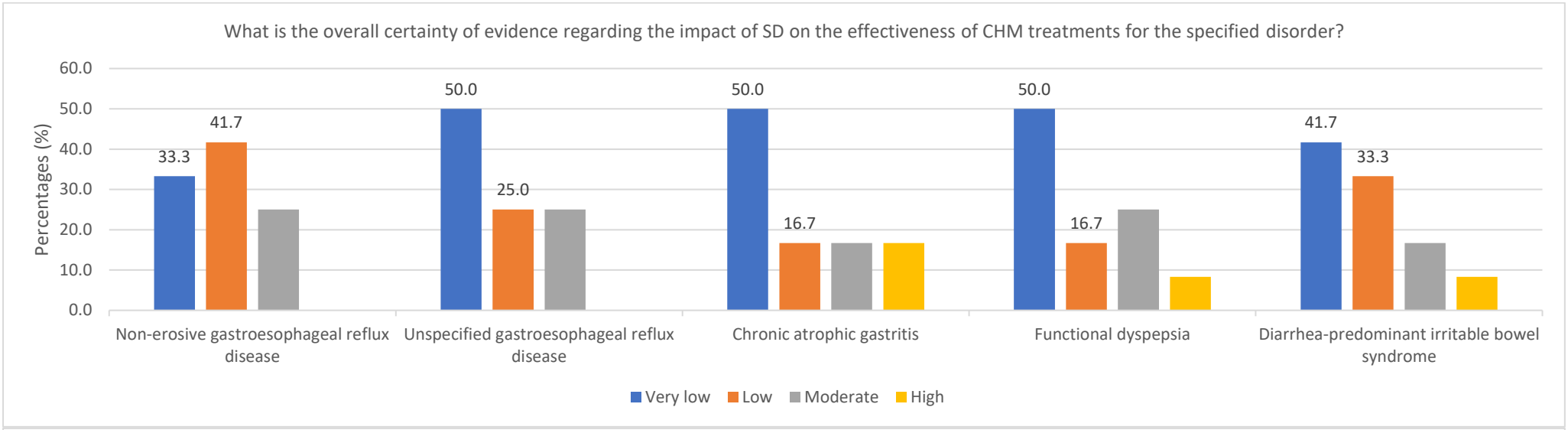
附錄 5 敏感性分析(調整了 RCT 的樣本量，資金支持類型和偏倚風險)

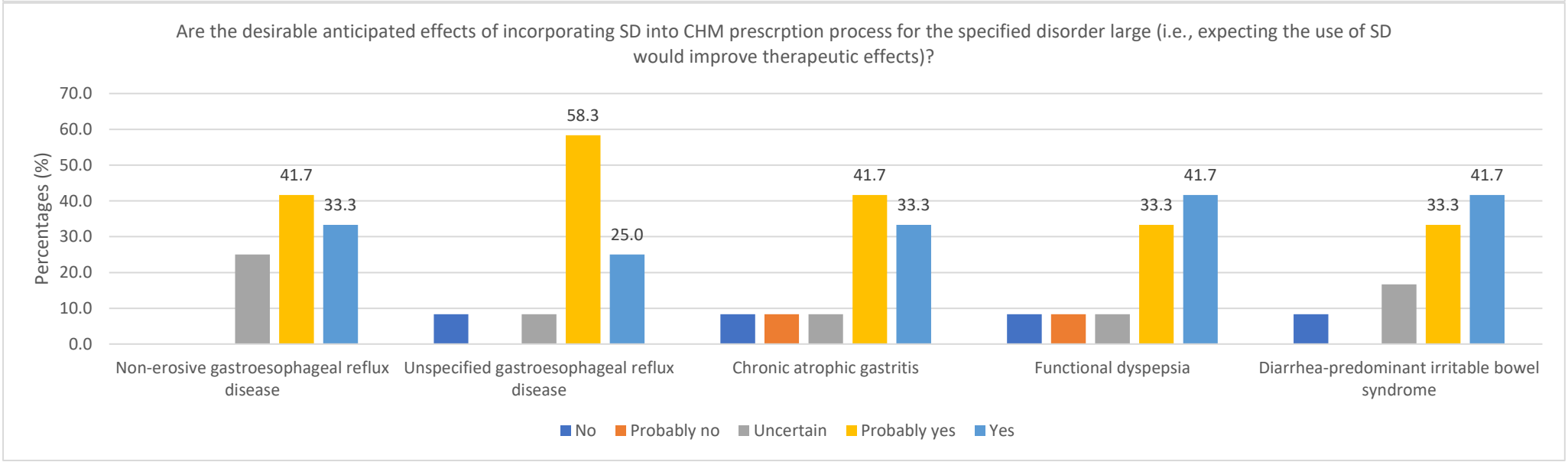
說明:ROR，比值比之比；RCT，隨機對照試驗；CI，置信區間；SR，系統綜述；dSMD，標準化均差之差。對於療效結局，當結局數據為二分類指標時，比值比之比小於 1 表明相對於採用辨證論治的 RCT，採用非辨證論治的 RCT 的中藥療效更大；當結局數據為連續性指標時，標準化均差之差小於 0 表明相對於採用辨證論治的 RCT，採用非辨證論治的 RCT 的中藥療效更大。對於安全性結局，比值比之比大於 1 表明相對於採用非辨證論治的 RCT，採用辨證論治的 RCT 的中藥不良反應風險更高。

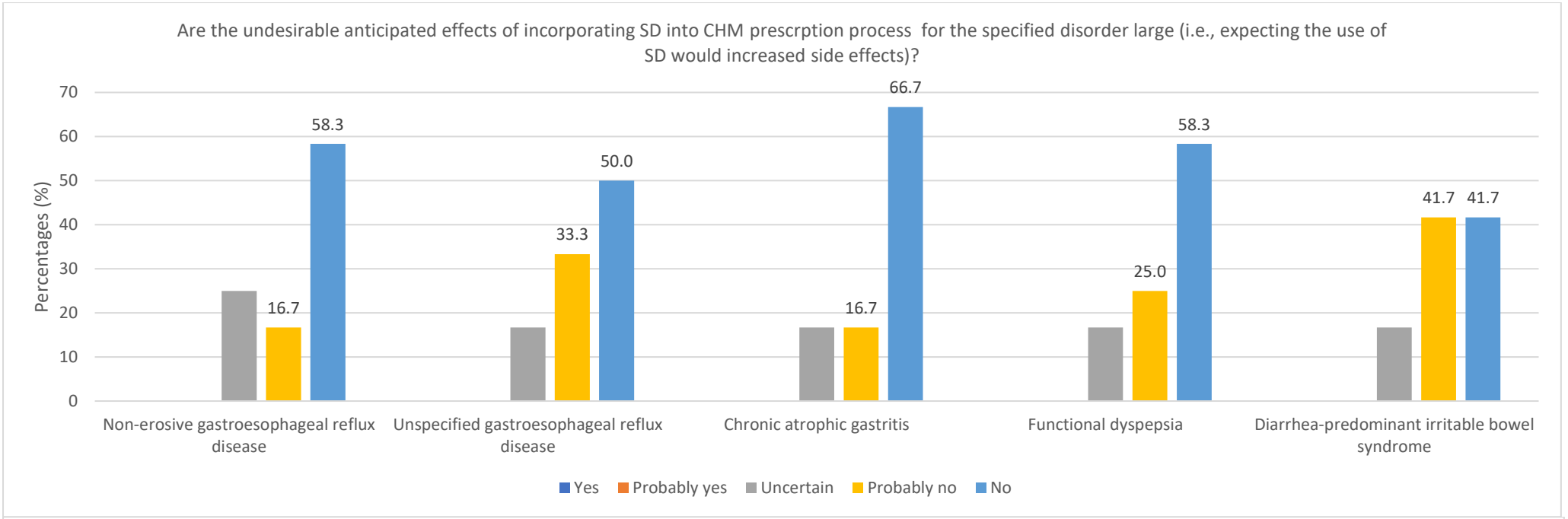
附錄 6 第一部分德爾菲研究：參與者人口學特征 (N=12)

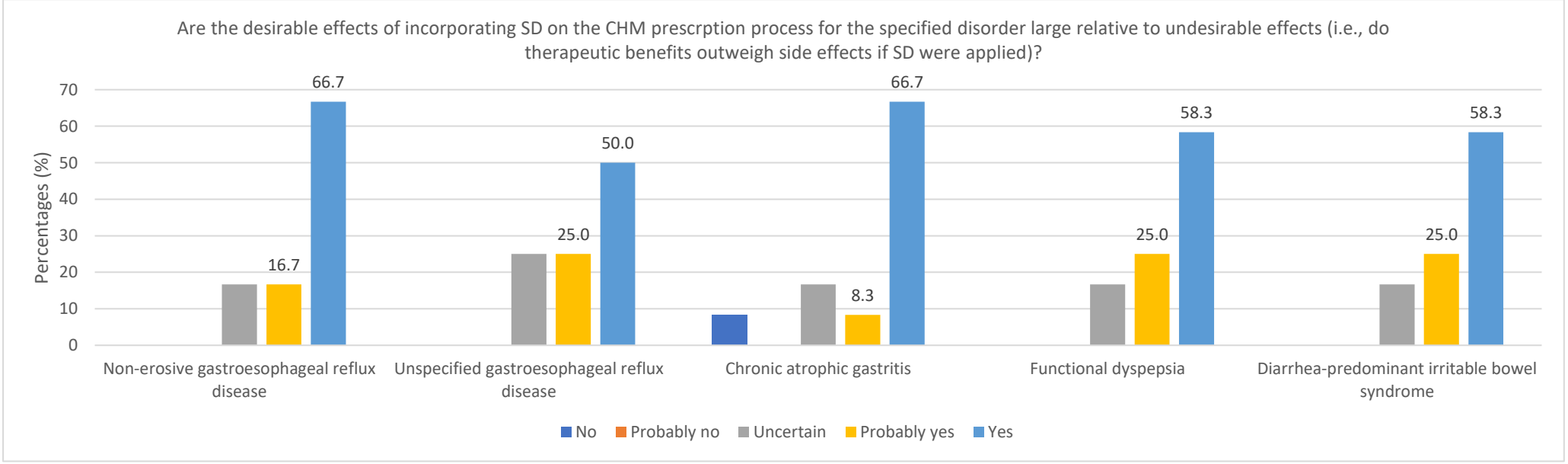
人口統計學特徵		n(%)
性別	男	5 (41.7)
	女	7 (58.3)
年齡	31-45	9 (75.0)
	46-60	3 (25.0)
教育水平	碩士學位	4 (33.3)
	博士學位	8 (66.7)
職業	中醫師	6 (50.0)
	接受過中醫培訓的西醫	4 (33.3)
	臨床研究方法學家	2 (16.7)
臨床/研究經驗		
	10-15 年	7 (58.3)
	16-20 年	2 (16.7)
	21 年以上	3 (25.0)

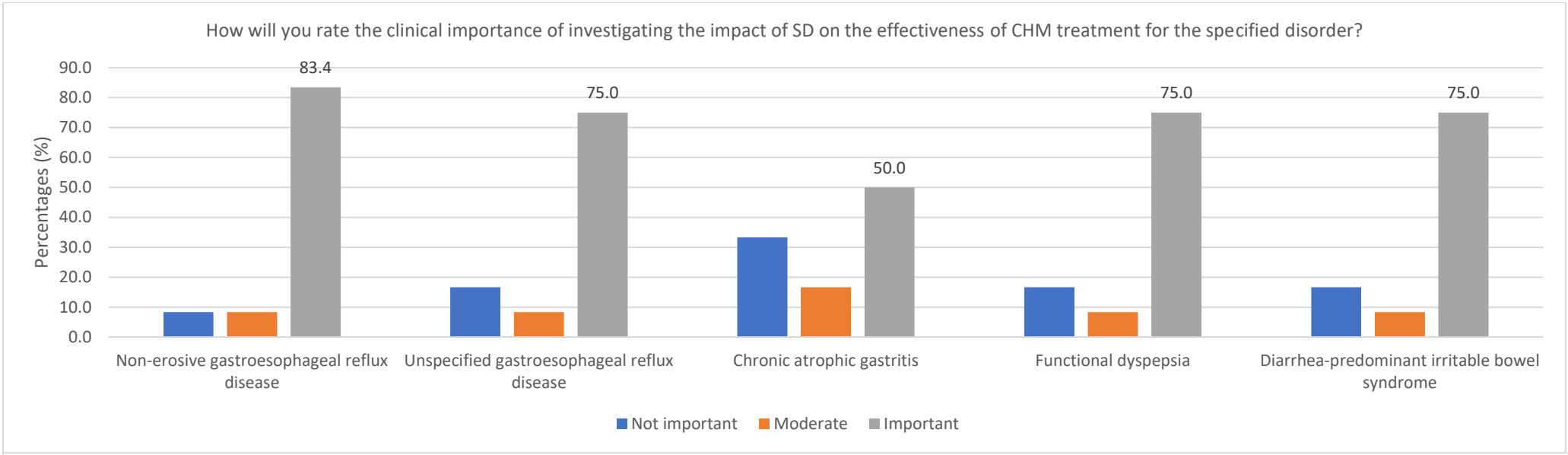
附錄 7 專家小組基於德爾菲問卷各條目對這些達成共識的臨床問題的評價結果

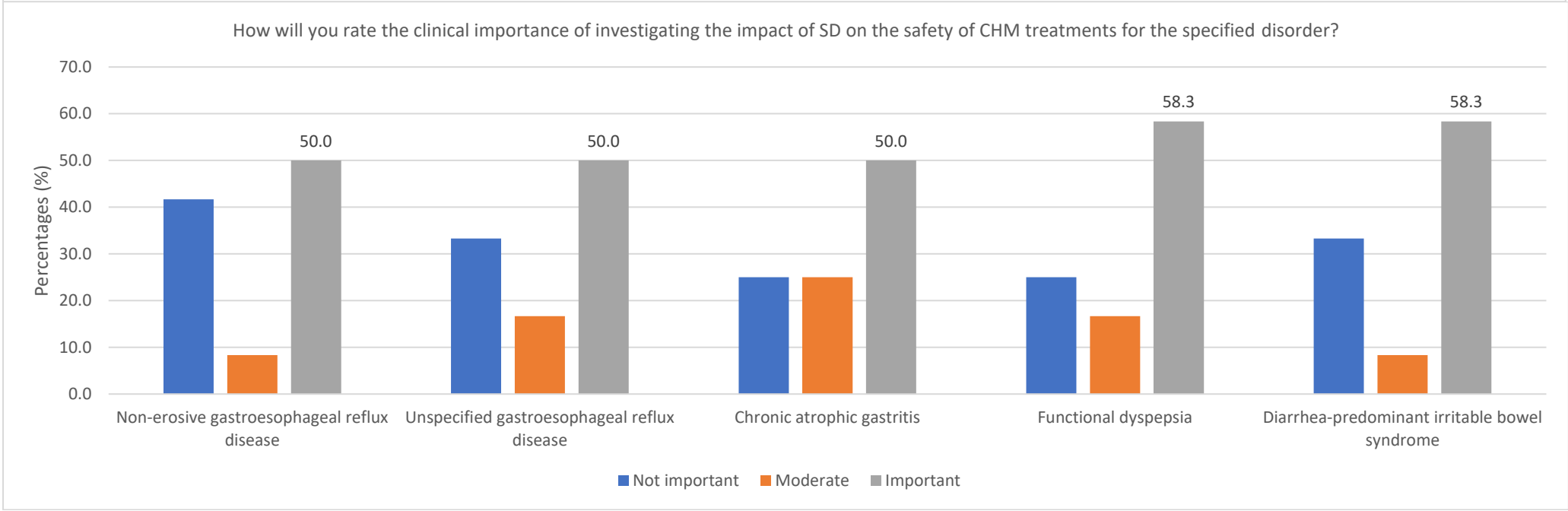





Are the undesirable anticipated effects of incorporating SD into CHM prescrption process for the specified disorder large (i.e., expecting the use of SD would increased side effects)?
Percentages (%)
70
60
50
40
30
20
10
0
58.3
16.7
50.0
33.3
66.7
16.7
58.3
25.0
41.7
41.7
Non-erosive gastroesophageal reflux disease
Unspecified gastroesophageal reflux disease
Chronic atrophic gastritis
Functional dyspepsia
Diarrhea-predominant irritable bowel syndrome
Yes
Probably yes
Uncertain
Probably no
No


Are the desirable effects of incorporating SD on the CHM prescrption process for the specified disorder large relative to undesirable effects (i.e., do therapeutic benefits outweigh side effects if SD were applied)?
Percentages (%)
70
60
50
40
30
20
10
0
66.7
16.7
50.0
25.0
66.7
8.3
58.3
25.0
58.3
25.0
Non-erosive gastroesophageal reflux disease
Unspecified gastroesophageal reflux disease
Chronic atrophic gastritis
Functional dyspepsia
Diarrhea-predominant irritable bowel syndrome
No
Probably no
Uncertain
Probably yes
Yes


How will you rate the clinical importance of investigating the impact of SD on the effectiveness of CHM treatment for the specified disorder?
Percentages (%)
90.0
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
83.4
75.0
50.0
75.0
75.0
Non-erosive gastroesophageal reflux disease
Unspecified gastroesophageal reflux disease
Chronic atrophic gastritis
Functional dyspepsia
Diarrhea-predominant irritable bowel syndrome
Not important
Moderate
Important


How will you rate the clinical importance of investigating the impact of SD on the safety of CHM treatments for the specified disorder?
Percentages (%)
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
50.0
50.0
50.0
58.3
58.3
Non-erosive gastroesophageal reflux disease
Unspecified gastroesophageal reflux disease
Chronic atrophic gastritis
Functional dyspepsia
Diarrhea-predominant irritable bowel syndrome
Not important
Moderate
Important

附錄 8 兩輪德爾菲問卷調查參與者的定性意見

a) 針對 5 種達成共識的臨床問題

Five endorsed disorders	Reasons for recommending/supporting	Reasons for not recommending/not supporting
1. Non-erosive gastroesophageal reflux disease 2. Unspecified gastresophageal reflux disease 3. Chronic atrophic gastritis 4. Functional dyspepsia 5. Irritable bowel syndrome with diarrhea	For all the five endorsed disorders: • Network meta-analysis helps quantify and rank the effects of different interventions, aiding in the selection of the best treatment approach for these disorders. (003) • It is necessary to conduct network meta-analysis as the findings can serve as guidance for clinical practice. In cases where no significant difference in efficacy exists between syndrome differentiation and non-syndrome differentiation, opting for non-syndrome differentiation treatment can enhance overall efficiency. (004) • Determining the necessity of syndrome differentiation treatment in traditional Chinese medicine is crucial. It's a key factor in the international acceptance and expansion of traditional Chinese medicine.(009)	For all the five endorsed disorders: • It's somewhat necessary to conduct a network meta-analysis, though not overly crucial. Some specific formulas have a clinical basis and align partially with syndrome differentiation principles. For instance, many formulas target liver-stomach disharmony, commonly seen in gastroesophageal reflux disease. (002). • In irritable bowel syndrome with diarrhea, the pathogenesis of liver depression and spleen deficiency is relatively prominent. Regardless of whether syndrome differentiation treatment is adopted, the basic approach is similar. (002) • There is no need for further network meta-analyses, as syndrome differentiation treatment is fundamental to traditional Chinese medicine and cannot be overlooked. Neglecting the understanding of disease mechanisms and the properties of herbs, and solely treating symptoms, can lead to the occurrence of deficiency or excess problems. (005) • The meta-epidemiological study suggests minimal disparities in treatment outcomes for all the disorders, thus, further comparisons may not be necessary. (012)

b) 針對 7 種未達成共識的臨床問題

Seven disorders that did not reach consensus	Reasons for recommending/supporting	Reasons for not recommending/not supporting
1. Helicobacter pylori induced gastritis 2. Superficial gastritis 3. Peptic ulcer disease 4. Crohn disease 5. Ulcerative colitis 6. Functional constipation in the elderly 7. Helicobacter pylori infection	For all the seven disorders that did not reach consensus: • Network meta-analysis helps quantify and rank the effects of different interventions, aiding in the selection of the best treatment approach for these disorders (003). • It is necessary to conduct a network meta-analysis as the findings can serve as guidance for clinical practice. In cases where no significant difference in efficacy exists between syndrome differentiation and non-syndrome differentiation, opting for non-syndrome differentiation treatment can enhance overall efficiency.(004)	• Helicobacter pylori induced gastritis is an infectious disease, and clinically, the impact of syndrome differentiation treatment is minimal. (009) • In Peptic ulcer disease, regardless of incorporating syndrome differentiation or not, the efficacy is not as good as conventional medicine, so there is no need for further comparison. (002) • In cases of superficial gastritis, Crohn's disease, ulcerative colitis, and helicobacter pylori infection, the associated syndromes are complex, and achieving effective treatment outcomes necessitates syndrome differentiation. Therefore, additional analysis is deemed unnecessary. (002) • Ulcerative colitis predominantly manifests as syndromes of spleen deficiency, damp-heat, and blood stasis. Therefore, Chinese herbal medicine formulas targeting these syndromes can effectively address the majority of patients in the acute phase. Consequently, there is not a strong necessity for further research. (002) • The meta-epidemiological study suggests minimal disparities in treatment outcomes for all the disorders, thus, further comparisons may not be necessary. (012)

附錄 9 納入的 34 個隨機對照試驗

1. 張冀. 降逆護膜湯治療胃食管反流病的臨床研究 [碩士]： 南京中醫藥大學; 2012.
2. 丁婷婷. 降逆護膜湯對反流性食管炎抑酸抗復發的實驗及臨床研究 [碩士]： 南京中醫藥大學; 2012.
3. 陳康遠， 劉暉， 曾韻萍， 駱潔恆. 半夏瀉心湯加味治療反流性食管炎 90 例臨床觀察. 新中醫. 2009（5）.
4. 張翔茹. 中醫辨證用藥治療胃食管反流病隨機對照臨床觀察 [碩士]： 北京中醫藥大學; 2013.
5. 周文真. 降逆調胃湯加減治療肝胃鬱熱型胃食管反流病的臨床觀察 [碩士]： 江西中醫藥大學; 2021.
6. 陳藝. 蝶螵左金湯治療肝胃鬱熱型反流性食管炎的臨床研究 [碩士]： 山東中醫藥大學; 2022a.
7. 李世琪， 鄭澤宇， 孫慧. 柴半湯治療肝胃不和型反流性食管炎的臨床研究. 中國民族醫藥雜誌. 2023; 29(05):14-5+34.
8.Shih Y-S， Tsai C-H， Li T-C， Yu C-J， Chou J-W， Feng C-L， et al. Wu chu yu tang 對胃食管反流病的影響：隨機、雙盲、安慰劑對照試驗。Phytomedicine：國際植物療法和植物藥理學雜誌。2019;56：118-25.
9. 孫彩芬， 趙瓊. 吳茱萸湯聯合柴胡疏肝散治療胃食管反流病臨床觀察. 山西中醫藥大學學報. 2019; 20(06):447-8+51.
10. 楊美豔. 加味瀉心湯聯合雷貝拉唑治療反流性食管炎（寒熱錯雜證）的臨床研究 [碩士]： 山西中醫藥大學; 2021.
11. 吳健， 劉淑敏. 加味下氣湯輔助治療中虛氣滯型胃食管反流病 60 例臨床觀察. 中國民族民間醫藥. 2019; 28(23):108-11.
12.Zhe L， Lin T， ShengSheng Z， XiaoHong S， SuNing C， Jing W. 改良小柴湖湯治療胃食管反流病：一項隨機雙重類比對照試驗。世界胃腸病學雜誌。2021;27（28）.
13. 卞銀燕， 童曉群， 徐燕芳， 張鷹. 半夏厚朴湯聯合雷貝拉唑治療反流性食管炎臨床研究. 新中醫. 2021; 53(14):18-21.
14. 陳莉麗， 徐致君， 楊振斌， 邱偉， 夏文娟. 加味六君子湯聯合雷貝拉唑治療中虛氣逆型胃食管反流病的臨床觀察. 中國藥物濫用防治雜誌. 2022b; 28(12):1859-61+69.
15. 陳振東. 加味半夏厚朴湯治療氣鬱痰阻型胃食管反流病的臨床研究 [碩士]： 甘肅中醫藥大學; 2020.
16. 鄧素萍， 孫豔. 小柴胡湯加減治療肝胃鬱熱型反流性食管炎 30 例臨床觀察. 湖南中醫雜誌. 2016; 32(07):48-9.
17. 馮昊， 李明， 任曉靜， 張遜， 袁金斌， 郭曉雙. 吳茱萸湯聯合西藥治療反流性食管炎的臨床觀察. 西北藥學雜誌. 2021; 36(04):656-9.
18. 郭鋼強. 和胃益氣降逆湯聯合雷貝拉唑治療反流性食管炎的臨床觀察. 中國中醫藥科技. 2018; 25(01):142-3.
19. 侯太源. 柴胡加龍骨牡蠣湯治療肝胃鬱熱型胃食管反流病的臨床觀察 [碩士]： 廣西中醫藥大學; 2021.
20. 侯兆川. 調胃降逆湯聯合埃索美拉唑鎂腸溶片治療肝胃鬱熱型反流性食管炎的臨床觀察 [碩士]： 山西中醫藥大學; 2018.
21. 黃啟婷. 降逆護膜湯治療胃食管反流病的實驗及臨床研究 [碩士]： 南京中醫藥大學; 2015.
22. 黃曉靜， 張思容， 岑鳳蘭. 半夏瀉心湯加味治療寒熱錯雜型非糜爛性胃食管反流病的臨床觀察. 中國民間療法. 2021; 29(8).
23. 黃亞娟， 蔣士生， 鄧芳， 周立武， 胡錦洋， 王迪， et al. 柴芍六君子湯合烏貝散治療難治性胃食管反流病 96 例臨床觀察. 湖南中醫雜誌. 2017; 33(06):10-2.
24. 劉全喜， 陳亮， 賈玉， 王韶華， 朱燕. 旋覆代赭湯合六君子湯聯合雷貝拉唑鈉腸溶膠囊治療反流性食管炎中虛氣逆證臨床觀察. 河北中醫. 2021a; 43(08):1332-6+45.
25. 劉松. 術苓瀉心湯聯合艾司奧美拉唑治療寒熱錯雜型反流性食管炎的臨床研究 [碩士]： 山西中醫藥大學; 2021b.
26. 劉雅萍. 蠲郁降逆湯聯合雷貝拉唑治療反流性食管炎（氣鬱痰阻型）的臨床觀察 [碩士]： 天津中醫藥大學; 2022.

27. 馬思遠. 疏肝清胃降逆湯治療肝胃鬱熱型反流性食管炎的臨床研究 [碩士]： 寧夏醫科大學; 2022.
28. 王碩. 柴胡梔子豉湯治療肝胃鬱熱型非糜爛性胃食管反流病的臨床研究 [碩士]： 山東中醫藥大學; 2022.
29. 張超強. 四逆降逆湯聯合雷貝拉唑治療肝胃鬱熱型反流性食管炎的臨床研究 [碩士]： 山西中醫藥大學; 2020.
30. 楊勤， 李軍祥， 李曉紅. 半夏瀉心湯加減對非糜爛性反流病癥狀和生活質量的影響. 北京中醫藥大學學報. 2013; 36(04):280-5.
31. 鍾毅， 饒梅冰， 范明. 降逆和胃湯治療非糜爛性反流病 80 例臨床觀察. 新中醫. 2011; 43(01):31-3.
32. 朱燕， 劉大銘， 賈玉. 化肝煎合左金丸聯合雷貝拉唑鈉腸溶膠囊治療肝胃鬱熱型反流性食管炎臨床觀察. 河北中醫. 2021; 43(01):100-2+6.
33. 倪小燕， 安禎祥. 胃康 I.號方治療中虛氣逆型反流性食管炎臨床觀察. 山東中醫雜誌. 2018; 37(05):378-80.
34. 秦嫣， 范懷玲， 朱凌雲， 朱遠鎔， 宋宏偉， 楊芸峰. 升降並舉法治療非糜爛性胃食管反流病的臨床研究. 世界中醫藥. 2013; 8(04):409-11.

附錄 10 34 隨機對照試驗的特征

Author, year	Reported funding*	Clinical conditions	Interventions/ comparators	No. of patients in the group (A/R)	Age mean±SD / range (years)	RCTs incorporating syndrome differentiation or not	Primary outcome	Secondary outcomes			
							Symptom relief rate[#]	Change of score from baseline[#^]		Adverse event rate [#]	Recurrence rate [#]
							N(%)	Scale	Mean±SD	N(%)	N(%)
Zhang, 2012	NR	Non-erosive GERD	Jiangni Humo Tang+PPI	23/25	NR	Yes	22(95.7)	NR	NR	NR	NR
			PPI	25/28	NR	NA	22(88.0)		NR	NR	NR
		Erosive GERD	Jiangni Humo Tang+PPI	32/35	NR	Yes	30(93.8)	NR	NR	NR	NR
			PPI	29/32	NR	NA	26(89.7)		NR	NR	NR
Ding, 2012	NR	Erosive GERD	Jiangni Humo Tang+PPI	30/30	48.1±13.2	Yes	28(93.3)	NR	NR	NR	NR
			PPI	30/30	47.7±12.5	NA	27(90.0)		NR	NR	NR
Chen, 2009	NR	Erosive GERD	Banxia Xiexin Tang	90/90	47.3±12.0	No	88(97.8)	RDQ	-5.63±2.61	0(0.00)	NR
			PPI	60/60	46.9±12.5	NA	58(96.7)		-6.07±2.57	14(23.3)	NR
Zhang, 2013	Yes	Unspecified GERD	Prescriptions based on syndrome differentiation	91/93	49.2±11.0	Yes	64(70.3)	NR	NR	0(0.00)	NR
			PPI	88/94	47.8±11.3	NA	73(83.0)		NR	0(0.00)	NR
Zhou, 2021	NR	Unspecified GERD	Jiangni Tiaowei Tang	30/30	NR	Yes	NR	RDQ	-13.6±6.21	0(0.00)	NR
			PPI	30/30	NR	NA	NR		-10.3±5.25	0(0.00)	NR

Author, year	Reported funding*	Clinical conditions	Interventions/ comparators	No. of patients in the group (A/R)	Age mean±SD / range (years)	RCTs incorporating syndrome differentiation or not	Primary outcome	Secondary outcomes			
							Symptom relief rate[#]	Change of score from baseline[#^]		Adverse event rate [#]	Recurrence rate [#]
							N(%)	Scale	Mean±SD	N(%)	N(%)
Chen, 2022a	NR	Erosive GERD	Dieli Zuojin Tang+PPI	30/30	47.2±2.37	Yes	30(100)	RDQ	-20.7±4.16	NR	NR
			PPI	30/30	47.3±2.16	NA	25(83.3)		-17.2±4.04	NR	NR
Li, 2023	Yes	Erosive GERD	Chaiban Tang+PPI	34/38	48.4±7.78	Yes	NR	GERD Q	-5.09±2.29	NR	NR
			PPI	34/38	47.8±6.98	NA	NR		-4.28±4.29	NR	NR
Shih, 2019	Yes	Unspecified GERD	Wuzhuyu Tang	40/45	46.0±13.9	No	NR	RDQ/ GERD Q	-16.9 ± 9.92 / -4.50 ± 2.95	NR	NR
			PPI	37/45	47.0±13.4	NA	NR		-9.95 ± 10.6/ -2.38 ± 3.26	NR	NR
Sun, 2019	Yes	Unspecified GERD	Wuzhuyu Tang Lianhe Chaihu Shugan San+PPI	30/30	52.8±10.2	Yes	NR	NA	NR	NR	1(3.00)
			PPI	30/30	53.7±9.20	NA	NR		NR	NR	3(10.0)
Yang, 2021	NR	Erosive GERD	Xiexin Tang+PPI	39/39	44.2±10.2	Yes	36(92.3)	RDQ	-12.0±4.48	0(0.00)	NR
			PPI	38/39	43.1±13.0	NA	30(79.0)		-9.36±3.97	0(0.00)	NR
Wu, 2019	NR	Unspecified GERD	Xiaqi Tang+PPI	60/60	42.9±14.5	Yes	NR	NA	NR	1(1.67)	NR
			PPI	60/60	42.7±14.4	NA	NR		NR	3(5.00)	NR
Zhe, 2021	Yes	Unspecified GERD	Xiaochaihu Tang	116/144	50.3±10.4	No	NR	GERD Q	-2.94±2.11	NR	NR

Author, year	Reported funding*	Clinical conditions	Interventions/ comparators	No. of patients in the group (A/R)	Age mean±SD / range (years)	RCTs incorporating syndrome differentiation or not	Primary outcome	Secondary outcomes			
							Symptom relief rate[#]	Change of score from baseline[#^]		Adverse event rate [#]	Recurrence rate [#]
							N(%)	Scale	Mean±SD	N(%)	N(%)
			PPI	130/144	50.5±11.7	NA	NR		-2.51±1.92	NR	NR
Bian, 2021	NR	Erosive GERD	Banxia Houpo Tang+PPI	40/40	47.2±12.4	No	NR	RDQ	-14.4±4.61	NR	NR
			PPI	40/40	46.9±11.7	NA	NR		-9.08±4.56	NR	NR
Chen, 2022b	NR	Unspecified GERD	Liujunzi Tang+PPI	30/30	NR	Yes	NR	RDQ	-21.3±8.96	NR	NR
			PPI	30/30	NR	NA	NR		-16.0±9.31	NR	NR
Chen, 2020	NR	Unspecified GERD	Banxia Houpo Tang+PPI	40/42	23-65	Yes	NR	GERD Q	-2.25±1.42	0(0.00)	NR
			PPI	38/42	24-65	NA	NR		-1.06±1.22	0(0.00)	NR
Deng, 2016	NR	Erosive GERD	Xiaochaihu Tang	30/30	43.5±12.7	Yes	NR	RDQ	-8.82±6.39	NR	NR
			PPI	30/30	44.3±13.5	NA	NR		-5.76±5.06	NR	NR
Feng, 2021	Yes	Erosive GERD	Wuzhuyu Tang+PPI	507/513	54.8±6.21	Yes	NR	NR	NR	33(6.51)	NR
			PPI	509/513	55.1±6.38	NA	NR		NR	46(9.04)	NR
Guo, 2018	NR	Erosive GERD	Hewei Yiqi Jiangni Tang+PPI	60/60	NR	No	NR	NR	NR	1(1.70)	3(9.7)
			PPI	60/60	NR	NA	NR		NR	9(15.0)	15(75.0)
Hou, 2021	NR	Unspecified GERD	Chaihu Jia Longgu Muli Tang	30/30	48.7±12.4	Yes	28(93.3)	RDQ	-11.3±5.57	1(3.33)	NR

Author, year	Reported funding*	Clinical conditions	Interventions/ comparators	No. of patients in the group (A/R)	Age mean±SD / range (years)	RCTs incorporating syndrome differentiation or not	Primary outcome	Secondary outcomes			
							Symptom relief rate[#]	Change of score from baseline[#^]		Adverse event rate[#]	Recurrence rate[#]
							N(%)	Scale	Mean±SD	N(%)	N(%)
			PPI	30/30	50.6±14.5	NA	25(83.3)		-6.86±5.55	2(6.66)	NR
Hou, 2018	NR	Erosive GERD	Tiaowei Jiangni Tang+PPI	30/30	57.6±7.53	Yes	NR	RDQ	-10.2±1.60	0(0.00)	NR
			PPI	30/30	54.8±9.68	NA	NR		-6.46±1.66	0(0.00)	NR
Huang, 2015	NR	Unspecified GERD	Jiangni Humo Tang+PPI	30/32	44.6±14.2	Yes	30(100)	RDQ	-12.9±4.32	NR	NR
			PPI	30/32	50.4±12.2	NA	25(83.3)		-12.0±3.79	NR	NR
Huang, 2021	Yes	Non-erosive GERD	Banxia Xiexin Tang	30/30	43.6±3.72	Yes	NR	GERD Q	-6.51±2.17	1(3.30)	NR
			PPI	30/30	43.7±3.69	NA	NR		-4.38±2.39	7(23.3)	NR
Huang, 2017	Yes	Unspecified GERD	Chaishao Liujunzi Tang He Wubei San	96/96	41.3±13.0	No	92(95.8)	RDQ	-22.8±5.92	NR	NR
			PPI	96/96	42.1±12.3	NA	60(62.5)		-3.00±7.27	NR	NR
Liu, 2021a	NR	Erosive GERD	Xuanfu Daizhe Tang He Liujunzi Tang+PPI	40/40	47.5±10.6	Yes	NR	RDQ	-25.0±7.74	0(0.0)	NR
			PPI	40/40	40.1±11.6	NA	NR		-21.5±6.49	1(2.50)	NR
Liu, 2021b	NR	Erosive GERD	Zhuling Xiexin Tang+PPI	36/36	47.9±12.3	Yes	NR	GERD Q	-3.88±2.35	0(0.0)	NR
			PPI	36/36	47.2±12.6	NA	NR		-2.70±2.69	0(0.0)	NR

Author, year	Reported funding*	Clinical conditions	Interventions/ comparators	No. of patients in the group (A/R)	Age mean±SD / range (years)	RCTs incorporating syndrome differentiation or not	Primary outcome	Secondary outcomes			
							Symptom relief rate[#]	Change of score from baseline[#^]		Adverse event rate [#]	Recurrence rate [#]
							N(%)	Scale	Mean±SD	N(%)	N(%)
Liu, 2022	NR	Erosive GERD	Juanyu Jiangni Tang+PPI	30/30	49.9±13.2	Yes	NR	GERDQ	-5.64±2.90	0(0.0)	NR
			PPI	30/30	55.8±10.5	NA	NR		-4.14±3.12	0(0.0)	NR
Ma, 2022	NR	Erosive GERD	Shugan Qingwei Jiangni Tang	30/35	37-57	Yes	27(90.0)	RDQ	-13.7±5.41	0(0.0)	NR
			PPI	33/35	32-57	NA	27(81.8)		-11.5±3.88	0(0.0)	NR
Wang, 2022	NR	Non-erosive GERD	Chaihu Zhizi Chitang Fang	40/40	46.8±10.8	Yes	NR	RDQ	-18.1±3.92	0(0.0)	NR
			PPI	40/40	44.5±10.2	NA	NR		-17.2±3.05	0(0.00)	NR
Zhang, 2020	NR	Erosive GERD	Sini Jiangni Tang+PPI	56/56	42.8±12.0	Yes	54(96.4)	RDQ	-13.3±3.99	0(0.00)	NR
			PPI	56/56	43.3± 13.6	NA	45(80.4)		-10.5±3.87	0(0.00)	NR
Yang, 2013	Yes	Non-erosive GERD	Banxia Xiexin Tang	64/64	50.4±10.0	Yes	56(87.5)	RDQ	-9.03±4.63	NR	NR
			PPI	63/63	46.3±12.2	NA	57(90.5)		-10.2±3.16	NR	NR
Zhong, 2011	NR	Non-erosive GERD	Jiangni Hewei Tang	80/80	32.9±18.6	No	75(93.8)	RDQ	NR	0(0.00)	NR
			PPI	80/80	34.5±11.3	NA	64(80.0)		NR	10(12.5)	NR
Zhu, 2021	NR	Erosive GERD	HuaganJian He Zuojin Wan+PPI	43/43	47.5±12.7	Yes	NR	NR	NR	NR	8(18.6)
			PPI	43/43	46.3±15.7	NA	NR		NR	NR	15(34.9)

Author, year	Reported funding*	Clinical conditions	Interventions/ comparators	No. of patients in the group (A/R)	Age mean±SD / range (years)	RCTs incorporating syndrome differentiation or not	Primary outcome	Secondary outcomes			
							Symptom relief rate[#]	Change of score from baseline[#^]		Adverse event rate[#]	Recurrence rate[#]
							N(%)	Scale	Mean±SD	N(%)	N(%)
Ni, 2018	NR	Erosive GERD	Weikang Yihao Fang+PPI	30/30	21-62	Yes	NR	GERD Q	-9.00±2.89	0(0.00)	NR
			PPI	30/30	20-65	NA	NR		-7.14±3.69	0(0.00)	NR
Qin, 2013	Yes	Non-erosive GERD	Formulae based on Qi Balance Method	54/54	50.2±11.4	No	49(90.7)	RDQ	NR	NR	NR
			PPI	50/50	52.2±12.0	NA	46(92.0)		NR	NR	NR

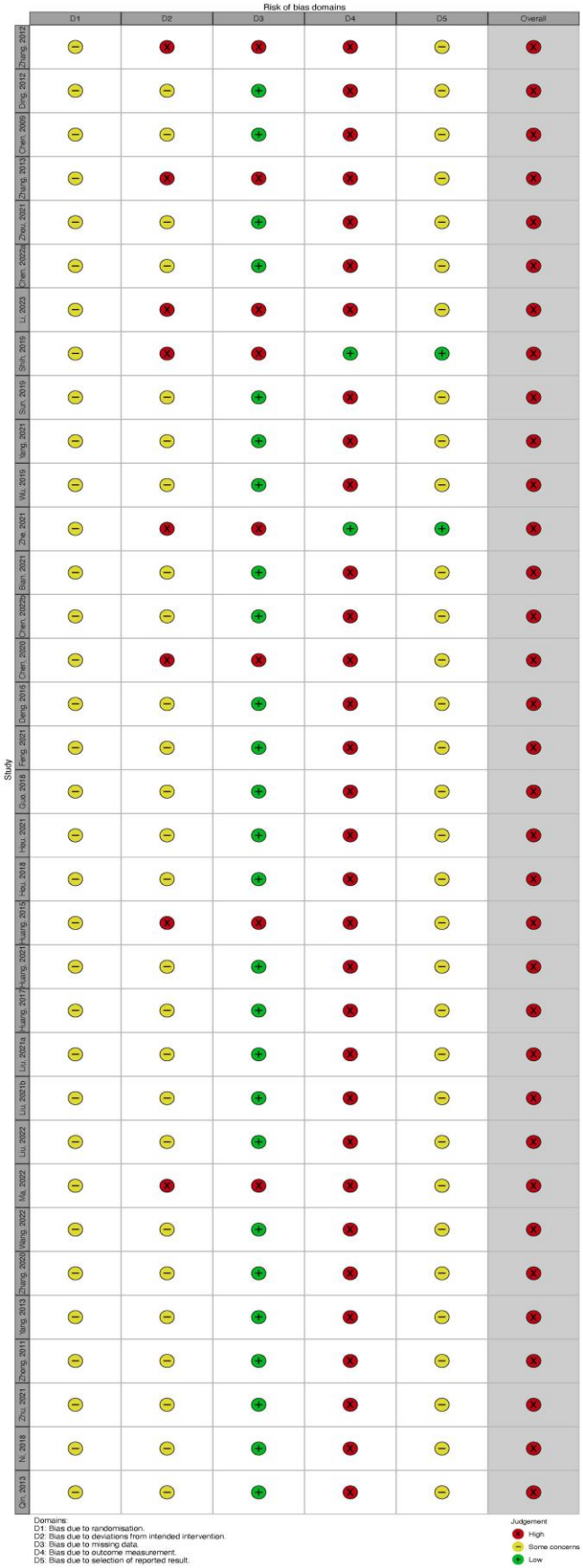
Notes: RCT: randomized controlled trials; GERD: gastroesophageal reflux disease; NR: not reported; A: number of patients analysed; R: number of patients randomized; SD: standard deviation; NA: not applicable; RDQ: reflux disease questionnaire; GERDQ: gastroesophageal reflux disease questionnaire.

*: For RCTs reporting funding support, the type of funding received was non-industry funding.

#: The follow-up duration for the assessment of symptom relief rate, change of score from baseline, or adverse event rate was 8 weeks, with the exception of Study Ni, 2018, which had a follow-up duration of 24 weeks. Meanwhile, the follow-up duration for recurrence rate was 12 months.

^: Change of score from baseline less than 0 indicated that patients' symptoms were relieved after treatment.

附錄 11 34 個隨機對照試驗的偏倚風險評估結果



附錄 12 34 個隨機對照試驗的偏倚風險評估具體表現：使用 Cochrane 偏倚風險評估工具

Domains	Items[a]	RCTs[b]																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Bias arising from the randomization process	**1.1**	NI	NI	NI	Y	NI	NI	NI	Y	NI	Y	Y	Y	NI	NI	Y	NI	NI
	1.2	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI
	1.3	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	RoB	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some
Bias due to deviations from intended interventions	**2.1**	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y	Y	Y
	2.2	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y	Y	Y
	2.3	NI	NI	NI	NI	NI	NI	NI	NA	NI	NI	NI	NA	NI	NI	NI	NI	NI
	2.4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	2.5	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	2.6	N	Y	Y	N	Y	Y	N	N	Y	N	Y	N	Y	Y	N	Y	N
	2.7	Y	NA	NA	Y	NA	NA	Y	Y	NA	N	NA	Y	NA	NA	Y	NA	N
	RoB	High	Some	Some	High	Some	Some	High	High	Some	Some	Some	High	Some	Some	High	Some	Some
Bias due to missing outcome data	**3.1**	N	Y	Y	N	Y	Y	N	N	Y	Y	Y	N	Y	Y	N	Y	Y
	3.2	N	NA	NA	N	NA	NA	N	N	NA	NA	NA	N	NA	NA	N	NA	NA
	3.3	NI	NA	NA	NI	NA	NA	NI	NI	NA	NA	NA	Y	NA	NA	NI	NA	NA
	3.4	NI	NA	NA	NI	NA	NA	NI	NI	NA	NA	NA	Y	NA	NA	NI	NA	NA
	RoB	High	Low	Low	High	Low	Low	High	High	Low	Low	Low	High	Low	Low	High	Low	Low
Bias in measurement of the outcome	**4.1**	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	4.2	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	4.3	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y	Y	Y
	4.4	Y	Y	Y	Y	Y	Y	Y	NA	Y	Y	Y	NA	Y	Y	Y	Y	Y
	4.5	Y	Y	Y	Y	Y	Y	Y	NA	Y	Y	Y	NA	Y	Y	Y	Y	Y
	RoB	High	High	High	High	High	High	High	Low	High	High	High	Low	High	High	High	High	High
Bias in selection of the reported result	**5.1**	NI	NI	NI	NI	NI	NI	NI	Y	NI	NI	NI	Y	NI	NI	NI	NI	NI
	5.2	NI	NI	NI	NI	NI	NI	NI	N	NI	NI	NI	N	NI	NI	NI	NI	NI
	5.3	NI	NI	NI	NI	NI	NI	NI	N	NI	NI	NI	N	NI	NI	NI	NI	NI
	RoB	Some	Some	Some	Some	Some	Some	Some	Low	Some	Some	Some	Low	Some	Some	Some	Some	Some
Overall bias		High	High	High	High	High	High	High	High	High	High	High	High	High	High	High	High	High

Domains	Items[a]	RCTs[b] (Continue)																
		18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
Bias arising from the randomization process	**1.1**	Y	Y	Y	NI	Y	NI	Y	Y	Y	NI	Y	Y	Y	Y	Y	Y	Y
	1.2	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI
	1.3	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	RoB	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some
Bias due to deviations from intended interventions	**2.1**	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	2.2	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	2.3	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI
	2.4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	2.5	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	2.6	Y	Y	Y	N	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y
	2.7	NA	NA	NA	Y	NA	NA	NA	NA	NA	Y	NA	NA	NA	NA	NA	NA	NA
	RoB	Some	Some	Some	High	Some	Some	Some	Some	Some	High	Some	Some	Some	Some	Some	Some	Some
Bias due to missing outcome data	**3.1**	Y	Y	Y	N	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y
	3.2	NA	NA	NA	N	NA	NA	NA	NA	NA	N	NA	NA	NA	NA	NA	NA	NA
	3.3	NA	NA	NA	NI	NA	NA	NA	NA	NA	NI	NA	NA	NA	NA	NA	NA	NA
	3.4	NA	NA	NA	NI	NA	NA	NA	NA	NA	NI	NA	NA	NA	NA	NA	NA	NA
	RoB	Low	Low	Low	High	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low
Bias in measurement of the outcome	**4.1**	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	4.2	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	4.3	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	4.4	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	4.5	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	RoB	High	High	High	High	High	High	High	High	High	High	High	High	High	High	High	High	High
Bias in selection of the reported result	**5.1**	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI
	5.2	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI
	5.3	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI
	RoB	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some	Some
Overall bias		High	High	High	High	High	High	High	High	High	High	High	High	High	High	High	High	High

Keys: RCTs: randomized controlled trials; Y: Yes; N: No; NI: no information; NA: not applicable. Low: Low risk of bias; Some: Some concerns; High: High risk of bias; RoB: risk of bias.

[a]: Detailed signaling questions for each item:

1.1 Was the allocation sequence random?
1.2 Was the allocation sequence concealed until participants were recruited and assigned to interventions?
1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?
2.1. Were participants aware of their assigned intervention during the trial?
2.2. Were carers and trial personnel aware of participants' assigned intervention during the trial?
2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention beyond what would be expected in usual practice?
2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?
2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?
2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?
2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?
3.1 Were outcome data available for all, or nearly all, participants randomized?
3.2 If N/PN/NI to 3.1: Is there evidence that the result was not biased by missing outcome data?
3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?
3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?
4.1 Was the method of measuring the outcome inappropriate?
4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?
4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants?
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?
5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?
5.2. ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?
5.3 ... multiple eligible analyses of the data?
[b]: 34 included RCTs as below, details can be found in Appendix 2.
1: Zhang, 2012; 2: Ding, 2012; 3: Chen, 2009; 4: Zhang, 2013; 5: Zhou, 2021; 6: Chen, 2022a; 7: Li, 2023; 8: Shih, 2019; 9: Sun, 2019; 10: Yang, 2021; 11: Wu, 2019; 12: Zhe, 2021; 13: Bian, 2021; 14: Chen, 2022b; 15: Chen, 2020; 16: Deng, 2016; 17: Feng, 2021; 18: Guo, 2018; 19: Hou, 2021; 20: Hou, 2018; 21: Huang, 2015; 22: Huang, 2021; 23: Huang, 2017; 24: Liu, 2021a; 25: Liu, 2021b; 26: Liu, 2022; 27: Ma, 2022; 28: Wang, 2022; 29: Zhang, 2020; 30: Yang, 2013; 31: Zhong, 2011; 32: Zhu, 2021; 33: Ni, 2018; 34: Qin, 2013.

附錄 13 不同的中藥方及是否採用辨證論治原則在癥狀緩解率上的 NMA 效應估計的證據質量等級

Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		RR (95%CI)	Quality of evidence	RR (95%CI)	Quality of evidence	RR (95%CI)	Quality of evidence
Formulae based on Qi Balance Method(NSD) vs Banxia Xiexin Tang(NSD)	2 (231)	-	-	0.98 (0.86; 1.11)	Very low∆§	0.98 (0.86; 1.11)	Very low
Formulae based on Qi Balance Method(NSD) vs Chaishao Liujunzi Tang He Wubei San(NSD)	2 (296)	-	-	0.64 (0.53; 0.79)	Very low∆§	0.64 (0.53; 0.79)	Very low
Formulae based on Qi Balance Method(NSD) vs Jiangni Hewei Tang(NSD)	2 (264)	-	-	0.84 (0.71; 1.00)	Low∆	0.84 (0.71; 1.00)	Low
Formulae based on Qi Balance Method(NSD) vs PPI	1 (104)	0.99 (0.88; 1.11)	Low*#	Not estimable[c]	Not estimable‡	0.99 (0.88; 1.11)	Low
Formulae based on Qi Balance Method(NSD) vs Xiexin Tang(SD)+PPI	2 (182)	-	-	0.82 (0.65; 1.03)	Very low∆§	0.82 (0.65; 1.03)	Very low
Formulae based on Qi Balance Method(NSD) vs Banxia Xiexin Tang(SD)	2 (254)	-	-	1.02 (0.86; 1.21)	Low∆	1.02 (0.86; 1.21)	Low
Formulae based on Qi Balance Method(NSD) vs Sini Jiangni Tang(SD)+PPI	2 (216)	-	-	0.82 (0.68; 0.99)	Very low∆§	0.82 (0.68; 0.99)	Very low
Formulae based on Qi Balance Method(NSD) vs Chaihu Jia Longgu Muli Tang(SD)	2 (164)	-	-	0.88 (0.71; 1.10)	Very low∆§	0.88 (0.71; 1.10)	Very low
Formulae based on Qi Balance Method(NSD) vs Shugan Qingwei Jiangni Tang(SD)	2 (174)	-	-	0.99 (0.74; 1.31)	Very low∆§	0.99 (0.74; 1.31)	Very low
Formulae based on Qi Balance Method(NSD) vs Prescriptions based on syndrome differentiation(SD)	2 (291)	-	-	1.11 (0.90; 1.37)	Very low∆§	1.11 (0.90; 1.37)	Very low

Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		RR (95%CI)	Quality of evidence	RR (95%CI)	Quality of evidence	RR (95%CI)	Quality of evidence
Formulae based on Qi Balance Method(NSD) vs Dieli Zuojin Tang(SD)+PPI	2 (164)	-	-	0.82 (0.68; 1.00)	Very low∆§	0.82 (0.68; 1.00)	Very low
Formulae based on Qi Balance Method(NSD) vs Jiangni Humo Tang(SD)+PPI	5 (348)	-	-	0.91 (0.78; 1.06)	Very low∆∆	0.91 (0.78; 1.06)	Very low
Banxia Xiexin Tang(NSD) vs Chaishao Liujunzi Tang He Wubei San(NSD)	2 (319)	-	-	0.66 (0.56; 0.78)	Low∆	0.66 (0.56; 0.78)	Low
Banxia Xiexin Tang(NSD) vs Jiangni Hewei Tang(NSD)	2 (287)	-	-	0.86 (0.75; 0.99)	Very low∆§	0.86 (0.75; 0.99)	Very low
Banxia Xiexin Tang(NSD) vs PPI	1 (127)	1.01 (0.96; 1.07)	Low*#	Not estimable[c]	Not estimable‡	1.01 (0.96; 1.07)	Low
Banxia Xiexin Tang(NSD) vs Xiexin Tang(SD)+PPI	2 (205)	-	-	0.84 (0.69; 1.03)	Low∆	0.84 (0.69; 1.03)	Low
Banxia Xiexin Tang(NSD) vs Banxia Xiexin Tang(SD)	2 (277)	-	-	1.05 (0.91; 1.20)	Very low∆§	1.05 (0.91; 1.20)	Very low
Banxia Xiexin Tang(NSD) vs Sini Jiangni Tang(SD)+PPI	2 (239)	-	-	0.84 (0.73; 0.98)	Low∆	0.84 (0.73; 0.98)	Low
Banxia Xiexin Tang(NSD) vs Chaihu Jia Longgu Muli Tang(SD)	2 (187)	-	-	0.90 (0.74; 1.10)	Very low∆§	0.90 (0.74; 1.10)	Very low
Banxia Xiexin Tang(NSD) vs Shugan Qingwei Jiangni Tang(SD)	2 (197)	-	-	1.01 (0.78; 1.31)	Low∆	1.01 (0.78; 1.31)	Low
Banxia Xiexin Tang(NSD) vs Prescriptions based on syndrome differentiation(SD)	2 (314)	-	-	1.14 (0.95; 1.37)	Very low∆§	1.14 (0.95; 1.37)	Very low
Banxia Xiexin Tang(NSD) vs Dieli Zuojin Tang(SD)+PPI	2 (187)	-	-	0.85 (0.72; 1.00)	Low∆	0.85 (0.72; 1.00)	Low
Banxia Xiexin Tang(NSD) vs Jiangni Humo Tang(SD)+PPI	5 (371)	-	-	0.93 (0.83; 1.04)	Very low∆∆	0.93 (0.83; 1.04)	Very low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Jiangni Hewei Tang(NSD)	2 (352)	-	-	1.31 (1.07; 1.60)	Very low∆§	1.31 (1.07; 1.60)	Very low

Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		RR (95%CI)	Quality of evidence	RR (95%CI)	Quality of evidence	RR (95%CI)	Quality of evidence
Chaishao Liujunzi Tang He Wubei San(NSD) vs PPI	1 (192)	1.53 (1.31; 1.80)	Low*#	Not estimable[c]	Not estimable‡	1.53 (1.31; 1.80)	Low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Xiexin Tang(SD)+PPI	2 (270)	-	-	1.28 (0.99; 1.64)	LowΔ	1.28 (0.99; 1.64)	Low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Banxia Xiexin Tang(SD)	2 (342)	-	-	1.59 (1.30; 1.94)	Very lowΔ§	1.59 (1.30; 1.94)	Very low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Sini Jiangni Tang(SD)+PPI	2 (304)	-	-	1.28 (1.03; 1.58)	LowΔ	1.28 (1.03; 1.58)	Low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Chaihu Jia Longgu Muli Tang(SD)	2 (252)	-	-	1.37 (1.07; 1.75)	Very lowΔ§	1.37 (1.07; 1.75)	Very low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Shugan Qingwei Jiangni Tang(SD)	2 (262)	-	-	1.53 (1.13; 2.07)	LowΔ	1.53 (1.13; 2.07)	Low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Prescriptions based on syndrome differentiation(SD)	2 (379)	-	-	1.73 (1.37; 2.19)	Very lowΔ§	1.73 (1.37; 2.19)	Very low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Dieli Zuojin Tang(SD)+PPI	2 (252)	-	-	1.28 (1.02; 1.60)	LowΔ	1.28 (1.02; 1.60)	Low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Jiangni Humo Tang(SD)+PPI	5 (436)	-	-	1.41 (1.17; 1.70)	Very lowΔΔ	1.41 (1.17; 1.70)	Very low
Jiangni Hewei Tang(NSD) vs PPI	1 (160)	1.17 (1.04; 1.33)	Low*#	Not estimable[c]	Not estimable‡	1.17 (1.04; 1.33)	Low
Jiangni Hewei Tang(NSD) vs Xiexin Tang(SD)+PPI	2 (238)	-	-	0.98 (0.78; 1.23)	Very lowΔ§	0.98 (0.78; 1.23)	Very low
Jiangni Hewei Tang(NSD) vs Banxia Xiexin Tang(SD)	2 (310)	-	-	1.21 (1.02; 1.44)	LowΔ	1.21 (1.02; 1.44)	Low
Jiangni Hewei Tang(NSD) vs Sini Jiangni Tang(SD)+PPI	2 (272)	-	-	0.98 (0.81; 1.18)	Very lowΔ§	0.98 (0.81; 1.18)	Very low

Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		RR (95%CI)	Quality of evidence	RR (95%CI)	Quality of evidence	RR (95%CI)	Quality of evidence
Jiangni Hewei Tang(NSD) vs Chaihu Jia Longgu Muli Tang(SD)	2 (220)	-	-	1.05 (0.84; 1.31)	Very lowΔ§	1.05 (0.84; 1.31)	Very low
Jiangni Hewei Tang(NSD) vs Shugan Qingwei Jiangni Tang(SD)	2 (230)	-	-	1.17 (0.88; 1.56)	Very lowΔ§	1.17 (0.88; 1.56)	Very low
Jiangni Hewei Tang(NSD) vs Prescriptions based on syndrome differentiation(SD)	2 (347)	-	-	1.32 (1.07; 1.64)	Very lowΔ§	1.32 (1.07; 1.64)	Very low
Jiangni Hewei Tang(NSD) vs Dieli Zuojin Tang(SD)+PPI	2 (220)	-	-	0.98 (0.80; 1.20)	Very lowΔ§	0.98 (0.80; 1.20)	Very low
Jiangni Hewei Tang(NSD) vs Jiangni Humo Tang(SD)+PPI	5 (404)	-	-	1.08 (0.92; 1.26)	Very lowΔΔ§	1.08 (0.92; 1.26)	Very low
PPI vs Xiexin Tang(SD)+PPI	1 (78)	0.83 (0.69; 1.01)	Low*#	Not estimable[c]	Not estimable‡	0.83 (0.69; 1.01)	Low
PPI vs Banxia Xiexin Tang(SD)	1 (150)	1.03 (0.91; 1.17)	Low*#	Not estimable[c]	Not estimable‡	1.03 (0.91; 1.17)	Low
PPI vs Sini Jiangni Tang(SD)+PPI	1 (112)	0.83 (0.73; 0.96)	Low*#	Not estimable[c]	Not estimable‡	0.83 (0.73; 0.96)	Low
PPI vs Chaihu Jia Longgu Muli Tang(SD)	1 (60)	0.89 (0.74; 1.08)	Low*#	Not estimable[c]	Not estimable‡	0.89 (0.74; 1.08)	Low
PPI vs Shugan Qingwei Jiangni Tang(SD)	1 (70)	1.00 (0.77; 1.29)	Low*#	Not estimable[c]	Not estimable‡	1.00 (0.77; 1.29)	Low
PPI vs Prescriptions based on syndrome differentiation(SD)	1 (187)	1.13 (0.95; 1.34)	Low*#	Not estimable[c]	Not estimable‡	1.13 (0.95; 1.34)	Low
PPI vs Dieli Zuojin Tang(SD)+PPI	1 (60)	0.84 (0.71; 0.98)	Low*#	Not estimable[c]	Not estimable‡	0.84 (0.71; 0.98)	Low
PPI vs Jiangni Humo Tang(SD)+PPI	4 (244)	0.92 (0.83; 1.01)	Very low*#†	Not estimable[c]	Not estimable‡	0.92 (0.83; 1.01)	Very low
Xiexin Tang(SD)+PPI vs Banxia Xiexin Tang(SD)	2 (228)	-	-	1.24 (0.99; 1.56)	Very lowΔ§	1.24 (0.99; 1.56)	Very low
Xiexin Tang(SD)+PPI vs Sini Jiangni Tang(SD)+PPI	2 (190)	-	-	1.00 (0.79; 1.27)	LowΔ	1.00 (0.79; 1.27)	Low
Xiexin Tang(SD)+PPI vs Chaihu Jia Longgu Muli Tang(SD)	2 (138)	-	-	1.07 (0.82; 1.40)	Very lowΔ§	1.07 (0.82; 1.40)	Very low

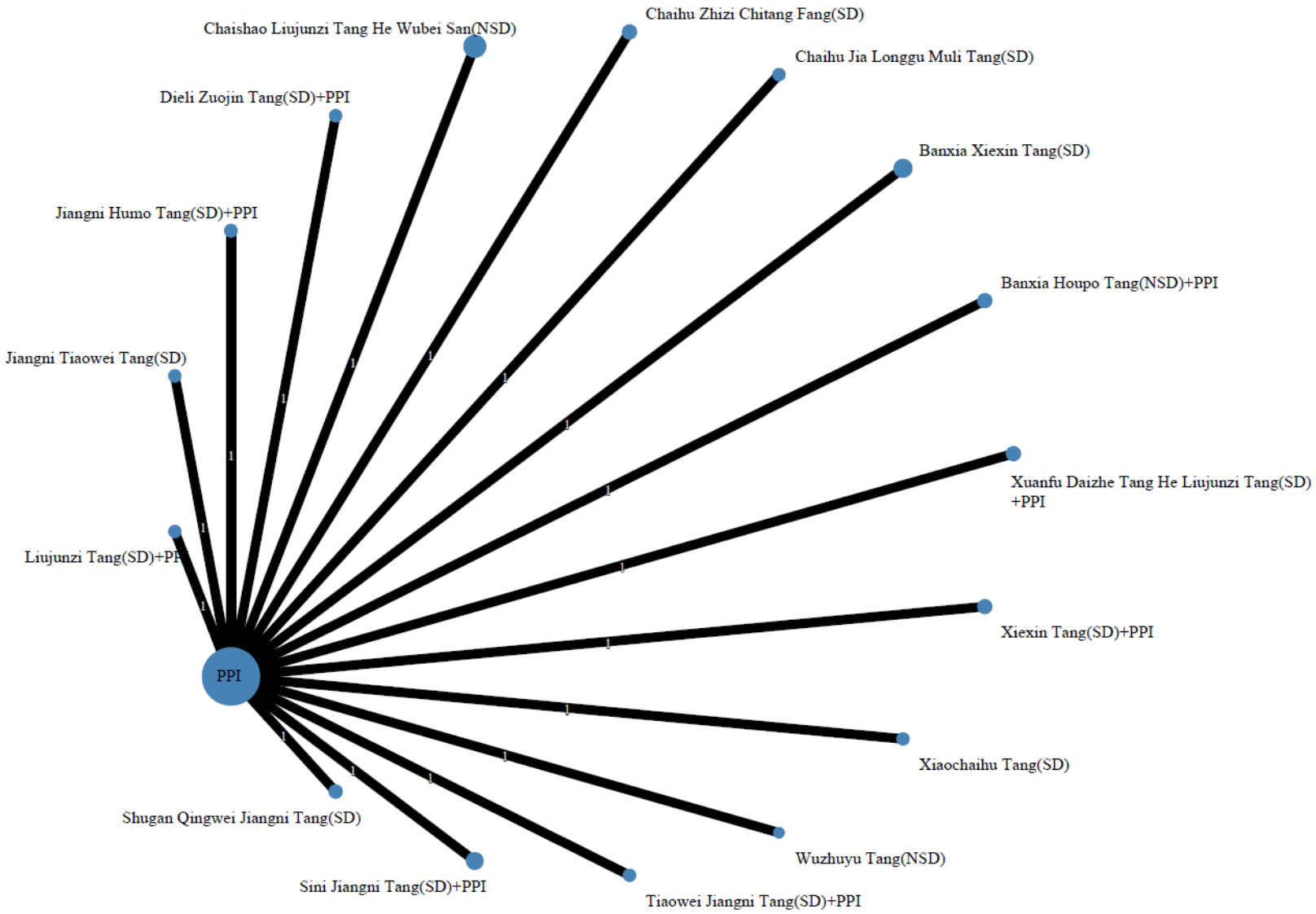
Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		RR (95%CI)	Quality of evidence	RR (95%CI)	Quality of evidence	RR (95%CI)	Quality of evidence
Xiexin Tang(SD)+PPI vs Shugan Qingwei Jiangni Tang(SD)	2 (148)	-	-	1.20 (0.87; 1.65)	LowΔ	1.20 (0.87; 1.65)	Low
Xiexin Tang(SD)+PPI vs Prescriptions based on syndrome differentiation(SD)	2 (265)	-	-	1.35 (1.04; 1.76)	Very lowΔ§	1.35 (1.04; 1.76)	Very low
Xiexin Tang(SD)+PPI vs Dieli Zuojin Tang(SD)+PPI	2 (138)	-	-	1.00 (0.78; 1.29)	LowΔ	1.00 (0.78; 1.29)	Low
Xiexin Tang(SD)+PPI vs Jiangni Humo Tang(SD)+PPI	5 (322)	-	-	1.10 (0.89; 1.37)	Very lowΔΔ§	1.10 (0.89; 1.37)	Very low
Banxia Xiexin Tang(SD) vs Sini Jiangni Tang(SD)+PPI	2 (262)	-	-	0.81 (0.67; 0.97)	Very lowΔ§	0.81 (0.67; 0.97)	Very low
Banxia Xiexin Tang(SD) vs Chaihu Jia Longgu Muli Tang(SD)	2 (210)	-	-	0.86 (0.69; 1.08)	Very lowΔ§	0.86 (0.69; 1.08)	Very low
Banxia Xiexin Tang(SD) vs Shugan Qingwei Jiangni Tang(SD)	2 (220)	-	-	0.97 (0.73; 1.28)	Very lowΔ§	0.97 (0.73; 1.28)	Very low
Banxia Xiexin Tang(SD) vs Prescriptions based on syndrome differentiation(SD)	2 (337)	-	-	1.09 (0.88; 1.35)	Very lowΔ§	1.09 (0.88; 1.35)	Very low
Banxia Xiexin Tang(SD) vs Dieli Zuojin Tang(SD)+PPI	2 (210)	-	-	0.81 (0.66; 0.99)	Very lowΔ§	0.81 (0.66; 0.99)	Very low
Banxia Xiexin Tang(SD) vs Jiangni Humo Tang(SD)+PPI	5 (394)	-	-	0.89 (0.76; 1.04)	Very lowΔΔ§	0.89 (0.76; 1.04)	Very low
Sini Jiangni Tang(SD)+PPI vs Chaihu Jia Longgu Muli Tang(SD)	2 (172)	-	-	1.07 (0.85; 1.35)	Very lowΔ§	1.07 (0.85; 1.35)	Very low
Sini Jiangni Tang(SD)+PPI vs Shugan Qingwei Jiangni Tang(SD)	2 (182)	-	-	1.20 (0.90; 1.60)	LowΔ	1.20 (0.90; 1.60)	Low
Sini Jiangni Tang(SD)+PPI vs Prescriptions based on syndrome differentiation(SD)	2 (299)	-	-	1.35 (1.08; 1.69)	Very lowΔ§	1.35 (1.08; 1.69)	Very low
Sini Jiangni Tang(SD)+PPI vs Dieli Zuojin Tang(SD)+PPI	2 (172)	-	-	1.00 (0.81; 1.24)	LowΔ	1.00 (0.81; 1.24)	Low
Sini Jiangni Tang(SD)+Jiangni Humo Tang(SD)+PPI	5 (356)	-	-	1.10 (0.93; 1.31)	Very lowΔΔ§	1.10 (0.93; 1.31)	Very low

Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		RR (95%CI)	Quality of evidence	RR (95%CI)	Quality of evidence	RR (95%CI)	Quality of evidence
Chaihu Jia Longgu Muli Tang(SD) vs Shugan Qingwei Jiangni Tang(SD)	2 (130)	-	-	1.12 (0.82; 1.54)	Very low∆§	1.12 (0.82; 1.54)	Very low
Chaihu Jia Longgu Muli Tang(SD) vs Prescriptions based on syndrome differentiation(SD)	2 (247)	-	-	1.26 (0.98; 1.63)	Very low∆§	1.26 (0.98; 1.63)	Very low
Chaihu Jia Longgu Muli Tang(SD) vs Dieli Zuojin Tang(SD)+PPI	2 (120)	-	-	0.94 (0.73; 1.19)	Low∆	0.94 (0.73; 1.19)	Low
Chaihu Jia Longgu Muli Tang(SD) vs Jiangni Humo Tang(SD)+PPI	5 (304)	-	-	1.03 (0.83; 1.27)	Very low∆∆	1.03 (0.83; 1.27)	Very low
Shugan Qingwei Jiangni Tang(SD) vs Prescriptions based on syndrome differentiation(SD)	2 (257)	-	-	1.13 (0.83; 1.54)	Very low∆§	1.13 (0.83; 1.54)	Very low
Shugan Qingwei Jiangni Tang(SD) vs Dieli Zuojin Tang(SD)+PPI	2 (130)	-	-	0.84 (0.62; 1.13)	Low∆	0.84 (0.62; 1.13)	Low
Shugan Qingwei Jiangni Tang(SD) vs Jiangni Humo Tang(SD)+PPI	5 (314)	-	-	0.92 (0.70; 1.21)	Very low∆∆§	0.92 (0.70; 1.21)	Very low
Prescriptions based on syndrome differentiation(SD) vs Dieli Zuojin Tang(SD)+PPI	2 (247)	-	-	0.74 (0.59; 0.94)	Very low∆§	0.74 (0.59; 0.94)	Very low
Prescriptions based on syndrome differentiation(SD) vs Jiangni Humo Tang(SD)+PPI	5 (431)	-	-	0.81 (0.67; 0.99)	Very low∆∆	0.81 (0.67; 0.99)	Very low
Dieli Zuojin Tang(SD)+PPI vs Jiangni Humo Tang(SD)+PPI	5 (304)	-	-	1.10 (0.91; 1.32)	Very low∆∆§	1.10 (0.91; 1.32)	Very low

Note: NSD, Non-Syndrome Differentiation; SD, Syndrome Differentiation; RCT, Randomized Controlled Trial; RR, Risk Ratio; CI: confidence interval;PPI, Proton Pump Inhibitors. -: Not applicable. *Serious study limitations (crucial risk of bias for one criterion, or multiple criteria, and likely to seriously alter the results); # Serious impresison (small sample size less than 400); † Serious indirectness(populations with various conditions, such as Erosive Gatroesophageal Reflux Disease(GERD), Non-Erosive GERD, and Unspecified GERD.); ∆Contributing direct evidence of low quality; ∆∆Contributing direct evidence of low or very low quality; ‡: Cannot be estimated because the drug was not connected in a loop in the evidence network; §Indirectness because of questionable comparability of trial populations or because of intransitivity.

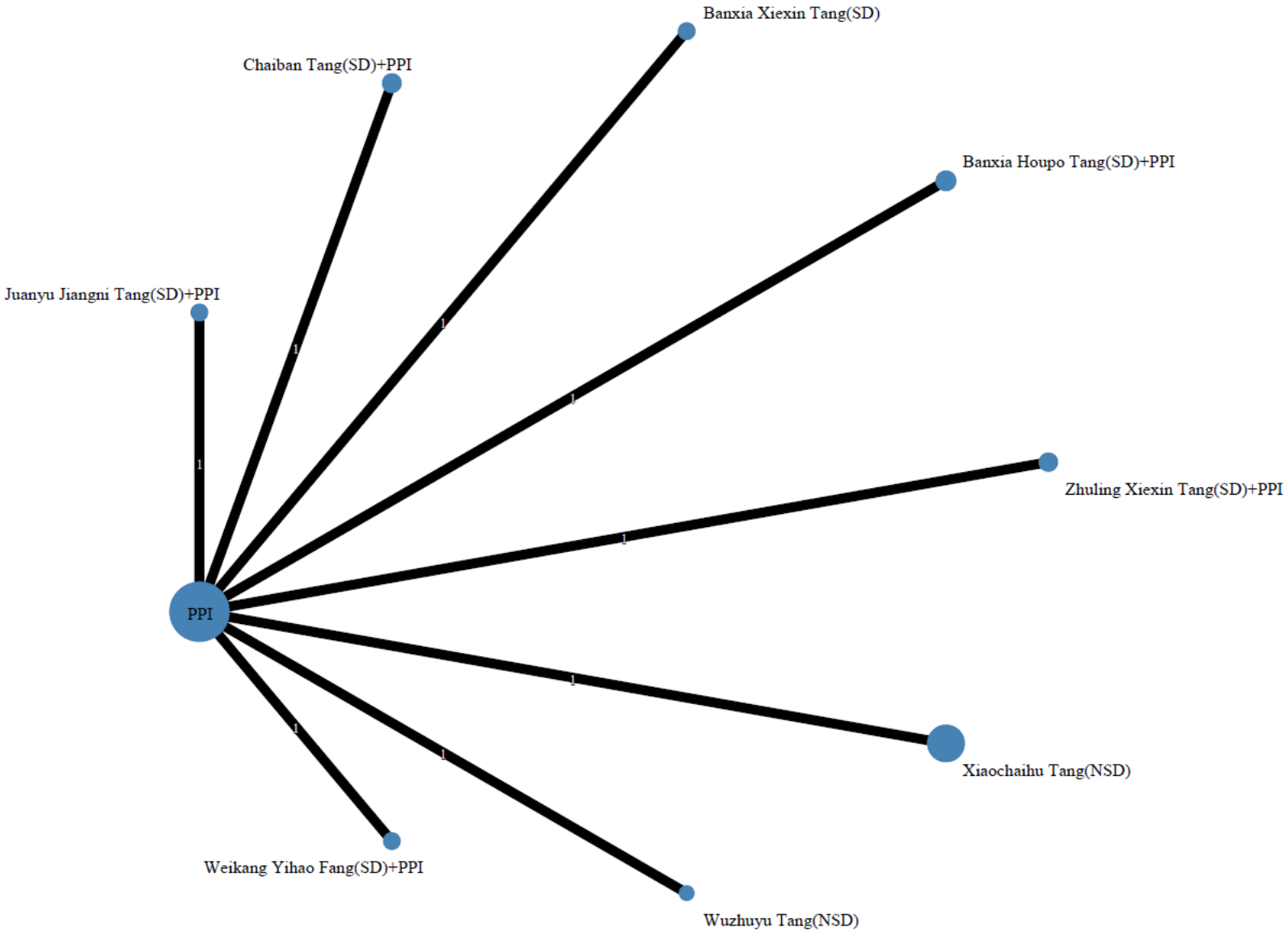
附錄 14 不同中藥方及是否採用辨證論治幹預模式的網絡幾何圖

(1) 結局： RDQ 評分變化



Notes: RDQ, reflux disease questionnaire; NSD, non-syndrome differentiation; SD, syndrome differentiation; PPI, proton pump inhibitors. Nodes represent the interventions, node sizes correspond to the number of studies involved, lines connecting nodes represent direct comparisons between pairs of interventions. Width of the lines represents the proportion of the number of trials for each comparison as compared to total number of trials. Number on the line represents the count of the direct comparisons.

(2) 結局: GERDQ 評分變化



Notes: GERDQ, gastroesophageal reflux disease questionnaire; NSD, non-syndrome differentiation; SD, syndrome differentiation; PPI, proton pump inhibitors. Nodes represent the interventions, node sizes correspond to the number of studies involved, lines connecting nodes represent direct comparisons between pairs of interventions. Width of the lines represents the proportion of the number of trials for each comparison as compared to total number of trials. Number on the line represents the count of the direct comparisons.

附錄 **15** 不同的中藥方及是否採用辨證論治原則在癥狀評分變化上兩兩比較的結果

(1) 結局：**RDQ** 評分變化

Chaishao Liujunzi Tang He Wubei San(NSD)																	
-12.9 (-19.2; -6.50)	Wuzhuyu Tang(NSD)																
-14.5 (-17.2; -11.7)	-1.63 (-8.02; 4.76)	Banxia Houpo Tang(NSD)+PPI															
-14.5 (-19.5; -9.52)	-1.66 (-9.29; 5.97)	-0.03 (-5.07; 5.01)	Liujunzi Tang(SD)+PPI														
-15.4 (-18.7; -12.0)	-2.51 (-9.20; 4.18)	-0.88 (-4.34; 2.58)	-0.85 (-6.26; 4.56)	Chaihu Jia Longgu Muli Tang(SD)													
-16.0 (-18.1; -14.0)	-3.18 (-9.31; 2.95)	-1.55 (-3.72; 0.62)	-1.52 (-6.22; 3.18)	-0.67 (-3.60; 2.26)	Tiaowei Jiangni Tang(SD)+PPI												
-16.2 (-19.0; -13.4)	-3.39 (-9.80; 3.02)	-1.76 (-4.65; 1.13)	-1.73 (-6.80; 3.34)	-0.88 (-4.38; 2.62)	-0.21 (-2.44; 2.02)	Dieli Zuojin Tang(SD)+PPI											
-16.3 (-19.9; -12.6)	-3.42 (-10.2; 3.41)	-1.79 (-5.51; 1.93)	-1.76 (-7.34; 3.82)	-0.91 (-5.12; 3.30)	-0.24 (-3.48; 3.00)	-0.03 (-3.79; 3.73)	Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI										
-16.4 (-19.9; -13.0)	-3.59 (-10.3; 3.14)	-1.96 (-5.50; 1.57)	-1.93 (-7.40; 3.53)	-1.08 (-5.13; 2.96)	-0.41 (-3.44; 2.61)	-0.20 (-3.78; 3.37)	-0.17 (-4.45; 4.10)	Jiangni Tiaowei Tang(SD)									
-16.7 (-20.2; -13.3)	-3.89 (-10.6; 2.84)	-2.26 (-5.80; 1.28)	-2.23 (-7.70; 3.24)	-1.38 (-5.43; 2.67)	-0.71 (-3.74; 2.32)	-0.50 (-4.08; 3.08)	-0.47 (-4.75; 3.81)	-0.30 (-4.42; 3.82)	Xiaochaihu Tang(SD)								
-17.0 (-19.3; -14.6)	-4.12 (-10.4; 2.12)	**-2.49 (-4.97; -0.01)**	-2.46 (-7.31; 2.39)	-1.61 (-4.78; 1.56)	-0.94 (-2.61; 0.73)	-0.73 (-3.27; 1.81)	-0.70 (-4.15; 2.75)	-0.53 (-3.78; 2.73)	-0.23 (-3.49; 3.03)	Sini Jiangni Tang(SD)+PPI							
-17.2 (-19.8; -14.5)	-4.34 (-10.7; 2.01)	-2.71 (-5.46; 0.04)	-2.68 (-7.67; 2.31)	-1.83 (-5.21; 1.55)	-1.16 (-3.21; 0.89)	-0.95 (-3.75; 1.85)	-0.92 (-4.57; 2.73)	-0.75 (-4.21; 2.72)	-0.45 (-3.92; 3.02)	-0.22 (-2.60; 2.16)	Xiexin Tang(SD)+PPI						
-17.6 (-20.5; -14.7)	-4.74 (-11.2; 1.72)	**-3.11 (-6.09; -0.13)**	-3.08 (-8.20; 2.04)	-2.23 (-5.81; 1.35)	-1.56 (-3.92; 0.80)	-1.35 (-4.38; 1.68)	-1.32 (-5.15; 2.51)	-1.15 (-4.80; 2.51)	-0.85 (-4.51; 2.81)	-0.62 (-3.26; 2.02)	-0.40 (-3.30; 2.50)	Shugan Qingwei Jiangni Tang(SD)					
-18.9 (-21.6; -16.2)	-6.05 (-12.4; 0.34)	**-4.42 (-7.25; -1.59)**	-4.39 (-9.42; 0.64)	**-3.54 (-6.99; -0.09)**	**-2.87 (-5.03; -0.72)**	-2.66 (-5.54; 0.22)	-2.63 (-6.34; 1.08)	-2.46 (-5.98; 1.07)	-2.16 (-5.69; 1.37)	-1.93 (-4.40; 0.54)	-1.71 (-4.45; 1.03)	-1.31 (-4.28; 1.66)	Jiangni Humo Tang(SD)+PPI				
-18.9 (-21.3; -16.5)	-6.05 (-12.3; 0.21)	**-4.42 (-6.95; -1.89)**	-4.39 (-9.26; 0.48)	**-3.54 (-6.75; -0.33)**	**-2.87 (-4.62; -1.12)**	**-2.66 (-5.24; -0.08)**	-2.63 (-6.12; 0.86)	-2.46 (-5.75; 0.84)	-2.16 (-5.46; 1.14)	-1.93 (-4.05; 0.19)	-1.71 (-4.14; 0.72)	-1.31 (-4.00; 1.38)	-0.00 (-2.52; 2.52)	Chaihu Zhizi Chitang Fang(SD)			
-19.8 (-21.7; -17.9)	**-6.95 (-13.0; -0.88)**	**-5.32 (-7.33; -3.31)**	**-5.29 (-9.91; -0.67)**	**-4.44 (-7.25; -1.63)**	**-3.77 (-4.60; -2.95)**	**-3.56 (-5.64; -1.49)**	**-3.53 (-6.66; -0.40)**	**-3.36 (-6.27; -0.45)**	**-3.06 (-5.98; -0.14)**	**-2.83 (-4.29; -1.37)**	**-2.61 (-4.49; -0.73)**	**-2.21 (-4.42; -0.00)**	-0.90 (-2.89; 1.09)	-0.90 (-2.44; 0.64)	PPI		
-21.0 (-23.3; -18.7)	**-8.16 (-14.4; -1.94)**	**-6.53 (-8.97; -4.09)**	**-6.50 (-11.3; -1.68)**	**-5.65 (-8.78; -2.52)**	**-4.98 (-6.59; -3.38)**	**-4.77 (-7.26; -2.28)**	**-4.74 (-8.16; -1.32)**	**-4.57 (-7.79; -1.35)**	**-4.27 (-7.50; -1.05)**	**-4.04 (-6.04; -2.04)**	**-3.82 (-6.15; -1.49)**	**-3.42 (-6.02; -0.82)**	-2.11 (-4.53; 0.31)	**-2.11 (-4.18; -0.05)**	-1.21 (-2.59; 0.17)	Banxia Xiexin Tang(SD)	

Note: RDQ, reflux disease questionnaire; NSD, non-syndrome differentiation; SD, syndrome differentiation; PPI, proton pump inhibitors. The values in each cell represent weighted mean differences (WMDs), and 95% confidence intervals(CIs) of the interventions at the top, compared to the comparators on the left. A negative WMD (<0) indicates a preference for the column intervention, suggesting that the column intervention performs better than the row intervention, with a higher RDQ score changes. Conversely, a positive WMD (>0) indicates a preference for the row intervention. Significant results are in bold and underlined.

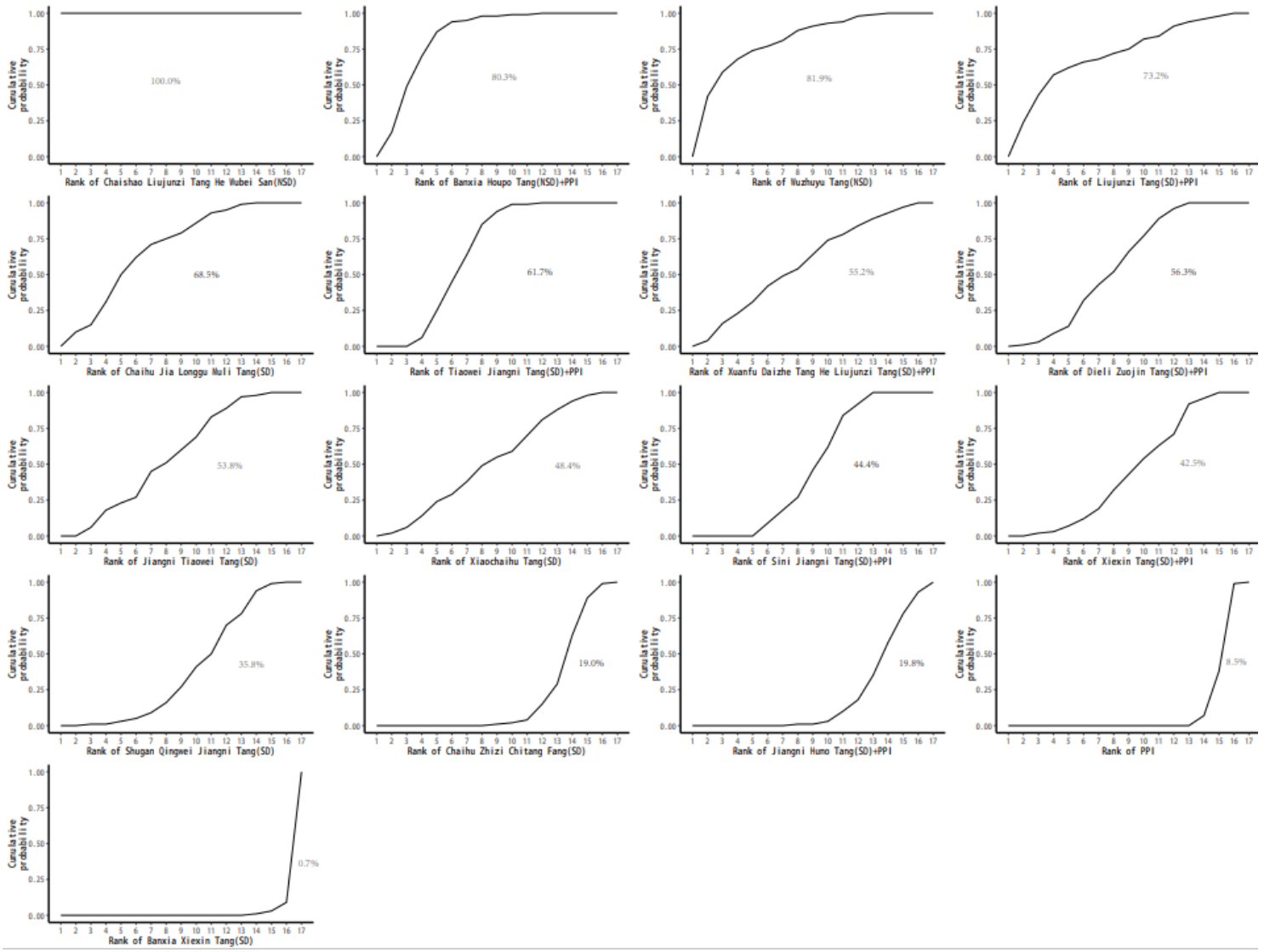
(2) 結局：**GERDQ** 評分變化

Banxia Xiexin Tang(SD)								
-0.01 (-2.15; 2.13)	Wuzhuyu Tang(NSD)							
-0.27 (-2.31; 1.77)	-0.26 (-2.72; 2.20)	Weikang Yihao Fang(SD)+PPI						
-0.63 (-2.54; 1.28)	-0.62 (-2.98; 1.74)	-0.36 (-2.63; 1.91)	Juanyu Jiangni Tang(SD)+PPI					
-0.94 (-2.23; 0.35)	-0.93 (-2.81; 0.95)	-0.67 (-2.44; 1.10)	-0.31 (-1.94; 1.32)	Banxia Houpo Tang(SD)+PPI				
-0.95 (-2.59; 0.69)	-0.94 (-3.08; 1.20)	-0.68 (-2.72; 1.36)	-0.32 (-2.24; 1.60)	-0.01 (-1.31; 1.29)	Zhuling Xiexin Tang(SD)+PPI			
-1.32 (-3.25; 0.61)	-1.31 (-3.68; 1.06)	-1.05 (-3.33; 1.23)	-0.69 (-2.86; 1.48)	-0.38 (-2.03; 1.27)	-0.37 (-2.31; 1.57)	Chaiban Tang(SD)+PPI		
-1.70 (-2.95; -0.45)	-1.69 (-3.55; 0.17)	-1.43 (-3.17; 0.31)	-1.07 (-2.66; 0.52)	**-0.76 (-1.49; -0.03)**	-0.75 (-2.01; 0.51)	-0.38 (-1.99; 1.23)	Xiaochaihu Tang(NSD)	
-2.13 (-3.29; -0.97)	**-2.12 (-3.92; -0.32)**	**-1.86 (-3.54; -0.18)**	-1.50 (-3.02; 0.02)	**-1.19 (-1.76; -0.62)**	**-1.18 (-2.35; -0.01)**	-0.81 (-2.36; 0.74)	-0.43 (-0.90; 0.04)	PPI

Note: GERDQ, gastroesophageal reflux disease questionnaire; NSD, non-syndrome differentiation; SD, syndrome differentiation; PPI, proton pump inhibitors. The values in each cell represent weighted mean differences (WMDs), and 95% confidence intervals(CIs) of the interventions at the top, compared to the comparators on the left. A negative WMD (<0) indicates a preference for the column intervention, suggesting that the column intervention performs better than the row intervention, with a higher GERDQ score changes. Conversely, a positive WMD (>0) indicates a preference for the row intervention. Significant results are in bold and underlined.

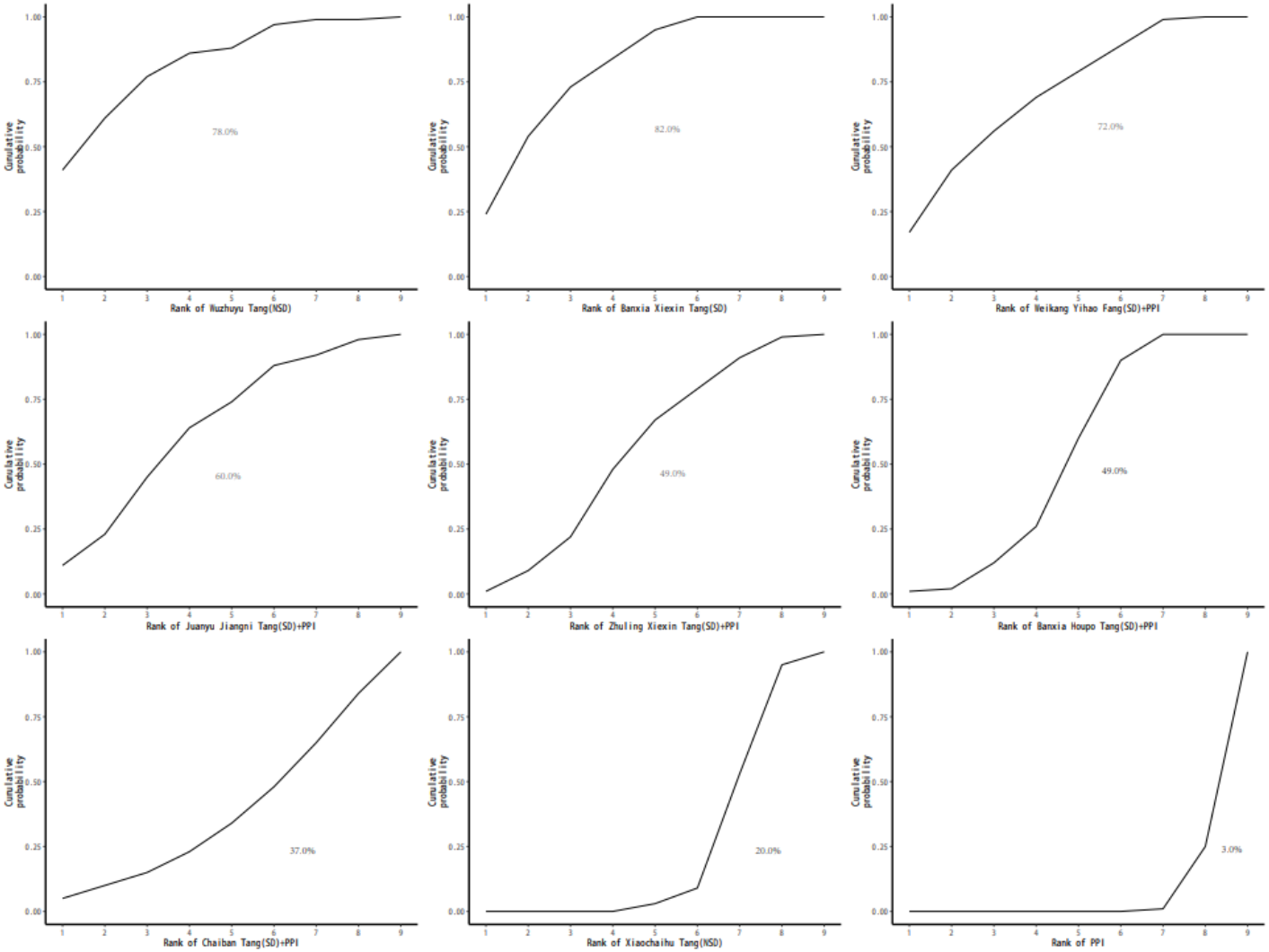
附錄 16 不同的中藥方及是否採用辨證論治原則在癥狀評分變化上的 SUCRA 排名

(1) 結局：**RDQ** 評分變化



Note: RDQ, reflux disease questionnaire; NSD, non-syndrome differentiation; SD, syndrome differentiation; PPI, proton pump inhibitors. The x-axis represents the possible rank of each intervention on the RDQ score changes(from the first best rank to the worst). The y-axis indicates the cumulative probability for each intervention to be the best intervention, the second-best intervention, the third best intervention, and so on. In this graphical approach, rankings are presented through examining the area under the curve. The bigger the area under the curve, the higher the likelihood that an intervention is in the top rank or one of the top ranks in achieving a higher RDQ score changes.

(2) 結局： **GERDQ** 評分變化



Note: GERDQ, gastroesophageal reflux disease questionnaire; NSD, non-syndrome differentiation; SD, syndrome differentiation; PPI, proton pump inhibitors. The x-axis represents the possible rank of each intervention on the GERDQ score changes (from the first best rank to the worst). The y-axis indicates the cumulative probability for each intervention to be the best intervention, the second-best intervention, the third best intervention, and so on. In this graphical approach, rankings are presented through examining the area under the curve. The bigger the area under the curve, the higher the likelihood that an intervention is in the top rank or one of the top ranks in achieving a higher GERDQ score changes.

附錄 17 不同的中藥方及是否採用辨證論治原則在 GERD 患者癥狀評分變化 NMA 效應估計的證據質量等級

(1) 結局：**RDQ** 評分變化

Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence
Chaishao Liujunzi Tang He Wubei San(NSD) vs Wuzhuyu Tang(NSD)	2 (282)	-	-	-12.9 (-19.2; -6.50)	Very low∆§	-12.9 (-19.2; -6.50)	Very low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Banxia Houpo Tang(NSD)+PPI	2 (272)	-	-	-14.5 (-17.2; -11.7)	Low∆	-14.5 (-17.2; -11.7)	Low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Liujunzi Tang(SD)+PPI	2 (252)	-	-	-14.5 (-19.5; -9.52)	Very low∆§	-14.5 (-19.5; -9.52)	Very low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Chaihu Jia Longgu Muli Tang(SD)	2 (252)	-	-	-15.4 (-18.7; -12.0)	Very low∆§	-15.4 (-18.7; -12.0)	Very low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Tiaowei Jiangni Tang(SD)+PPI	2 (252)	-	-	-16.0 (-18.1; -14.0)	Low∆	-16.0 (-18.1; -14.0)	Low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Dieli Zuojin Tang(SD)+PPI	2 (252)	-	-	-16.2 (-19.0; -13.4)	Low∆	-16.2 (-19.0; -13.4)	Low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI	2 (272)	-	-	-16.3 (-19.9; -12.6)	Low∆	-16.3 (-19.9; -12.6)	Low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Jiangni Tiaowei Tang(SD)	2 (252)	-	-	-16.4 (-19.9; -13.0)	Very low∆§	-16.4 (-19.9; -13.0)	Very low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Xiaochaihu Tang(SD)	2 (252)	-	-	-16.7 (-20.2; -13.3)	Low∆	-16.7 (-20.2; -13.3)	Low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Sini Jiangni Tang(SD)+PPI	2 (304)	-	-	-17.0 (-19.3; -14.6)	Low∆	-17.0 (-19.3; -14.6)	Low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Xiexin Tang(SD)+PPI	2 (270)	-	-	-17.2 (-19.8; -14.5)	Low∆	-17.2 (-19.8; -14.5)	Low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Shugan Qingwei Jiangni Tang(SD)	2 (262)	-	-	-17.6 (-20.5; -14.7)	Low∆	-17.6 (-20.5; -14.7)	Low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Jiangni Humo Tang(SD)+PPI	2 (256)	-	-	-18.9 (-21.6; -16.2)	Very low∆§	-18.9 (-21.6; -16.2)	Very low
Chaishao Liujunzi Tang He Wubei San(NSD) vs Chaihu Zhizi Chitang Fang(SD)	2 (272)	-	-	-18.9 (-21.3; -16.5)	Very low∆§	-18.9 (-21.3; -16.5)	Very low
Chaishao Liujunzi Tang He Wubei San(NSD) vs PPI	1 (192)	-19.8 (-21.7; -17.9)	Low*#	Not estimable‡	Not estimable‡	-19.8 (-21.7; -17.9)	Low

Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence
Chaishao Liujunzi Tang He Wubei San(NSD) vs Banxia Xiexin Tang(SD)	2 (319)	-	-	-21.0 (-23.3; -18.7)	Very low∆§	-21.0 (-23.3; -18.7)	Very low
Wuzhuyu Tang(NSD) vs Banxia Houpo Tang(NSD)+PPI	2 (170)	-	-	-1.63 (-8.02; 4.76)	Very low∆§	-1.63 (-8.02; 4.76)	Very low
Wuzhuyu Tang(NSD) vs Liujunzi Tang(SD)+PPI	2 (150)	-	-	-1.66 (-9.29; 5.97)	Low∆	-1.66 (-9.29; 5.97)	Low
Wuzhuyu Tang(NSD) vs Chaihu Jia Longgu Muli Tang(SD)	2 (150)	-	-	-2.51 (-9.20; 4.18)	Low∆	-2.51 (-9.20; 4.18)	Low
Wuzhuyu Tang(NSD) vs Tiaowei Jiangni Tang(SD)+PPI	2 (150)	-	-	-3.18 (-9.31; 2.95)	Very low∆§	-3.18 (-9.31; 2.95)	Very low
Wuzhuyu Tang(NSD) vs Dieli Zuojin Tang(SD)+PPI	2 (150)	-	-	-3.39 (-9.80; 3.02)	Very low∆§	-3.39 (-9.80; 3.02)	Very low
Wuzhuyu Tang(NSD) vs Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI	2 (170)	-	-	-3.42 (-10.2; 3.41)	Very low∆§	-3.42 (-10.2; 3.41)	Very low
Wuzhuyu Tang(NSD) vs Jiangni Tiaowei Tang(SD)	2 (150)	-	-	-3.59 (-10.3; 3.14)	Low∆	-3.59 (-10.3; 3.14)	Low
Wuzhuyu Tang(NSD) vs Xiaochaihu Tang(SD)	2 (150)	-	-	-3.89 (-10.6; 2.84)	Very low∆§	-3.89 (-10.6; 2.84)	Very low
Wuzhuyu Tang(NSD) vs Sini Jiangni Tang(SD)+PPI	2 (202)	-	-	-4.12 (-10.4; 2.12)	Very low∆§	-4.12 (-10.4; 2.12)	Very low
Wuzhuyu Tang(NSD) vs Xiexin Tang(SD)+PPI	2 (168)	-	-	-4.34 (-10.7; 2.01)	Very low∆§	-4.34 (-10.7; 2.01)	Very low
Wuzhuyu Tang(NSD) vs Shugan Qingwei Jiangni Tang(SD)	2 (160)	-	-	-4.74 (-11.2; 1.72)	Very low∆§	-4.74 (-11.2; 1.72)	Very low
Wuzhuyu Tang(NSD) vs Jiangni Humo Tang(SD)+PPI	2 (154)	-	-	-6.05 (-12.4; 0.34)	Low∆	-6.05 (-12.4; 0.34)	Low
Wuzhuyu Tang(NSD) vs Chaihu Zhizi Chitang Fang(SD)	2 (170)	-	-	-6.05 (-12.3; 0.21)	Very low∆§	-6.05 (-12.3; 0.21)	Very low
Wuzhuyu Tang(NSD) vs PPI	1 (90)	-6.95 (-13.0; -0.88)	Low*#	Not estimable‡	Not estimable‡	-6.95 (-13.0; -0.88)	Low
Wuzhuyu Tang(NSD) vs Banxia Xiexin Tang(SD)	2 (217)	-	-	-8.16 (-14.4; -1.94)	Very low∆§	-8.16 (-14.4; -1.94)	Very low
Banxia Houpo Tang(NSD)+PPI vs Liujunzi Tang(SD)+PPI	2 (140)	-	-	-0.03 (-5.07; 5.01)	Very low∆§	-0.03 (-5.07; 5.01)	Very low

Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence
Banxia Houpo Tang(NSD)+PPI vs Chaihu Jia Longgu Muli Tang(SD)	2 (140)	-	-	-0.88 (-4.34; 2.58)	Very low∆§	-0.88 (-4.34; 2.58)	Very low
Banxia Houpo Tang(NSD)+PPI vs Tiaowei Jiangni Tang(SD)+PPI	2 (140)	-	-	-1.55 (-3.72; 0.62)	Low∆	-1.55 (-3.72; 0.62)	Low
Banxia Houpo Tang(NSD)+PPI vs Dieli Zuojin Tang(SD)+PPI	2 (140)	-	-	-1.76 (-4.65; 1.13)	Low∆	-1.76 (-4.65; 1.13)	Low
Banxia Houpo Tang(NSD)+PPI vs Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI	2 (160)	-	-	-1.79 (-5.51; 1.93)	Low∆	-1.79 (-5.51; 1.93)	Low
Banxia Houpo Tang(NSD)+PPI vs Jiangni Tiaowei Tang(SD)	2 (140)	-	-	-1.96 (-5.50; 1.57)	Very low∆§	-1.96 (-5.50; 1.57)	Very low
Banxia Houpo Tang(NSD)+PPI vs Xiaochaihu Tang(SD)	2 (140)	-	-	-2.26 (-5.80; 1.28)	Low∆	-2.26 (-5.80; 1.28)	Low
Banxia Houpo Tang(NSD)+PPI vs Sini Jiangni Tang(SD)+PPI	2 (192)	-	-	-2.49 (-4.97; -0.01)	Low∆	-2.49 (-4.97; -0.01)	Low
Banxia Houpo Tang(NSD)+PPI vs Xiexin Tang(SD)+PPI	2 (158)	-	-	-2.71 (-5.46; 0.04)	Low∆	-2.71 (-5.46; 0.04)	Low
Banxia Houpo Tang(NSD)+PPI vs Shugan Qingwei Jiangni Tang(SD)	2 (150)	-	-	-3.11 (-6.09; -0.13)	Low∆	-3.11 (-6.09; -0.13)	Low
Banxia Houpo Tang(NSD)+PPI vs Jiangni Humo Tang(SD)+PPI	2 (144)	-	-	-4.42 (-7.25; -1.59)	Very low∆§	-4.42 (-7.25; -1.59)	Very low
Banxia Houpo Tang(NSD)+PPI vs Chaihu Zhizi Chitang Fang(SD)	2 (160)	-	-	-4.42 (-6.95; -1.89)	Very low∆§	-4.42 (-6.95; -1.89)	Very low
Banxia Houpo Tang(NSD)+PPI vs PPI	1 (80)	-5.32 (-7.33; -3.31)	Low*#	Not estimable‡	Not estimable‡	-5.32 (-7.33; -3.31)	Low
Banxia Houpo Tang(NSD)+PPI vs Banxia Xiexin Tang(SD)	2 (207)	-	-	-6.53 (-8.97; -4.09)	Very low∆§	-6.53 (-8.97; -4.09)	Very low
Liujunzi Tang(SD)+PPI vs Chaihu Jia Longgu Muli Tang(SD)	2 (120)	-	-	-0.85 (-6.26; 4.56)	Low∆	-0.85 (-6.26; 4.56)	Low
Liujunzi Tang(SD)+PPI vs Tiaowei Jiangni Tang(SD)+PPI	2 (120)	-	-	-1.52 (-6.22; 3.18)	Very low∆§	-1.52 (-6.22; 3.18)	Very low
Liujunzi Tang(SD)+PPI vs Dieli Zuojin Tang(SD)+PPI	2 (120)	-	-	-1.73 (-6.80; 3.34)	Very low∆§	-1.73 (-6.80; 3.34)	Very low
Liujunzi Tang(SD)+PPI vs Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI	2 (140)	-	-	-1.76 (-7.34; 3.82)	Very low∆§	-1.76 (-7.34; 3.82)	Very low

Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence
Liujunzi Tang(SD)+PPI vs Jiangni Tiaowei Tang(SD)	2 (120)	-	-	-1.93 (-7.40; 3.53)	LowΔ	-1.93 (-7.40; 3.53)	Low
Liujunzi Tang(SD)+PPI vs Xiaochaihu Tang(SD)	2 (120)	-	-	-2.23 (-7.70; 3.24)	Very lowΔ§	-2.23 (-7.70; 3.24)	Very low
Liujunzi Tang(SD)+PPI vs Sini Jiangni Tang(SD)+PPI	2 (172)	-	-	-2.46 (-7.31; 2.39)	Very lowΔ§	-2.46 (-7.31; 2.39)	Very low
Liujunzi Tang(SD)+PPI vs Xiexin Tang(SD)+PPI	2 (138)	-	-	-2.68 (-7.67; 2.31)	Very lowΔ§	-2.68 (-7.67; 2.31)	Very low
Liujunzi Tang(SD)+PPI vs Shugan Qingwei Jiangni Tang(SD)	2 (130)	-	-	-3.08 (-8.20; 2.04)	Very lowΔ§	-3.08 (-8.20; 2.04)	Very low
Liujunzi Tang(SD)+PPI vs Jiangni Humo Tang(SD)+PPI	2 (124)	-	-	-4.39 (-9.42; 0.64)	LowΔ	-4.39 (-9.42; 0.64)	Low
Liujunzi Tang(SD)+PPI vs Chaihu Zhizi Chitang Fang(SD)	2 (140)	-	-	-4.39 (-9.26; 0.48)	Very lowΔ§	-4.39 (-9.26; 0.48)	Very low
Liujunzi Tang(SD)+PPI vs PPI	1 (60)	-5.29 (-9.91; -0.67)	Low*#	Not estimable‡	Not estimable‡	-5.29 (-9.91; -0.67)	Low
Liujunzi Tang(SD)+PPI vs Banxia Xiexin Tang(SD)	2 (187)	-	-	-6.50 (-11.3; -1.68)	Very lowΔ§	-6.50 (-11.3; -1.68)	Very low
Chaihu Jia Longgu Muli Tang(SD) vs Tiaowei Jiangni Tang(SD)+PPI	2 (120)	-	-	-0.67 (-3.60; 2.26)	Very lowΔ§	-0.67 (-3.60; 2.26)	Very low
Chaihu Jia Longgu Muli Tang(SD) vs Dieli Zuojin Tang(SD)+PPI	2 (120)	-	-	-0.88 (-4.38; 2.62)	Very lowΔ§	-0.88 (-4.38; 2.62)	Very low
Chaihu Jia Longgu Muli Tang(SD) vs Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI	2 (140)	-	-	-0.91 (-5.12; 3.30)	Very lowΔ§	-0.91 (-5.12; 3.30)	Very low
Chaihu Jia Longgu Muli Tang(SD) vs Jiangni Tiaowei Tang(SD)	2 (120)	-	-	-1.08 (-5.13; 2.96)	LowΔ	-1.08 (-5.13; 2.96)	Low
Chaihu Jia Longgu Muli Tang(SD) vs Xiaochaihu Tang(SD)	2 (120)	-	-	-1.38 (-5.43; 2.67)	Very lowΔ§	-1.38 (-5.43; 2.67)	Very low
Chaihu Jia Longgu Muli Tang(SD) vs Sini Jiangni Tang(SD)+PPI	2 (172)	-	-	-1.61 (-4.78; 1.56)	Very lowΔ§	-1.61 (-4.78; 1.56)	Very low
Chaihu Jia Longgu Muli Tang(SD) vs Xiexin Tang(SD)+PPI	2 (138)	-	-	-1.83 (-5.21; 1.55)	Very lowΔ§	-1.83 (-5.21; 1.55)	Very low
Chaihu Jia Longgu Muli Tang(SD) vs Shugan Qingwei Jiangni Tang(SD)	2 (130)	-	-	-2.23 (-5.81; 1.35)	Very lowΔ§	-2.23 (-5.81; 1.35)	Very low

Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence
Chaihu Jia Longgu Muli Tang(SD) vs Jiangni Humo Tang(SD)+PPI	2 (124)	-	-	-3.54 (-6.99; -0.09)	LowΔ	-3.54 (-6.99; -0.09)	Low
Chaihu Jia Longgu Muli Tang(SD) vs Chaihu Zhizi Chitang Fang(SD)	2 (140)	-	-	-3.54 (-6.75; -0.33)	Very lowΔ§	-3.54 (-6.75; -0.33)	Very low
Chaihu Jia Longgu Muli Tang(SD) vs PPI	1 (60)	-4.44 (-7.25; -1.63)	Low*#	Not estimable‡	Not estimable‡	-4.44 (-7.25; -1.63)	Low
Chaihu Jia Longgu Muli Tang(SD) vs Banxia Xiexin Tang(SD)	2 (187)	-	-	-5.65 (-8.78; -2.52)	Very lowΔ§	-5.65 (-8.78; -2.52)	Very low
Tiaowei Jiangni Tang(SD)+PPI vs Dieli Zuojin Tang(SD)+PPI	2 (120)	-	-	-0.21 (-2.44; 2.02)	LowΔ	-0.21 (-2.44; 2.02)	Low
Tiaowei Jiangni Tang(SD)+PPI vs Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI	2 (140)	-	-	-0.24 (-3.48; 3.00)	LowΔ	-0.24 (-3.48; 3.00)	Low
Tiaowei Jiangni Tang(SD)+PPI vs Jiangni Tiaowei Tang(SD)	2 (120)	-	-	-0.41 (-3.44; 2.61)	Very lowΔ§	-0.41 (-3.44; 2.61)	Very low
Tiaowei Jiangni Tang(SD)+PPI vs Xiaochaihu Tang(SD)	2 (120)	-	-	-0.71 (-3.74; 2.32)	LowΔ	-0.71 (-3.74; 2.32)	Low
Tiaowei Jiangni Tang(SD)+PPI vs Sini Jiangni Tang(SD)+PPI	2 (172)	-	-	-0.94 (-2.61; 0.73)	LowΔ	-0.94 (-2.61; 0.73)	Low
Tiaowei Jiangni Tang(SD)+PPI vs Xiexin Tang(SD)+PPI	2 (138)	-	-	-1.16 (-3.21; 0.89)	LowΔ	-1.16 (-3.21; 0.89)	Low
Tiaowei Jiangni Tang(SD)+PPI vs Shugan Qingwei Jiangni Tang(SD)	2 (130)	-	-	-1.56 (-3.92; 0.80)	LowΔ	-1.56 (-3.92; 0.80)	Low
Tiaowei Jiangni Tang(SD)+PPI vs Jiangni Humo Tang(SD)+PPI	2 (124)	-	-	-2.87 (-5.03; -0.72)	Very lowΔ§	-2.87 (-5.03; -0.72)	Very low
Tiaowei Jiangni Tang(SD)+PPI vs Chaihu Zhizi Chitang Fang(SD)	2 (140)	-	-	-2.87 (-4.62; -1.12)	Very lowΔ§	-2.87 (-4.62; -1.12)	Very low
Tiaowei Jiangni Tang(SD)+PPI vs PPI	1 (60)	-3.77 (-4.60; -2.95)	Low*#	Not estimable‡	Not estimable‡	-3.77 (-4.60; -2.95)	Low
Tiaowei Jiangni Tang(SD)+PPI vs Banxia Xiexin Tang(SD)	2 (187)	-	-	-4.98 (-6.59; -3.38)	Very lowΔ§	-4.98 (-6.59; -3.38)	Very low
Dieli Zuojin Tang(SD)+PPI vs Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI	2 (140)	-	-	-0.03 (-3.79; 3.73)	LowΔ	-0.03 (-3.79; 3.73)	Low
Dieli Zuojin Tang(SD)+PPI vs Jiangni Tiaowei Tang(SD)	2 (120)	-	-	-0.20 (-3.78; 3.37)	Very lowΔ§	-0.20 (-3.78; 3.37)	Very low

Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence
Dieli Zuojin Tang(SD)+PPI vs Xiaochaihu Tang(SD)	2 (120)	-	-	-0.50 (-4.08; 3.08)	LowΔ	-0.50 (-4.08; 3.08)	Low
Dieli Zuojin Tang(SD)+PPI vs Sini Jiangni Tang(SD)+PPI	2 (172)	-	-	-0.73 (-3.27; 1.81)	LowΔ	-0.73 (-3.27; 1.81)	Low
Dieli Zuojin Tang(SD)+PPI vs Xiexin Tang(SD)+PPI	2 (138)	-	-	-0.95 (-3.75; 1.85)	LowΔ	-0.95 (-3.75; 1.85)	Low
Dieli Zuojin Tang(SD)+PPI vs Shugan Qingwei Jiangni Tang(SD)	2 (130)	-	-	-1.35 (-4.38; 1.68)	LowΔ	-1.35 (-4.38; 1.68)	Low
Dieli Zuojin Tang(SD)+PPI vs Jiangni Humo Tang(SD)+PPI	2 (124)	-	-	-2.66 (-5.54; 0.22)	Very lowΔ§	-2.66 (-5.54; 0.22)	Very low
Dieli Zuojin Tang(SD)+PPI vs Chaihu Zhizi Chitang Fang(SD)	2 (140)	-	-	-2.66 (-5.24; -0.08)	Very lowΔ§	-2.66 (-5.24; -0.08)	Very low
Dieli Zuojin Tang(SD)+PPI vs PPI	1 (60)	-3.56 (-5.64; -1.49)	Low*#	Not estimable‡	Not estimable‡	-3.56 (-5.64; -1.49)	Low
Dieli Zuojin Tang(SD)+PPI vs Banxia Xiexin Tang(SD)	2 (187)	-	-	-4.77 (-7.26; -2.28)	Very lowΔ§	-4.77 (-7.26; -2.28)	Very low
Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI vs Jiangni Tiaowei Tang(SD)	2 (140)	-	-	-0.17 (-4.45; 4.10)	Very lowΔ§	-0.17 (-4.45; 4.10)	Very low
Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI vs Xiaochaihu Tang(SD)	2 (140)	-	-	-0.47 (-4.75; 3.81)	LowΔ	-0.47 (-4.75; 3.81)	Low
Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI vs Sini Jiangni Tang(SD)+PPI	2 (192)	-	-	-0.70 (-4.15; 2.75)	LowΔ	-0.70 (-4.15; 2.75)	Low
Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI vs Xiexin Tang(SD)+PPI	2 (158)	-	-	-0.92 (-4.57; 2.73)	LowΔ	-0.92 (-4.57; 2.73)	Low
Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI vs Shugan Qingwei Jiangni Tang(SD)	2 (150)	-	-	-1.32 (-5.15; 2.51)	LowΔ	-1.32 (-5.15; 2.51)	Low
Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI vs Jiangni Humo Tang(SD)+PPI	2 (144)	-	-	-2.63 (-6.34; 1.08)	Very lowΔ§	-2.63 (-6.34; 1.08)	Very low
Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI vs Chaihu Zhizi Chitang Fang(SD)	2 (160)	-	-	-2.63 (-6.12; 0.86)	Very lowΔ§	-2.63 (-6.12; 0.86)	Very low
Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI vs PPI	1 (80)	-3.53 (-6.66; -0.40)	Low*#	Not estimable‡	Not estimable‡	-3.53 (-6.66; -0.40)	Low
Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI vs Banxia Xiexin Tang(SD)	2 (207)	-	-	-4.74 (-8.16; -1.32)	Very lowΔ§	-4.74 (-8.16; -1.32)	Very low

Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence
Jiangni Tiaowei Tang(SD) vs Xiaochaihu Tang(SD)	2 (120)	-	-	-0.30 (-4.42; 3.82)	Very low∆§	-0.30 (-4.42; 3.82)	Very low
Jiangni Tiaowei Tang(SD) vs Sini Jiangni Tang(SD)+PPI	2 (172)	-	-	-0.53 (-3.78; 2.73)	Very low∆§	-0.53 (-3.78; 2.73)	Very low
Jiangni Tiaowei Tang(SD) vs Xiexin Tang(SD)+PPI	2 (138)	-	-	-0.75 (-4.21; 2.72)	Very low∆§	-0.75 (-4.21; 2.72)	Very low
Jiangni Tiaowei Tang(SD) vs Shugan Qingwei Jiangni Tang(SD)	2 (130)	-	-	-1.15 (-4.80; 2.51)	Very low∆§	-1.15 (-4.80; 2.51)	Very low
Jiangni Tiaowei Tang(SD) vs Jiangni Humo Tang(SD)+PPI	2 (124)	-	-	-2.46 (-5.98; 1.07)	Low∆	-2.46 (-5.98; 1.07)	Low
Jiangni Tiaowei Tang(SD) vs Chaihu Zhizi Chitang Fang(SD)	2 (140)	-	-	-2.46 (-5.75; 0.84)	Very low∆§	-2.46 (-5.75; 0.84)	Very low
Jiangni Tiaowei Tang(SD) vs PPI	1 (60)	-3.36 (-6.27; -0.45)	Low*#	Not estimable‡	Not estimable‡	-3.36 (-6.27; -0.45)	Low
Jiangni Tiaowei Tang(SD) vs Banxia Xiexin Tang(SD)	2 (187)	-	-	-4.57 (-7.79; -1.35)	Very low∆§	-4.57 (-7.79; -1.35)	Very low
Xiaochaihu Tang(SD) vs Sini Jiangni Tang(SD)+PPI	2 (172)	-	-	-0.23 (-3.49; 3.03)	Low∆	-0.23 (-3.49; 3.03)	Low
Xiaochaihu Tang(SD) vs Xiexin Tang(SD)+PPI	2 (138)	-	-	-0.45 (-3.92; 3.02)	Low∆	-0.45 (-3.92; 3.02)	Low
Xiaochaihu Tang(SD) vs Shugan Qingwei Jiangni Tang(SD)	2 (130)	-	-	-0.85 (-4.51; 2.81)	Low∆	-0.85 (-4.51; 2.81)	Low
Xiaochaihu Tang(SD) vs Jiangni Humo Tang(SD)+PPI	2 (124)	-	-	-2.16 (-5.69; 1.37)	Very low∆§	-2.16 (-5.69; 1.37)	Very low
Xiaochaihu Tang(SD) vs Chaihu Zhizi Chitang Fang(SD)	2 (140)	-	-	-2.16 (-5.46; 1.14)	Very low∆§	-2.16 (-5.46; 1.14)	Very low
Xiaochaihu Tang(SD) vs PPI	1 (60)	-3.06 (-5.98; -0.14)	Low*#	Not estimable‡	Not estimable‡	-3.06 (-5.98; -0.14)	Low
Xiaochaihu Tang(SD) vs Banxia Xiexin Tang(SD)	2 (187)	-	-	-4.27 (-7.50; -1.05)	Very low∆§	-4.27 (-7.50; -1.05)	Very low
Sini Jiangni Tang(SD)+PPI vs Xiexin Tang(SD)+PPI	2 (190)	-	-	-0.22 (-2.60; 2.16)	Low∆	-0.22 (-2.60; 2.16)	Low
Sini Jiangni Tang(SD)+PPI vs Shugan Qingwei Jiangni Tang(SD)	2 (182)	-	-	-0.62 (-3.26; 2.02)	Low∆	-0.62 (-3.26; 2.02)	Low

Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence
Sini Jiangni Tang(SD)+PPI vs Jiangni Humo Tang(SD)+PPI	2 (176)	-	-	-1.93 (-4.40; 0.54)	Very low∆§	-1.93 (-4.40; 0.54)	Very low
Sini Jiangni Tang(SD)+PPI vs Chaihu Zhizi Chitang Fang(SD)	2 (192)	-	-	-1.93 (-4.05; 0.19)	Very low∆§	-1.93 (-4.05; 0.19)	Very low
Sini Jiangni Tang(SD)+PPI vs PPI	1 (112)	-2.83 (-4.29; -1.37)	Low*#	Not estimable‡	Not estimable‡	-2.83 (-4.29; -1.37)	Low
Sini Jiangni Tang(SD)+PPI vs Banxia Xiexin Tang(SD)	2 (239)	-	-	-4.04 (-6.04; -2.04)	Very low∆§	-4.04 (-6.04; -2.04)	Very low
Xiexin Tang(SD)+PPI vs Shugan Qingwei Jiangni Tang(SD)	2 (148)	-	-	-0.40 (-3.30; 2.50)	Low∆	-0.40 (-3.30; 2.50)	Low
Xiexin Tang(SD)+PPI vs Jiangni Humo Tang(SD)+PPI	2 (142)	-	-	-1.71 (-4.45; 1.03)	Very low∆§	-1.71 (-4.45; 1.03)	Very low
Xiexin Tang(SD)+PPI vs Chaihu Zhizi Chitang Fang(SD)	2 (158)	-	-	-1.71 (-4.14; 0.72)	Very low∆§	-1.71 (-4.14; 0.72)	Very low
Xiexin Tang(SD)+PPI vs PPI	1 (78)	-2.61 (-4.49; -0.73)	Low*#	Not estimable‡	Not estimable‡	-2.61 (-4.49; -0.73)	Low
Xiexin Tang(SD)+PPI vs Banxia Xiexin Tang(SD)	2 (205)	-	-	-3.82 (-6.15; -1.49)	Very low∆§	-3.82 (-6.15; -1.49)	Very low
Shugan Qingwei Jiangni Tang(SD) vs Jiangni Humo Tang(SD)+PPI	2 (134)	-	-	-1.31 (-4.28; 1.66)	Very low∆§	-1.31 (-4.28; 1.66)	Very low
Shugan Qingwei Jiangni Tang(SD) vs Chaihu Zhizi Chitang Fang(SD)	2 (150)	-	-	-1.31 (-4.00; 1.38)	Very low∆§	-1.31 (-4.00; 1.38)	Very low
Shugan Qingwei Jiangni Tang(SD) vs PPI	1 (70)	-2.21 (-4.42; -0.00)	Low*#	Not estimable‡	Not estimable‡	-2.21 (-4.42; -0.00)	Low
Shugan Qingwei Jiangni Tang(SD) vs Banxia Xiexin Tang(SD)	2 (197)	-	-	-3.42 (-6.02; -0.82)	Very low∆§	-3.42 (-6.02; -0.82)	Very low
Jiangni Humo Tang(SD)+PPI vs Chaihu Zhizi Chitang Fang(SD)	2 (144)	-	-	-0.00 (-2.52; 2.52)	Very low∆§	-0.00 (-2.52; 2.52)	Very low
Jiangni Humo Tang(SD)+PPI vs PPI	1 (64)	-0.90 (-2.89; 1.09)	Low*#	Not estimable‡	Not estimable‡	-0.90 (-2.89; 1.09)	Low
Jiangni Humo Tang(SD)+PPI vs Banxia Xiexin Tang(SD)	2 (191)	-	-	-2.11 (-4.53; 0.31)	Very low∆§	-2.11 (-4.53; 0.31)	Very low
Chaihu Zhizi Chitang Fang(SD) vs PPI	1 (80)	-0.90 (-2.44; 0.64)	Low*#	Not estimable‡	Not estimable‡	-0.90 (-2.44; 0.64)	Low

Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence
Chaihu Zhizi Chitang Fang(SD) vs Banxia Xiexin Tang(SD)	2 (207)	-	-	-2.11 (-4.18; -0.05)	LowΔ	-2.11 (-4.18; -0.05)	Low
PPI vs Banxia Xiexin Tang(SD)	1 (127)	-1.21 (-2.59; 0.17)	Low*#	Not estimable‡	Not estimable‡	-1.21 (-2.59; 0.17)	Low

Note: NSD, Non-Syndrome Differentiation; SD, Syndrome Differentiation; RCT, Randomized Controlled Trial; WMD, Weighted Mean Difference; CI: confidence interval; PPI, Proton Pump Inhibitors. -: Not applicable. *Serious study limitations (crucial risk of bias for one criterion, or multiple criteria, and likely to seriously alter the results); # Serious impression (small sample size less than 400); ΔContributing direct evidence of low quality; ‡: Cannot be estimated because the drug was not connected in a loop in the evidence network; §Indirectness because of questionable comparability of trial populations or because of intransitivity.

(2) 結局：**GERDQ** 評分變化

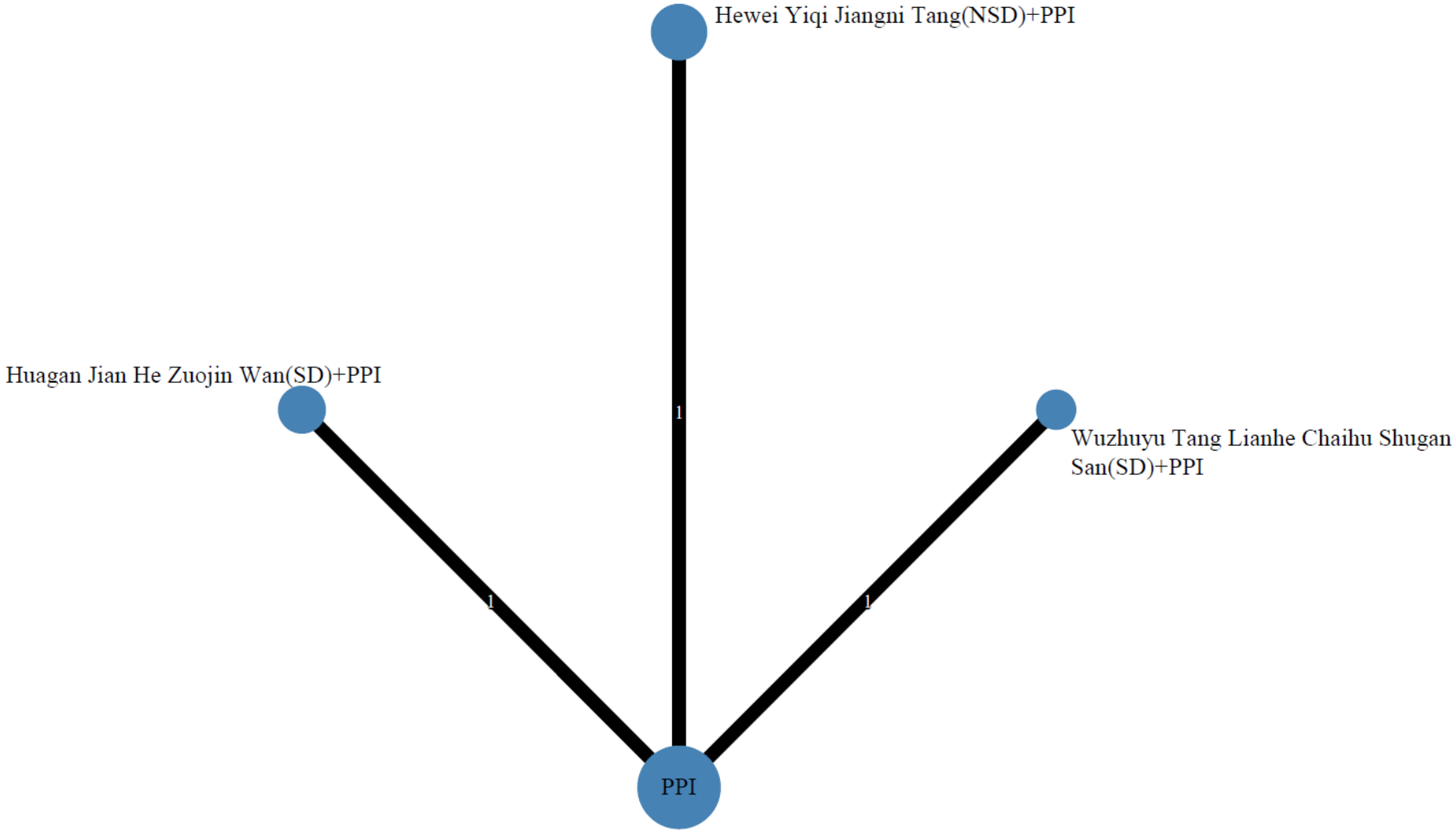
Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence
Banxia Xiexin Tang(SD) vs Wuzhuyu Tang(NSD)	2 (150)	-	-	-0.01 (-2.15; 2.13)	Very low∆§	-0.01 (-2.15; 2.13)	Very low
Banxia Xiexin Tang(SD) vs Weikang Yihao Fang(SD)+PPI	2 (120)	-	-	-0.27 (-2.31; 1.77)	Very low∆§	-0.27 (-2.31; 1.77)	Very low
Banxia Xiexin Tang(SD) vs Juanyu Jiangni Tang(SD)+PPI	2 (120)	-	-	-0.63 (-2.54; 1.28)	Very low∆§	-0.63 (-2.54; 1.28)	Very low
Banxia Xiexin Tang(SD) vs Banxia Houpo Tang(SD)+PPI	2 (144)	-	-	-0.94 (-2.23; 0.35)	Very low∆§	-0.94 (-2.23; 0.35)	Very low
Banxia Xiexin Tang(SD) vs Zhuling Xiexin Tang(SD)+PPI	2 (132)	-	-	-0.95 (-2.59; 0.69)	Very low∆§	-0.95 (-2.59; 0.69)	Very low
Banxia Xiexin Tang(SD) vs Chaiban Tang(SD)+PPI	2 (136)	-	-	-1.32 (-3.25; 0.61)	Very low∆§	-1.32 (-3.25; 0.61)	Very low
Banxia Xiexin Tang(SD) vs Xiaochaihu Tang(NSD)	2 (348)	-	-	-1.70 (-2.95; -0.45)	Very low∆§	-1.70 (-2.95; -0.45)	Very low
Banxia Xiexin Tang(SD) vs PPI	1 (60)	-2.13 (-3.29; -0.97)	Low*#	Not estimable‡	Not estimable‡	-2.13 (-3.29; -0.97)	Low
Wuzhuyu Tang(NSD) vs Weikang Yihao Fang(SD)+PPI	2 (150)	-	-	-0.26 (-2.72; 2.20)	Low∆	-0.26 (-2.72; 2.20)	Low
Wuzhuyu Tang(NSD) vs Juanyu Jiangni Tang(SD)+PPI	2 (150)	-	-	-0.62 (-2.98; 1.74)	Low∆	-0.62 (-2.98; 1.74)	Low
Wuzhuyu Tang(NSD) vs Banxia Houpo Tang(SD)+PPI	2 (174)	-	-	-0.93 (-2.81; 0.95)	Very low∆§	-0.93 (-2.81; 0.95)	Very low
Wuzhuyu Tang(NSD) vs Zhuling Xiexin Tang(SD)+PPI	2 (162)	-	-	-0.94 (-3.08; 1.20)	Low∆	-0.94 (-3.08; 1.20)	Low
Wuzhuyu Tang(NSD) vs Chaiban Tang(SD)+PPI	2 (166)	-	-	-1.31 (-3.68; 1.06)	Low∆	-1.31 (-3.68; 1.06)	Low
Wuzhuyu Tang(NSD) vs Xiaochaihu Tang(NSD)	2 (378)	-	-	-1.69 (-3.55; 0.17)	Very low∆§	-1.69 (-3.55; 0.17)	Very low
Wuzhuyu Tang(NSD) vs PPI	1 (90)	-2.12 (-3.92; -0.32)	Low*#	Not estimable‡	Not estimable‡	-2.12 (-3.92; -0.32)	Low
Weikang Yihao Fang(SD)+PPI vs Juanyu Jiangni Tang(SD)+PPI	2 (120)	-	-	-0.36 (-2.63; 1.91)	Low∆	-0.36 (-2.63; 1.91)	Low

Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence
Weikang Yihao Fang(SD)+PPI vs Banxia Houpo Tang(SD)+PPI	2 (144)	-	-	-0.67 (-2.44; 1.10)	Very low∆§	-0.67 (-2.44; 1.10)	Very low
Weikang Yihao Fang(SD)+PPI vs Zhuling Xiexin Tang(SD)+PPI	2 (132)	-	-	-0.68 (-2.72; 1.36)	Low∆	-0.68 (-2.72; 1.36)	Low
Weikang Yihao Fang(SD)+PPI vs Chaiban Tang(SD)+PPI	2 (136)	-	-	-1.05 (-3.33; 1.23)	Low∆	-1.05 (-3.33; 1.23)	Low
Weikang Yihao Fang(SD)+PPI vs Xiaochaihu Tang(NSD)	2 (348)	-	-	-1.43 (-3.17; 0.31)	Very low∆§	-1.43 (-3.17; 0.31)	Very low
Weikang Yihao Fang(SD)+PPI vs PPI	1 (60)	-1.86 (-3.54; -0.18)	Low*#	Not estimable‡	Not estimable‡	-1.86 (-3.54; -0.18)	Low
Juanyu Jiangni Tang(SD)+PPI vs Banxia Houpo Tang(SD)+PPI	2 (144)	-	-	-0.31 (-1.94; 1.32)	Very low∆§	-0.31 (-1.94; 1.32)	Very low
Juanyu Jiangni Tang(SD)+PPI vs Zhuling Xiexin Tang(SD)+PPI	2 (132)	-	-	-0.32 (-2.24; 1.60)	Low∆	-0.32 (-2.24; 1.60)	Low
Juanyu Jiangni Tang(SD)+PPI vs Chaiban Tang(SD)+PPI	2 (136)	-	-	-0.69 (-2.86; 1.48)	Low∆	-0.69 (-2.86; 1.48)	Low
Juanyu Jiangni Tang(SD)+PPI vs Xiaochaihu Tang(NSD)	2 (348)	-	-	-1.07 (-2.66; 0.52)	Very low∆§	-1.07 (-2.66; 0.52)	Very low
Juanyu Jiangni Tang(SD)+PPI vs PPI	1 (60)	-1.50 (-3.02; 0.02)	Low*#	Not estimable‡	Not estimable‡	-1.50 (-3.02; 0.02)	Low
Banxia Houpo Tang(SD)+PPI vs Zhuling Xiexin Tang(SD)+PPI	2 (156)	-	-	-0.01 (-1.31; 1.29)	Very low∆§	-0.01 (-1.31; 1.29)	Very low
Banxia Houpo Tang(SD)+PPI vs Chaiban Tang(SD)+PPI	2 (160)	-	-	-0.38 (-2.03; 1.27)	Very low∆§	-0.38 (-2.03; 1.27)	Very low
Banxia Houpo Tang(SD)+PPI vs Xiaochaihu Tang(NSD)	2 (372)	-	-	-0.76 (-1.49; -0.03)	Low∆	-0.76 (-1.49; -0.03)	Low
Banxia Houpo Tang(SD)+PPI vs PPI	1 (84)	-1.19 (-1.76; -0.62)	Low*#	Not estimable‡	Not estimable‡	-1.19 (-1.76; -0.62)	Low
Zhuling Xiexin Tang(SD)+PPI vs Chaiban Tang(SD)+PPI	2 (148)	-	-	-0.37 (-2.31; 1.57)	Low∆	-0.37 (-2.31; 1.57)	Low
Zhuling Xiexin Tang(SD)+PPI vs Xiaochaihu Tang(NSD)	2 (360)	-	-	-0.75 (-2.01; 0.51)	Very low∆§	-0.75 (-2.01; 0.51)	Very low
Zhuling Xiexin Tang(SD)+PPI vs PPI	1 (72)	-1.18 (-2.35; -0.01)	Low*#	Not estimable‡	Not estimable‡	-1.18 (-2.35; -0.01)	Low

Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence	WMD (95%CI)	Quality of evidence
Chaiban Tang(SD)+PPI vs Xiaochaihu Tang(NSD)	2 (364)	-	-	-0.38 (-1.99; 1.23)	LowΔ	-0.38 (-1.99; 1.23)	Low
Chaiban Tang(SD)+PPI vs PPI	1 (76)	-0.81 (-2.36; 0.74)	Low*#	Not estimable‡	Not estimable‡	-0.81 (-2.36; 0.74)	Low
Xiaochaihu Tang(NSD) vs PPI	1 (288)	-0.43 (-0.90; 0.04)	Low*#	Not estimable‡	Not estimable‡	-0.43 (-0.90; 0.04)	Low

Note: NSD, Non-Syndrome Differentiation; SD, Syndrome Differentiation; RCT, Randomized Controlled Trial; WMD, Weighted Mean Difference; CI: confidence interval; PPI, Proton Pump Inhibitors. -: Not applicable. *Serious study limitations (crucial risk of bias for one criterion, or multiple criteria, and likely to seriously alter the results); # Serious impression (small sample size less than 400); ΔContributing direct evidence of low quality; ‡: Cannot be estimated because the drug was not connected in a loop in the evidence network; §Indirectness because of questionable comparability of trial populations or because of intransitivity.

附錄 18 不同中藥方及是否採用辨證論治幹預模式在 GERD 患者復發率上的網絡幾何圖



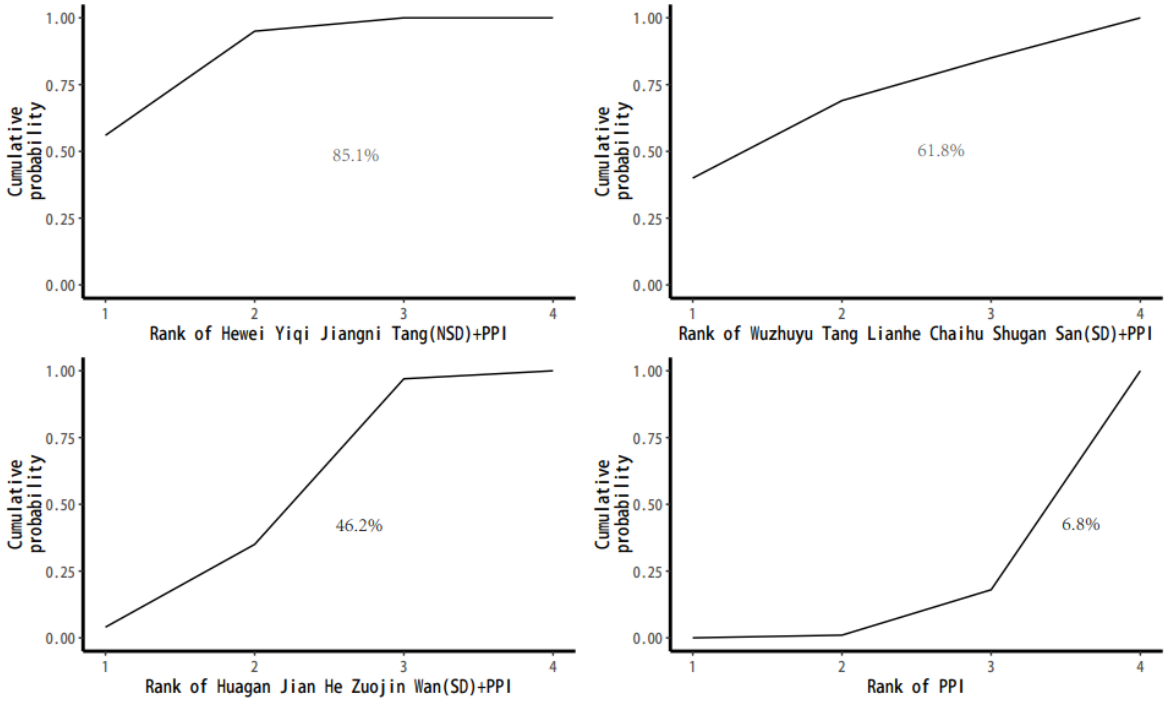
Notes: NSD, non-syndrome differentiation; SD, syndrome differentiation; PPI, proton pump inhibitors. Nodes represent the interventions, node sizes correspond to the number of studies involved, lines connecting nodes represent direct comparisons between pairs of interventions. Width of the lines represents the proportion of the number of trials for each comparison as compared to total number of trials. Number on the line represents the count of the direct comparisons.

附錄 **19** 不同的中藥方及是否採用辨證論治原則在 **GERD** 患者復發率上兩兩比較的結果

Hewei Yiqi Jiangni Tang(NSD) +PPI			
0.60 (0.05; 7.35)	Wuzhuyu Tang Lianhe Chaihu Shugan San(SD) +PPI		
0.38 (0.09; 1.52)	0.63 (0.06; 6.42)	Huagan Jian He Zuojin Wan(SD) +PPI	
<u>0.20 (0.06; 0.66)</u>	0.33 (0.04; 3.03)	0.53 (0.25; 1.13)	PPI

Note: NSD, non-syndrome differentiation; SD, syndrome differentiation; PPI, proton pump inhibitors. The values in each cell represent risk ratios (RRs), and 95% confidence intervals of the intervention at the top, compared to the comparator on the left. An RR less than 1 suggests a preference for the column intervention, indicating superior performance compared to the row intervention, with a lower recurrence rate for patients with gastroesophageal reflux disease. Significant results are in bold and underlined.

附錄 20 不同的中藥方及是否採用辨證論治原則在 GERD 患者復發率上的 SUCRA 結果



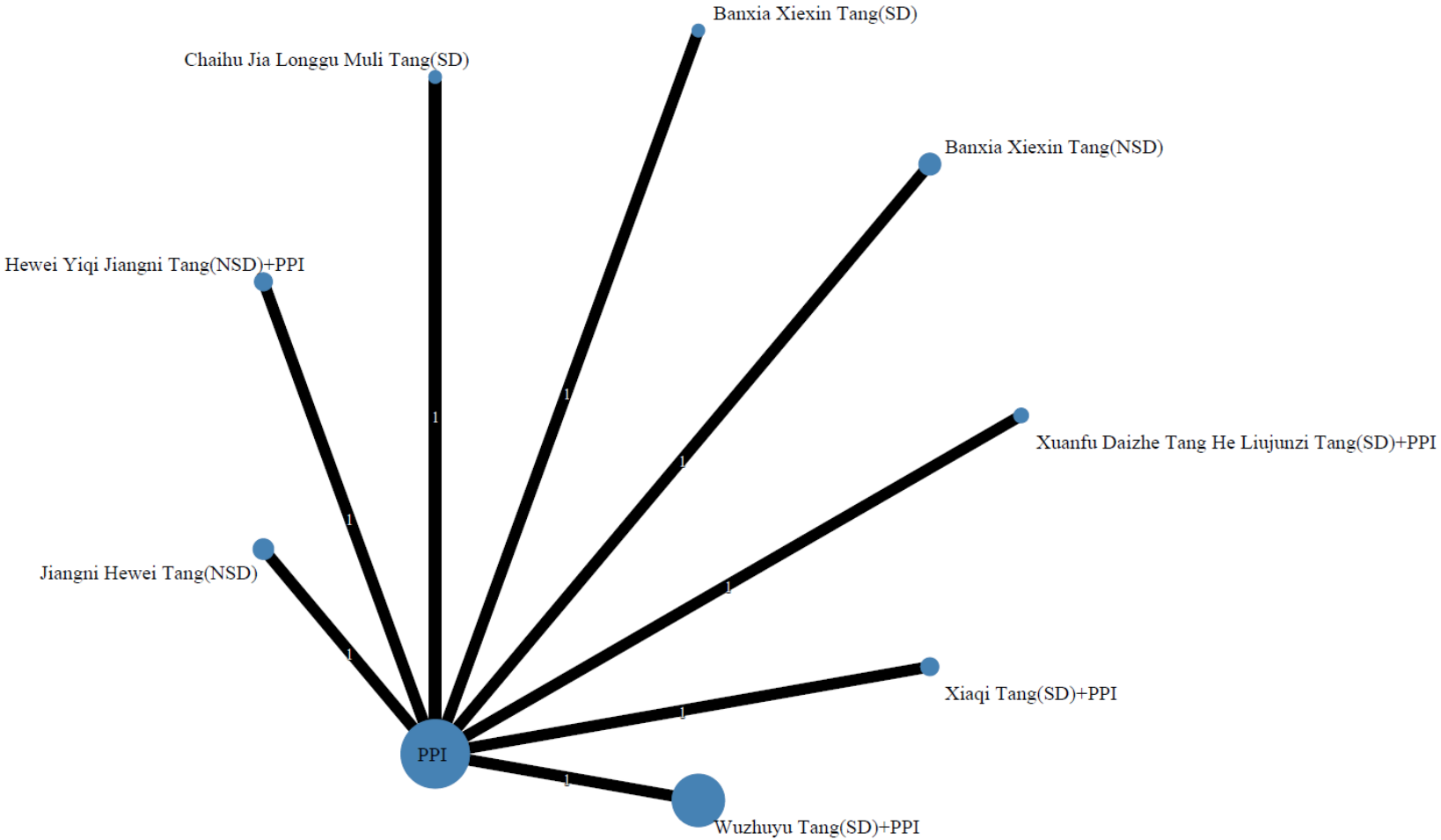
Note: NSD, non-syndrome differentiation; SD, syndrome differentiation; PPI, proton pump inhibitors. The x-axis represents the possible rank on recurrence rate of each intervention (from the first best rank to the worst). The y-axis indicates the cumulative probability for each intervention to be the best intervention, the second-best intervention, the third best intervention, and so on. In this graphical approach, rankings are presented through examining the area under the curve. The bigger the area under the curve, the higher the likelihood that an intervention is in the top rank or one of the top ranks in achieving a lower recurrence rate.

附錄 21 不同的中藥方及是否採用辨證論治原則在 GERD 患者復發率 NMA 效應估計的證據質量等級

Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		RR (95%CI)	Quality of evidence	RR (95%CI)	Quality of evidence	RR (95%CI)	Quality of evidence
Hewei Yiqi Jiangni Tang(NSD)+PPI vs PPI	1 (120)	0.20 (0.06; 0.66)	Low*#	Not estimable‡	Not estimable‡	0.20 (0.06; 0.66)	Low
Hewei Yiqi Jiangni Tang(NSD)+PPI vs Huagan Jian He Zuojin Wan(SD)+PPI	2 (206)	-	-	0.37 (0.09; 1.52)	Very lowΔ	0.37 (0.09; 1.52)	Very low
Hewei Yiqi Jiangni Tang(NSD)+PPI vs Wuzhuyu Tang Lianhe Chaihu Shugan San(SD)+PPI	2 (180)	-	-	0.60 (0.05; 7.35)	Very lowΔ§	0.60 (0.05; 7.35)	Very low
PPI vs Huagan Jian He Zuojin Wan(SD)+PPI	1 (86)	1.88 (0.89; 3.96)	Very low*##	Not estimable‡	Not estimable‡	1.88 (0.89; 3.96)	Very low
PPI vs PPI+Wuzhuyu Tang and Chaihushugan San(SD)	1 (60)	3.00 (0.33; 27.2)	Very low*##	Not estimable‡	Not estimable‡	3.00 (0.33; 27.2)	Very low
Huagan Jian He Zuojin Wan(SD)+PPI vs Wuzhuyu Tang Lianhe Chaihu Shugan San(SD)+PPI	2 (146)	-	-	1.60 (0.16; 16.4)	Very lowΔΔ§	1.60 (0.16; 16.4)	Very low

Note: NSD, Non-Syndrome Differentiation; SD, Syndrome Differentiation; RCT, Randomized Controlled Trial; RR, Risk Ratio; CI: confidence interval; PPI, Proton Pump Inhibitors. -: Not applicable. *Serious study limitations (crucial risk of bias for one criterion, or multiple criteria, and likely to seriously alter the results); # Serious impression (small sample size less than 400); ## Very serious impression (95% CI overlaps the RR of 1.0, but includes important benefit or important harm (RR estimates below 0.5 and above 2.0 are considered clinically important); and the small sample size (less than 400)); ΔContributing direct evidence of low or very low quality; ΔΔContributing direct evidence of very low quality; ‡: Cannot be estimated because the drug was not connected in a loop in the evidence network; §Indirectness because of questionable comparability of trial populations or because of intransitivity.

附錄 22 不同中藥方及是否採用辨證論治幹預模式在 GERD 患者不良反應發生率上的網絡幾何圖



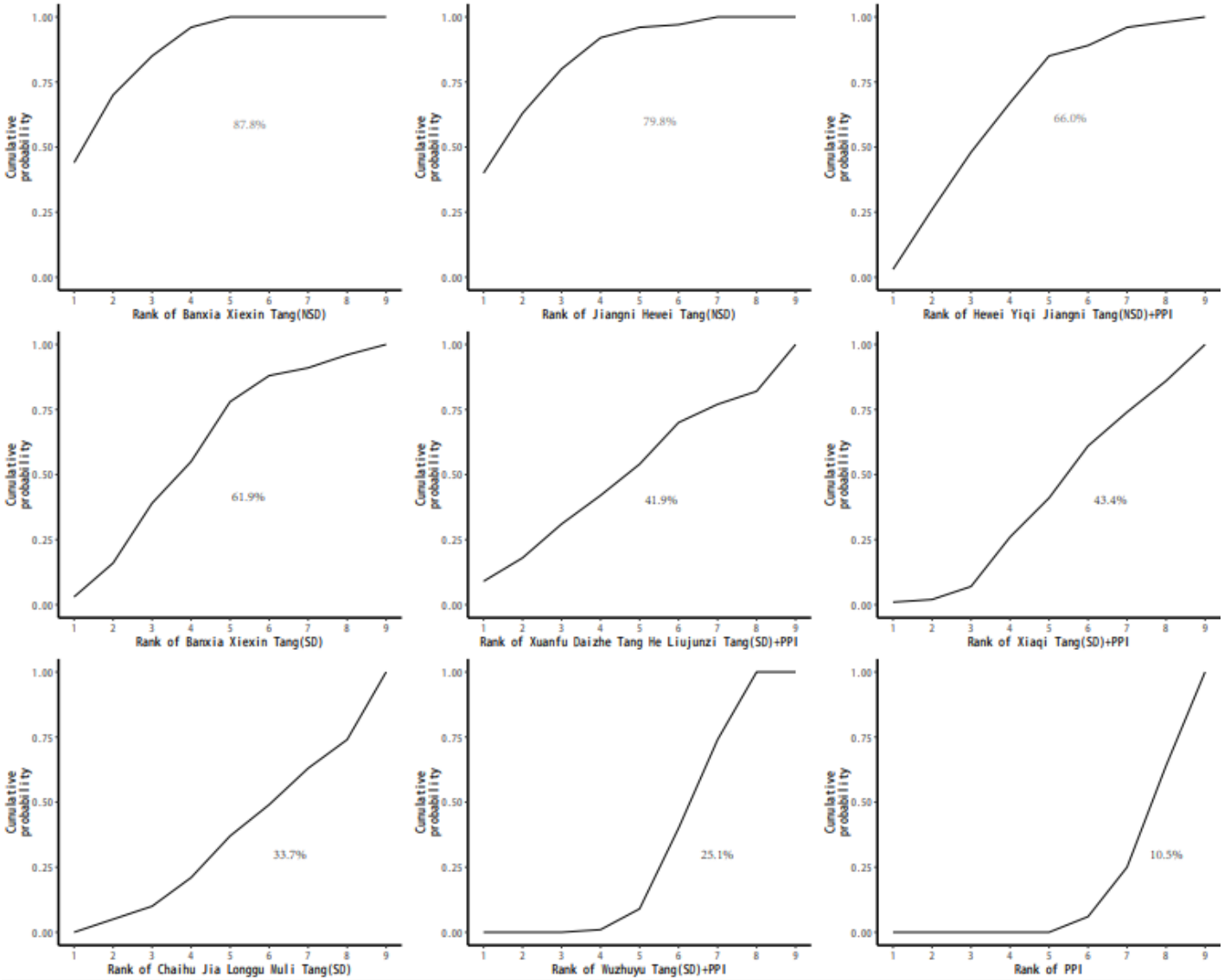
Notes: NSD, non-syndrome differentiation; SD, syndrome differentiation; PPI, proton pump inhibitors. Nodes represent the interventions, node sizes correspond to the number of studies involved, lines connecting nodes represent direct comparisons between pairs of interventions. Width of the lines represents the proportion of the number of trials for each comparison as compared to total number of trials. Number on the line represents the count of the direct comparisons.

附錄 **23** 不同的中藥方及是否採用辨證論治原則在 **GERD** 患者不良反應發生率上兩兩比較的結果

Banxia Xiexin Tang(NSD)								
0.16(0.01;5.14)	Banxia Xiexin Tang(SD)							
0.05(0.00;1.78)	0.29(0.01;6.37)	Chaihu Jia Longgu Muli Tang(SD)						
0.21(0.01;6.61)	1.29(0.07;22.8)	4.50(0.20;100)	Hewei Yiqi Jiangni Tang(NSD) +PPI					
0.48(0.01;25.8)	3.00(0.09;97.1)	10.5(0.27;411)	2.33(0.07;75.6)	Jiangni Hewei Tang(NSD)				
0.02(0.00;0.38)	0.14(0.02;1.09)	0.50(0.05;5.22)	**0.11(0.01;0.85)**	**0.05(0.00;0.80)**	PPI			
0.03(0.00;0.55)	0.20(0.02;1.59)	0.70(0.06;7.57)	0.15(0.02;1.24)	0.07(0.00;1.15)	1.39(0.91;2.14)	Wuzhuyu Tang(SD)+PPI		
0.07(0.00;2.49)	0.43(0.02;8.79)	1.50(0.06;38.3)	0.33(0.02;6.85)	0.14(0.00;5.22)	3.00(0.32;28.0)	2.15(0.22;21.0)	Xiaqi Tang(SD)+PPI	
0.07(0.00;4.75)	0.43(0.01;18.5)	1.50(0.03;77.5)	0.33(0.01;14.4)	0.14(0.00;9.95)	3.00(0.13;71.5)	2.15(0.09;52.8)	1.00(0.02;48.4)	Xuanfu Daizhe Tang He Liujunzi Tang(SD) +PPI

Note: NSD, non-syndrome differentiation; SD, syndrome differentiation; PPI, proton pump inhibitors. The values in each cell represent risk ratio (RR, and 95% confidence interval) of the intervention at the top, compared to the comparator on the left. An RR less than 1 suggests a preference for the column intervention, indicating superior performance compared to the row intervention, with a lower adverse event rate for patients with gastroesophageal reflux disease. Significant results are in bold and underlined.

附錄 24 不同的中藥方及是否採用辨證論治原則在 GERD 患者不良反應發生率上的 SUCRA 結果



Note: NSD, non-syndrome differentiation; SD, syndrome differentiation; PPI, proton pump inhibitors. The x-axis represents the possible rank on adverse event rate of each intervention (from the first best rank to the worst). The y-axis indicates the cumulative probability for each intervention to be the best intervention, the second-best intervention, the third best intervention, and so on. In this graphical approach, rankings are presented through examining the area under the curve. The bigger the area under the curve, the higher the likelihood that an intervention is in the top rank or one of the top ranks in achieving a lower adverse event rate.

附錄 25 不同的中藥方及是否採用辨證論治原則在 GERD 患者不良反應發生率 NMA 結果證據質量等級

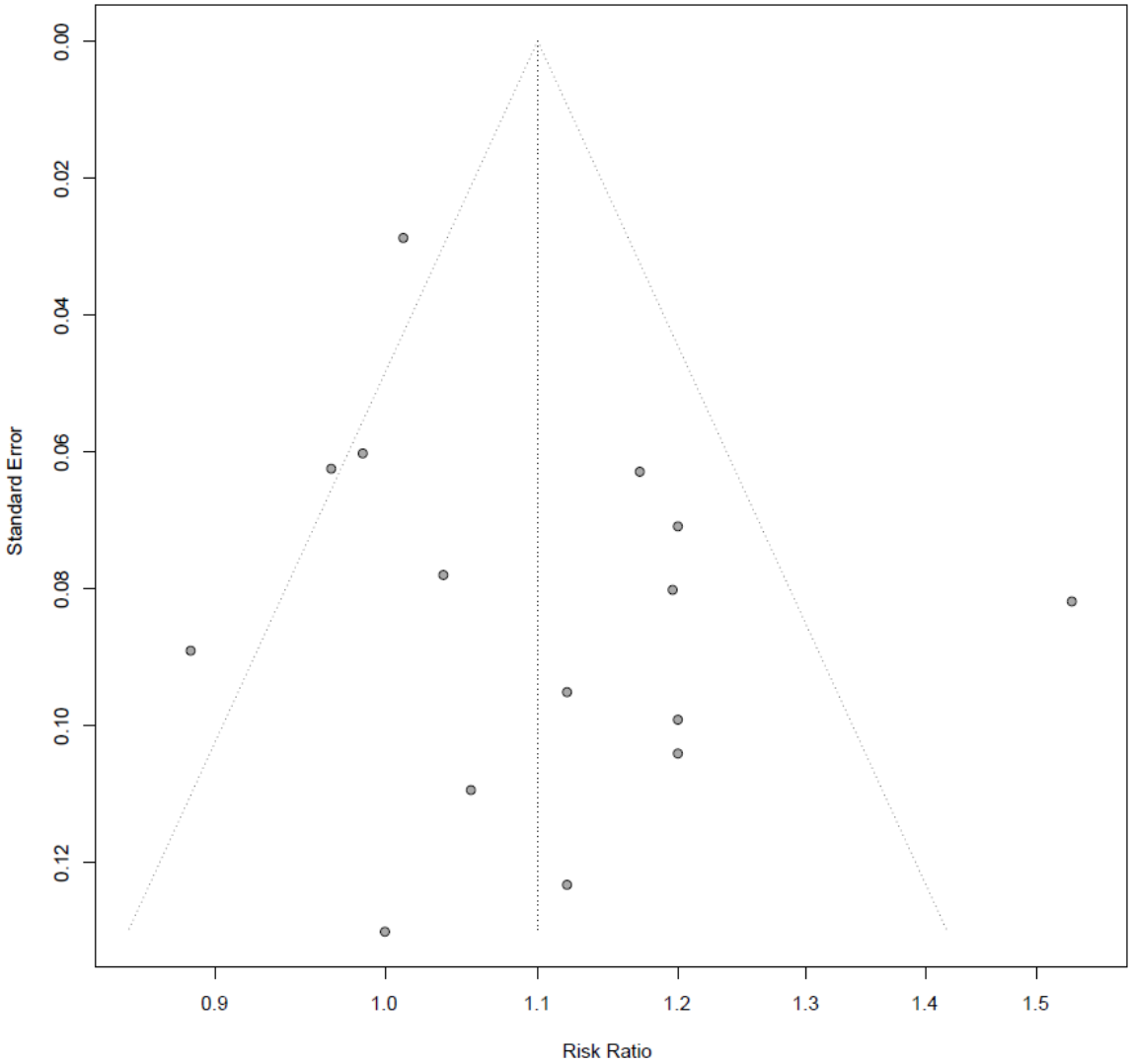
Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		RR (95%CI)	Quality of evidence	RR (95%CI)	Quality of evidence	RR (95%CI)	Quality of evidence
Banxia Xiexin Tang(NSD) vs Hewei Yiqi Jiangni Tang(NSD)+PPI	2 (270)	-	-	0.21 (0.01; 6.61)	LowΔ	0.21 (0.01; 6.61)	Low
Banxia Xiexin Tang(NSD) vs Jiangni Hewei Tang(NSD)	2 (310)	-	-	0.48 (0.01;25.76)	Very lowΔ§	0.48 (0.01; 25.76)	Very low
Banxia Xiexin Tang(NSD) vs PPI	1 (150)	0.02 (0.00;0.38)	Low*#	Not estimable‡	Not estimable‡	0.02 (0.00;0.38)	Low
Banxia Xiexin Tang(NSD) vs Xiaqi Tang(SD)+PPI	2 (270)	-	-	0.07 (0.00;2.49)	Very lowΔΔ§	0.07 (0.00;2.49)	Very low
Banxia Xiexin Tang(NSD) vs Banxia Xiexin Tang(SD)	2 (210)	-	-	0.16 (0.01;5.14)	Very lowΔΔ§	0.16 (0.01;5.14)	Very low
Banxia Xiexin Tang(NSD) vs Wuzhuyu Tang(SD)+PPI	2 (1,176)	-	-	0.03 (0.00;0.55)	LowΔ	0.03 (0.00;0.55)	Low
Banxia Xiexin Tang(NSD) vs Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI	2 (230)	-	-	0.07 (0.00;4.75)	Very lowΔΔ	0.07 (0.00;4.75)	Very low
Banxia Xiexin Tang(NSD) vs Chaihu Jia Longgu Muli Tang(SD)	2 (210)	-	-	0.05 (0.00;1.78)	Very lowΔΔ§	0.05 (0.00;1.78)	Very low
Hewei Yiqi Jiangni Tang(NSD)+PPI vs Jiangni Hewei Tang(NSD)	2 (280)	-	-	2.33 (0.07; 75.56)	Very lowΔ§	2.33 (0.07; 75.56)	Very low
Hewei Yiqi Jiangni Tang(NSD)+PPI vs PPI	1 (120)	0.11 (0.01;0.85)	Low*#	Not estimable‡	Not estimable‡	0.11 (0.01;0.85)	Low
Hewei Yiqi Jiangni Tang(NSD)+PPI vs Xiaqi Tang(SD)+PPI	2 (240)	-	-	0.33 (0.02;6.85)	Very lowΔΔ§	0.33 (0.02;6.85)	Very low
Hewei Yiqi Jiangni Tang(NSD)+PPI vs Banxia Xiexin Tang(SD)	2 (180)	-	-	0.78 (0.04; 13.81)	Very lowΔΔ§	0.78 (0.04; 13.81)	Very low

Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		RR (95%CI)	Quality of evidence	RR (95%CI)	Quality of evidence	RR (95%CI)	Quality of evidence
Hewei Yiqi Jiangni Tang(NSD)+PPI vs Wuzhuyu Tang(SD)+PPI	2 (1,146)	-	-	0.15 (0.02;1.24)	LowΔ	0.15 (0.02;1.24)	Low
Hewei Yiqi Jiangni Tang(NSD)+PPI vs Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI	2 (200)	-	-	0.33 (0.01; 14.42)	Very lowΔΔ	0.33 (0.01; 14.42)	Very low
Hewei Yiqi Jiangni Tang(NSD)+PPI vs Chaihu Jia Longgu Muli Tang(SD)	2 (180)	-	-	0.22 (0.01;4.96)	Very lowΔΔ§	0.22 (0.01;4.96)	Very low
Jiangni Hewei Tang(NSD) vs PPI	1 (160)	0.05 (0.00;0.80)	Low*#	Not estimable‡	Not estimable‡	0.05 (0.00;0.80)	Low
Jiangni Hewei Tang(NSD) vs Xiaqi Tang(SD)+PPI	2 (280)	-	-	0.14 (0.00;5.22)	Very lowΔΔ§	0.14 (0.00;5.22)	Very low
Jiangni Hewei Tang(NSD) vs Banxia Xiexin Tang(SD)	2 (220)	-	-	0.33 (0.01; 10.78)	Very lowΔΔ	0.33 (0.01; 10.78)	Very low
Jiangni Hewei Tang(NSD) vs Wuzhuyu Tang(SD)+PPI	2 (1,186)	-	-	0.07 (0.00;1.15)	Very lowΔ§	0.07 (0.00;1.15)	Very low
Jiangni Hewei Tang(NSD) vs Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI	2 (240)	-	-	0.14 (0.00;9.95)	Very lowΔΔ§	0.14 (0.00;9.95)	Very low
Jiangni Hewei Tang(NSD) vs Chaihu Jia Longgu Muli Tang(SD)	2 (220)	-	-	0.10 (0.00;3.73)	Very lowΔΔ§	0.10 (0.00;3.73)	Very low
PPI vs Xiaqi Tang(SD)+PPI	1 (120)	3.00 (0.32; 28.03)	Very low*##	Not estimable‡	Not estimable‡	3.00 (0.32; 28.03)	Very low
PPI vs Banxia Xiexin Tang(SD)	1 (60)	7.00 (0.92; 53.47)	Very low*##	Not estimable‡	Not estimable‡	7.00 (0.92; 53.47)	Very low
PPI vs Wuzhuyu Tang(SD)+PPI	1 (1,026)	1.39 (0.91;2.14)	Low*#	Not estimable‡	Not estimable‡	1.39 (0.91;2.14)	Low
PPI vs Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI	1 (80)	3.00 (0.13; 71.48)	Very low*##	Not estimable‡	Not estimable‡	3.00 (0.13; 71.48)	Very low

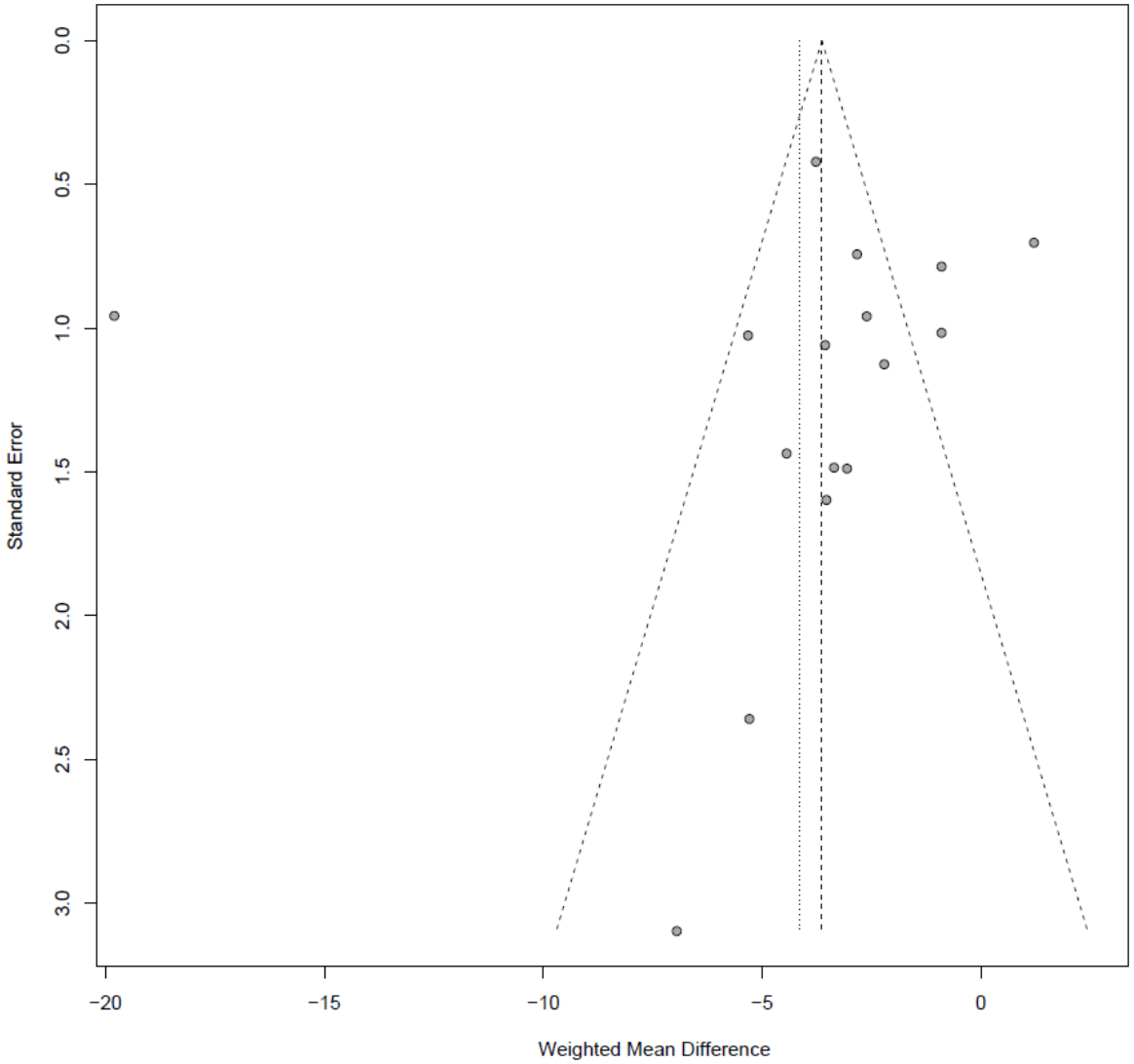
Comparisons	Number of RCTs (participants)	Direct evidence		Indirect evidence		Network meta-analysis	
		RR (95%CI)	Quality of evidence	RR (95%CI)	Quality of evidence	RR (95%CI)	Quality of evidence
PPI vs Chaihu Jia Longgu Muli Tang(SD)	1 (60)	2.00 (0.19; 20.90)	Very low*##	Not estimable‡	Not estimable‡	2.00 (0.19; 20.90)	Very low
Xiaqi Tang(SD)+PPI vs Banxia Xiexin Tang(SD)	2 (180)	-	-	2.33 (0.11; 47.87)	Very lowΔΔ§	2.33 (0.11; 47.87)	Very low
Xiaqi Tang(SD)+PPI vs Wuzhuyu Tang(SD)+PPI	2 (1,146)	-	-	0.46 (0.05;4.52)	Very lowΔΔ§	0.46 (0.05;4.52)	Very low
Xiaqi Tang(SD)+PPI vs Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI	2 (200)	-	-	1.00 (0.02; 48.38)	Very lowΔΔ§	1.00 (0.02; 48.38)	Very low
Xiaqi Tang(SD)+PPI vs Chaihu Jia Longgu Muli Tang(SD)	2 (180)	-	-	0.67 (0.03; 17.03)	Very lowΔΔ	0.67 (0.03; 17.03)	Very low
Banxia Xiexin Tang(SD) vs Wuzhuyu Tang(SD)+PPI	2 (1,086)	-	-	0.20 (0.02;1.59)	Very lowΔΔ§	0.20 (0.02;1.59)	Very low
Banxia Xiexin Tang(SD) vs Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI	2 (140)	-	-	0.43 (0.01; 18.53)	Very lowΔΔ§	0.43 (0.01; 18.53)	Very low
Banxia Xiexin Tang(SD) vs Chaihu Jia Longgu Muli Tang(SD)	2 (120)	-	-	0.29 (0.01;6.37)	Very lowΔΔ§	0.29 (0.01;6.37)	Very low
Wuzhuyu Tang(SD)+PPI vs Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI	2 (1,106)	-	-	2.15 (0.09; 52.79)	Very lowΔΔ	2.15 (0.09; 52.79)	Very low
Wuzhuyu Tang(SD)+PPI vs Chaihu Jia Longgu Muli Tang(SD)	2 (1,086)	-	-	1.43 (0.13; 15.59)	Very lowΔΔ§	1.43 (0.13; 15.59)	Very low
Xuanfu Daizhe Tang He Liujunzi Tang(SD)+PPI vs Chaihu Jia Longgu Muli Tang(SD)	2 (140)	-	-	0.67 (0.01; 34.44)	Very lowΔΔ§	0.67 (0.01; 34.44)	Very low

Note: NSD, Non-Syndrome Differentiation; SD, Syndrome Differentiation; RCT, Randomized Controlled Trial; RR, Risk Ratio; CI: confidence interval; PPI, Proton Pump Inhibitors. -: Not applicable. *Serious study limitations (crucial risk of bias for one criterion, or multiple criteria, and likely to seriously alter the results); # Serious impression (small sample size less than 400 or 95% CI overlaps the RR of 1.0, but includes important benefit or important harm (RR estimates below 0.5 and above 2.0 are considered clinically important)); ## Very serious impression (95% CI overlaps the RR of 1.0, but includes important benefit or important harm (RR estimates below 0.5 and above 2.0 are considered clinically important); and the small sample size (less than 400)); ΔContributing direct evidence of low quality; ΔΔContributing direct evidence of low or very low quality; ‡: Cannot be estimated because the drug was not connected in a loop in the evidence network; §Indirectness because of questionable comparability of trial populations or because of intransitivity.

附錄 26 漏斗圖, 結局：癥狀緩解率



附錄 27 漏斗圖, 結局：RDQ 評分變化



Note: RDQ: Reflux Disease Questionnaire