

学校代号 10162

学 号 201610100019



遼寧中醫藥大學

Liaoning University of Traditional Chinese Medicine

博士学位论文

小柴胡汤及针灸干预慢性乙型肝炎的循证
评价与证据质量可视化工具的应用探索

学位申请人	孔德昭
指导教师姓名	刘建平
专业名称	中西医结合基础
申请学位类型	学术学位
论文提交日期	二 一九年三月

辽宁中医药大学学位论文独创性与真实性声明

本人郑重声明：在我呈交的学位论文中，本人经研究所得的数据、原理、结论等一切内容均真实，且经本人认真核对无误；我呈交的与学位论文相关的资料（原始记录、照片、录像片、检查化验报告单等）确为完整的学位论文研究中的原始资料，本人没作任何修改；我呈交的学位论文是我本人在指导教师的指导下进行的研究工作及取得的研究成果。除文中特别加以标注和致谢的地方外，论文中没有其他人已经发表或撰写过的研究成果，其他人对我完成此论文的帮助以及对本研究所做的任何贡献均已在论文中作了明确说明并表示谢意。

我呈交的学位论文及与该论文相关的资料若有不实之处，本人承担一切相关责任。

学位论文作者签名 孔德昭

日期：2019年6月11日

关于学位论文使用授权的声明

本人完全了解辽宁中医药大学有关保留使用学位论文的规定，同意学校保留或向国家有关部门机构送交论文的印刷版和电子版，允许被查阅和借阅。本人授权辽宁中医药大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或其他复印手段保存和汇编本学位论文。

（保密论文在解密后应遵守此规定）

学位论文作者签名 孔德昭 学位论文导师签名 孔克平

日期：2019年6月11日

目 录

一、摘要	
中文摘要.	1
英文摘要.	8
二、英文缩略词表.	17
三、正文	
前言.	19
论文一 中药复方小柴胡汤干预慢性乙型肝炎的循证评价	
资料与方法.	24
研究结果.	33
讨论.	60
小结.	63
论文二 针灸疗法干预慢性乙型肝炎的循证评价	
资料与方法.	65
研究结果.	73
讨论.	98
小结.	102
论文三 三维误差证据矩阵在中医临床证据评价中应用的探索分析	
资料与方法.	105
研究结果.	111
讨论.	144
小结.	146
结论.	147
四、本研究创新性的自我评价.	149
五、参考文献.	151
六、附表.	176

七、附录

综述一 高级别循证医学证据的评价与综合的重要性和方法学进展.....	182
综述二 证据评价的工具现状及研制新的多维质量评价工具的必要性.....	187
个人简介.....	195
在学期间科研成绩.....	197
致谢.....	199

摘要

目的:

1. 评估中药复方小柴胡汤和针灸疗法（包括手针、电针、激光针刺、灸法、药物穴位注射、药物贴敷、穴位按压）分别干预慢性乙型肝炎的获益和风险，为临床决策者提供临床决策依据。
2. 评估中药复方小柴胡汤和针灸疗法分别干预慢性乙型肝炎的临床研究的方法学质量，为中医药未来临床试验研究指明方向。
3. 以中药复方小柴胡汤和针灸疗法分别干预慢性乙型肝炎为例，构建评估中医临床证据体的可视化工具-三维误差证据矩阵，通过系统误差、随机误差、设计误差三个维度全面评估中医临床证据体的质量，补充常用的证据质量评估方法对设计误差重视不足的缺陷，从系统误差、随机误差、设计误差三个维度提出中医临床研究设计改进建议，提高中医临床研究质量，促进中医疗法的临床实施。

资料与方法:

1. 中药复方小柴胡汤干预慢性乙型肝炎的循证评价

检索全球范围内运用中药复方小柴胡汤干预慢性乙型肝炎的临床研究，构建Cochrane系统综述分析符合纳入标准的临床研究以评估中药复方小柴胡汤干预慢性乙型肝炎的获益和风险。运用试验序贯分析（TSA）探索要获得“充分证据”所需要的样本量大小（IS），并对获得的证据进行统计显著性检验和无效性检验，以判断是否需要更多的临床试验以进一步提供证据，指明未来临床研究方向。通过GRADE分级评估所获得证据质量的确定性，为临床决策制定提供依据。

2. 针灸疗法干预慢性乙型肝炎的循证评价

检索全球范围内运用针灸疗法（包括手针、电针、激光针刺、灸法、药物穴位注射、药物贴敷、穴位按压）干预慢性乙型肝炎的临床研究，构建Cochrane系统综述分析符合纳入标准的临床研究以评估针灸疗法干预慢性乙型肝炎的获益和风险。运用试验序贯分析（TSA）探索要获得“充分证据”所需要的样本量大小（IS），并对获得的证据进行统计显著性检验和无效性检验，以判断是否需要更多的临床试验以进一步提供证据，指明未来临床研究方向。通过GRADE分级评估所获得证据质量的确定性，为临床决策制定提供依据。

3. 三维误差证据矩阵在中医临床证据评价中应用的探索分析

以小柴胡汤和针灸疗法（包括手针、电针、激光针刺、灸法、药物穴位注射、药物贴敷、穴位按压）干预慢性乙型肝炎为例，评估获得的临床证据体（Evidence body）的质量。通过系统误差（Systematic error）、随机误差（Random error）和设计误差（Design error）评估获得证据的内部真实性和外部真实性。运用证据等级水平对临床证据的系统误差风险大小进行排序，运用‘Risk of bias tool’评估临床证据的具体系统误差。运用标准误（Standard Error）评估临床证据的随机误差风险大小。根据“结局指标”的重要性对临床证据的设计误差风险大小进行分级，并对影响临床证据的设计误差的其他因素进行描述性统计分析。最终从系统误差、随机误差、设计误差三个纬度构建可视化“三维误差证据矩阵”（error matrix），将获得的临床证据按照这三个纬度进行等级排序，提出临床证据质量改进方法及建议。

结 果：

1. 中药复方小柴胡汤干预慢性乙型肝炎的循证评价

我们纳入了 10 项临床随机对照试验(RCT)共随机分配 926 例参与者。所有纳入研究均比较小柴胡汤联合常规疗法对经常规疗法。纳入的研究包含传统小柴胡汤及小柴胡汤加减方，疗程为 3-8 个月左右。其中有一项研究纳入合并肺结核的人群，有一项研究纳入合并肝硬化人群，其余研究只纳入患有慢性乙肝人群。所有纳入研究均被评价为高风险偏倚研究，并且所有结局指标的证据的确定性均极低。

没有一项研究对全因死亡率、严重不良事件、健康相关生存质量、乙肝相关死亡率、或乙肝相关并发症这些结局指标进行评估和报告。对于非严重不良事件发生率（综合报告）这一结局指标，小柴胡汤联合常规治疗组与常规治疗组相比较，干预效应差异无统计学意义（RR=0.43, 95%CI [0.02, 11.98]； $I^2=69\%$ ； $P=0.62$ ；纳入两项研究；240 例参与者；证据确定性极低）。对于 HBV-DNA 阳性检测率改变这一结局指标，小柴胡汤联合常规治疗与常规治疗相比，小柴胡汤可降低 HBV-DNA 阳性人群比例（RR=0.62, 95%CI [0.45, 0.85]； $I^2=0\%$ ； $P=0.003$ ；纳入三项研究；222 例参与者；证据确定性极低）。对于 HBeAg 阳性检测率改变这一结局指标，小柴胡汤联合常规治疗组与常规治疗组相比，干预效应差异无统计学意义（RR=0.72, 95%CI [0.50, 1.02]； $I^2=38\%$ ； $P=0.003$ ；纳入两项研究；160 例参与者；证据确定性极低）。对于分开报告的非严重不良事件的发生率，小柴胡汤联合常规治疗组与常规治疗组相比较，干预效应差异无统计学意义：不良事件

恶心的发生率 (RR=5, 95%CI[0.24, 102.53]; P>0.05; 仅纳入一项研究); 不良事件恶心及呕吐的发生率 (RR=2, 95%CI[0.19, 21.18]; P>0.05; 仅纳入一项研究); 不良事件头晕及失眠的发生率 (RR=0.11, 95%CI[0.01, 2.03]; P>0.05; 仅纳入一项研究); 不良事件头晕及乏力的发生率 (RR=1, 95%CI[0.06, 15.44]; P>0.05; 仅纳入一项研究); 不良事件口干及口苦的发生率 (RR=1, 95%CI[0.06, 15.44]; P>0.05; 仅纳入一项研究); 不良事件腹胀及嗝气的发生率 (RR=1, 95%CI[0.06, 15.44]; P>0.05; 仅纳入一项研究); 不良事件食欲不振的发生率 (RR=1, 95%CI[0.06, 15.44]; P>0.05; 仅纳入一项研究)。各不同小柴胡汤组成、疗程、剂型、剂量的亚组分析结果无统计学差异 (P>0.05)。对于参与者血清 ALT 水平这一结局指标, 小柴胡汤联合常规治疗组与常规治疗组相比, 小柴胡汤可降低参与者血清 ALT 水平 (MD=-18.94, 95%CI[-29.42, -8.45]; I²=97%; P=0.0004; 纳入七项研究; 633 例参与者; 证据确定性极低)。对于参与者血清 AST 水平这一结局指标, 小柴胡汤联合常规治疗组与常规治疗组相比, 小柴胡汤可降低参与者血清 AST 水平 (MD=-22.75, 95%CI[-44.64, -0.85]; I²=99%; P=0.04; 纳入六项研究; 867 例参与者; 证据确定性极低)。

10 项研究中有两项研究接受了政府或者医院的学术资金资助, 其余研究未报告资助资金来源。此外, 47 项研究由于无法获得必要的方法学信息, 不符合本研究的纳入标准而没有纳入本研究中。

2. 针灸疗法干预慢性乙型肝炎的循证评价

我们纳入了 8 项临床随机对照试验共随机分配 552 例参与者。所有纳入研究均比较针灸疗法 (包括手针、电针、激光针刺、灸法、药物穴位注射、药物贴敷、穴位按压) 联合常规疗法对照常规疗法, 有两项试验研究穴位注射疗法, 三项试验研究穴位贴敷疗法, 两项试验研究手针疗法, 一项试验研究艾灸疗法。有 7 项研究只纳入患有慢性乙肝人群, 有一项研究纳入合并肺结核人群, 纳入研究的疗程范围为 14 天-12 个月。所有纳入研究均被评价为高风险偏倚研究, 并且所有结局指标的证据的确定性均极低。

没有一项研究对全因死亡率、严重不良事、健康相关生存质量、乙肝相关死亡率、或乙肝相关并发症这些结局指标进行评估和报告。对于非严重不良事件发生率 (综合报告) 这一结局指标, 针灸疗法组与对照组相比较, 干预效应差异无统计学意义 (RR=0.67, 95%CI[0.43, 1.06]; I²=0%; P=0.09; 纳入 3 项研究; 203 例参与者; 证据确定性极低)。对于 HBV-DNA 阳性检测率这一结局指标, 针灸疗法联合常规治疗与常规治疗组相比, 针

灸疗法可降低 HBV-DNA 阳性人群比例 (RR=0.45, 95% CI[0.27, 0.74] ; 仅纳入一项研究; 58 例参与者; 证据确定性极低)。对于 HBeAg 阳性检测率改变这一结局指标, 针灸疗法联合常规治疗组与常规治疗组相比, 干预效应差异无统计学意义 (RR=0.64, 95%CI[0.11, 3.68]; $I^2=98\%$; $P=0.61$; 纳入两项研究; 158 例参与者; 证据确定性极低)。对于分开报告的非严重不良事件的发生率, 针灸疗法联合常规治疗组与常规治疗组相比较, 干预效应差异无统计学意义: 不良事件乏力的发生率 (RR=0.67, 95%CI[0.12, 3.78]; $P>0.05$; 仅纳入一项研究); 不良事件食欲不振的发生率 (RR=0.60, 95%CI[0.15, 2.34]; $P>0.05$; 仅纳入一项研究); 不良事件胁肋疼痛的发生率 (RR=0.67, 95%CI[0.26, 1.70]; $P>0.05$; 仅纳入一项研究); 不良事件心动过缓的发生率 (RR=2.81, 95%CI[0.12, 66.17]; $P>0.05$; 仅纳入一项研究); 不良事件血常规显著异常的发生率 (RR=0.62, 95%CI[0.30, 1.29]; $P>0.05$; 仅纳入一项研究)。针灸疗法联合常规治疗组与常规治疗组相比较, 可能增加不良事件腹胀的发生率 (RR=4.5, 95%CI[1.04, 19.54]; $P<0.05$; 仅纳入一项研究)。各不同针灸疗法、疗程、诊断标准、参与者合并疾病的亚组分析结果无统计学差异 ($P>0.05$)。对于参与者血清 ALT 水平这一结局指标, 针灸联合常规治疗组与常规治疗组相比, 针灸可降低参与者血清 ALT 水平 (MD=-21.35, 95%CI[-33.60, -9.09]; $I^2=81\%$; $P=0.0006$; 纳入四项研究; 275 例参与者; 证据确定性极低)。对于参与者血清 AST 水平这一结局指标, 针灸联合常规治疗组与常规治疗组相比, 针灸可降低参与者血清 AST 水平 (MD=-18.31, 95%CI[-35.44, -1.18]; $I^2=94\%$; $P=0.04$; 纳入四项研究; 275 例参与者; 证据确定性极低)。

8 项研究中有三项研究接受了政府或者医院的学术资金资助, 其余研究未报告资助资金来源。此外, 79 项研究由于缺少必要的方法学信息, 不符合本研究的纳入标准而没有纳入本研究中。

3. 三维误差证据矩阵在中医临床证据评价中应用的探索分析

我们以小柴胡汤和针灸疗法分别干预慢性乙型肝炎为例构建临床证据体质量评估工具-三维误差证据矩阵。

共有两项 meta 分析、52 项随机对照试验研究小柴胡汤干预慢性乙型肝炎的疗效。由于存在高风险的系统误差和设计误差, 和小风险的随机误差, 两项评估小柴胡汤干预慢性乙型肝炎的 meta 分析被评为低质量研究, 并且均不属于系统综述。17% (9/54) 的证据表现为高风险的随机误差。100% 的证据表现为高风险的系统误差。对于设计误差,

没有研究报告全因死亡率，2%（1/54）的证据报告了严重不良事件，20%（11/54）的证据报告了非严重不良事件，72%（39/54）的证据仅报告替代结局指标，43%（23/54）的证据清楚定义了诊断标准、纳入标准和排除标准，57%（31/54）的证据清楚报告了联合干预措施的实施方​​案，19%（10/54）的证据报告小柴胡汤的化裁组成，0%的证据可被评估为实用性研究，4%（2/54）的证据采用多中心试验设计。位于证据矩阵最西北角的证据代表了目前所获得的最高质量证据，为一项低质量的 meta 分析研究。

共有两项 meta 分析，77 项临床随机对照试验以及两项病例系列报告研究针灸疗法干预慢性乙型肝炎的疗效。由于存在高风险的系统误差和设计误差，和小风险的随机误差，两项评估针灸疗法干预慢性乙型肝炎的 meta 分析被评为低质量研究，并且均不属于系统综述。22%（18/81）的证据表现为高风险的随机误差。100%的证据表现为高风险的系统误差。对于设计误差，没有研究报告全因死亡率，1%（1/81）的证据既报告了全因死亡率又报告了乙肝相关并发症，19%（15/81）的证据报告的严重不良事件，36%（29/81）的证据报告了非严重不良事件，12%（10/81）的证据报告了健康相关生存质量，7%（6/81）报告了乙肝相关并发症。58%（47/81）的证据清楚定义了诊断标准、纳入标准和排除标准，44%（36/81）的证据清楚报告了联合干预措施的实施方​​案，9%（3/33）报告了针灸疗法的个体化裁，0%的证据报告了医疗措施实施者的资格标准，0%的研究可被评估为优效性研究或实用性研究，4%（3/81）的证据采用多中心试验设计。位于证据矩阵最西北角的证据代表了目前所获得的最高质量证据，为一项低质量的 RCT 研究。

结 论：

1. 获得的小柴胡汤干预慢性乙型肝炎的证据缺少对于全因死亡率、健康相关生存质量、严重不良事件、乙肝相关死亡率、乙肝相关发病率这些结局指标的评估和报告。通过评价低质量的证据，我们发现小柴胡汤联合常规疗法对比单用常规疗法可降低 HBV-DNA 阳性检测率，降低参与者血清 ALT、AST 水平，但对于非严重不良事件和 HBeAg 阳性检测率改变这两个结局指标，目前尚不能表明小柴胡汤联合常规疗法对比单用常规疗法具有显著有效性。由于检索到的研究具有高风险偏倚，HBV-DNA 阳性检测率、参与者血清 ALT、AST 水平是未经证实的替代结局指标，并且纳入的研究其确定性较低，未达到所需样本量，因此结论尚不能确证。

2. 获得的针灸疗法（穴位注射、中药贴敷、手针、艾灸疗法）干预慢性乙型肝炎的证据缺少对于全因死亡率、健康相关生存质量、严重不良事件、乙肝相关死亡率、乙肝相

关发病率这些结局指标的评估和报告。通过评价低质量的证据，我们发现手针疗法联合常规治疗对比常规治疗可降低 HBV-DNA 阳性检测率，手针、艾灸、穴位贴敷联合常规治疗对比常规治疗可降低参与者血清 ALT、AST 水平，但对于非严重不良事件和 HBeAg 阳性检测率改变这两个结局指标，目前尚不能表明针灸疗法联合常规疗法对比单用常规疗法具有显著有效性。针灸疗法联合常规治疗与常规治疗相比较，可能增加不良事件腹胀的发生率。但由于检索到的研究具有高风险偏倚，HBV-DNA 阳性检测率、参与者血清 ALT、AST 水平是未经证实的替代结局指标，且未达到所需样本量，纳入的研究其确定性较低，因此关于针灸疗法对慢性乙型肝炎的有益性结论尚不能确证。而对于针灸疗法增加慢性乙型肝炎患者腹胀事件的风险，同样由于只有一项研究能够提供数据，未达到所需样本量，且纳入的研究其确定性较低，其结论同样不能确证。

3. 纳入的小柴胡汤和针灸疗法干预慢性乙型肝炎的研究均具有高风险偏倚，研究质量均很低。未来的研究应该对于研究人群的诊断标准以及合并疾病进行更严格清晰的限定，并且应针对不同的合并疾病进行分层随机；应设计更长随访期的临床随机对照试验；应设计更多针对以患者为中心的结局指标（如死亡率、严重不良事件、健康相关生存质量等）的临床试验，并应设计临床随机对照试验以验证替代结局指标与以患者为中心的结局指标的关系。对于小柴胡汤干预慢性乙型肝炎的临床随机对照试验：应针对不同的小柴胡汤方剂组成、不同的剂量以及不同的疗程进行试验设计；更多设计对比小柴胡汤与安慰剂干预慢性乙型肝炎的临床试验。对于针灸疗法干预慢性乙型肝炎的临床随机对照试验：应设计如激光针刺等更多疗法，或者多种疗法联合应用干预慢性乙型肝炎的研究；更多设计针刺疗法对比假针刺疗法的临床试验。

4. 由于高质量的系统综述是证据体中证据等级最高者，我们建议在临床实践中，实施任何干预措施前，先从高质量的系统综述中获得各种误差最小化、质量最高的证据，并且在进行系统综述时运用序贯分析以及 GRADE 分级以辅助评价证据体的质量。

5. 运用本研究建立的可视化“三维误差证据矩阵”发现目前小柴胡汤和针灸疗法干预慢性乙型肝炎的临床证据在系统误差、随机误差、设计误差三个方面存在较严重问题。我们提出提高中医临床证据质量的相关建议：在未来的临床随机对照试验设计中应对导致高风险系统误差的试验要素（如随机化、盲法、随机分配隐藏、不完整的结局指标、选择性报告）进行严谨的设计、实施和报告；应设计足够样本量以及多样性参与者的随机对照试验，以减少随机误差；应更多设计以患者为中心的结局指标（如死亡率、严重不

良事件、健康相关生存质量等)的临床试验;应在研究设计阶段即明确疾病的诊断标准、参与者纳入标准及排除标准,并详细报告;应设计实施多中心临床试验,尽可能的进行实效性及优效性研究;针对中医辨证论治的特色,对于中医药疗法的临床研究,应当在试验设计和实施阶段清楚定义个体化裁的细节,并制定医疗干预措施实施者的资格标准。

6. “三维误差证据矩阵”可应用于中医临床研究的证据体质量的可视化及证据评估,与目前常用的偏倚评估工具“Risk of bias tool”以及 Grade 分级标准既相互支撑,又可弥补后两者对于设计误差具体影响因素评估的不足。

关键词: 小柴胡汤 针灸疗法 慢性乙型肝炎 循证评价 随机对照试验 证据矩阵

Abstract

Objective:

1. To evaluate the benefits and harms of traditional Chinese medicine (TCM) Xiao Chai Hu Tang formula and acupuncture therapies (manual needle acupuncture, electroacupuncture, laser acupuncture, moxibustion, acupoint injection, herbal patching, and acupoint pressure) separately intervening chronic hepatitis B, and to provide recommendations for clinical decision makers.
2. To evaluate the research methodology quality of Xiao Chai Hu Tang formula and acupuncture therapies separately intervening chronic hepatitis B, and to light direction for future clinical trial research in Chinese medicine.
3. Taking Xiao Chai Hu Tang formula and acupuncture therapies separately intervening chronic hepatitis B as examples, we constructed a visual tool, the three-dimensional error evidence matrix, to evaluate the quality of clinical evidence body for TCM through three dimensions of errors: systematic errors, random errors and design errors, to supplement the defects of the commonly used evidence quality assessment methods, which are insufficient attention to design errors, to make suggestions for TCM clinical research design from the three dimensions, to improve the TCM clinical evidence quality, and to further promote clinical implementation of Chinese medicine therapies.

Material and method:

1. Evidence-based evaluation of Xiao Chai Hu Tang formula for chronic hepatitis B
We searched clinical studies of Xiao Chai Hu Tang formula for chronic hepatitis B worldwide. We conducted Cochrane systematic reviews to evaluate the benefits and harms of Xiao Chai Hu Tang formula for chronic hepatitis B. We used the Trial Sequential Analysis (TSA) to explore the required information (IS) for “sufficient evidence”, and we test the statistical significance and invalidity of the retrieved clinical evidence to determine whether further clinical trials are needed, and to light the direction for future clinical research. We assessed the overall quality of the retrieved evidence by GRADE.
2. Evidence-based evaluation of acupuncture therapies for chronic hepatitis B

We searched clinical studies of acupuncture therapies (manual needle acupuncture, electroacupuncture, laser acupuncture, moxibustion, acupoint injection, herbal patching, and acupoint pressure) for chronic hepatitis B worldwide. We conducted Cochrane systematic reviews to evaluate the benefits and harms of acupuncture therapies for chronic hepatitis B. We used the Trial Sequential Analysis (TSA) to explore the required information (IS) for “sufficient evidence”, and we test the statistical significance and invalidity of the retrieved clinical evidence to determine whether further clinical trials are needed, and to light the direction for future clinical research. We assessed the overall certainty of the retrieved evidence by GRADE.

3. Exploratory Analysis of the Application of Three-Dimensional Error Evidence Matrix in the Evaluation of TCM Clinical Evidence

Taking Xiao Chai Hu Tang formula and acupuncture therapies (manual needle acupuncture, electroacupuncture, laser acupuncture, moxibustion, acupoint injection, herbal patching, and acupoint pressure) separately intervening chronic hepatitis B as examples to evaluate the quality of retrieved evidence body. We evaluate the internal and external validity by systematic errors, random errors, and design errors. We use evidence level to rank evidence body from the view of systematic errors, and we use the 'Risk of bias tool' to assess the specific systematic errors of retrieved evidence. We used standard error to assess the risk of random errors. We rank the evidence from the view of the risk of design errors according to the importance of outcomes, and we described other risk factors of design errors in details. We conducted a visualised ‘three-dimensional error evidence matrix’ (error matrix) from three dimensions of systematic errors, random errors and design errors. We ranked retrieved evidence by the error matrix, and we make suggestions improving the quality of clinical evidence.

Results:

1. Evidence-based evaluation of Xiao Chai Hu Tang formula for chronic hepatitis B

We included ten randomised clinical trials with 926 participants randomised. All the trials compared Xiao Chai Hu Tang formula combined with conventional therapies with

conventional therapies. The trials assessed heterogeneous forms of Xiao Chai Hu Tang formula, administered for 3 to 8 months. One trial included participants with hepatitis B and comorbid tuberculosis, and one trial included participants with hepatitis B and liver cirrhosis. The remaining trials included participants with hepatitis B only. All the trials were assessed at high risk of bias, and the certainty of evidence for all outcomes was very low.

None of the trials aimed at assessing all-cause mortality, serious adverse events, health-related quality of life, hepatitis B-related mortality, or morbidity. We are uncertain whether Xiao Chai Hu Tang formula has an effect regarding adverse events considered 'not to be serious' (RR 0.43, 95% CI 0.02 to 11.98; $I^2 = 69\%$; two trials; 240 participants; very low certainty). Xiao Chai Hu Tang formula showed a reduction of the proportion of people with detectable HBV-DNA (RR 0.62, 95% CI 0.45 to 0.85; $I^2 = 0\%$; three trials; 222 participants; very low certainty). We are uncertain whether Xiao Chai Hu Tang formula has an effect regarding the proportion of people with detectable HBeAg (RR 0.72, 95% CI 0.50 to 1.02; $I^2 = 38\%$; two trials; 160 participants; very low certainty). For separately reported adverse events considered 'not to be serious', the intervention effect was not statistically significant in Xiao Chai Hu Tang group compared with the no intervention group: proportion of participants with nausea (RR 5.00; 95% CI 0.24 to 102.53; $P > 0.05$; only one trial), proportion of participants with nausea and vomiting (RR 2.00, 95% CI 0.19 to 21.18; $P > 0.05$; only one trial), proportion of participants with dizziness and sleep disorders (RR 0.11, 95% CI 0.01 to 2.03; $P > 0.05$; only one trial), proportion of participants with dizziness and fatigue (RR 1.00, 95% CI 0.06 to 15.44; $P > 0.05$; only one trial), proportion of participants with a dry feeling or bitter taste in the mouth (RR 1.00, 95% CI 0.06 to 15.44; $P > 0.05$; only one trial), proportion of participants with bloating and belching (RR 1.00, 95% CI 0.06 to 15.44; $P > 0.05$; only one trial), proportion of participants with loss of appetite (RR 1.00, 95% CI 0.06 to 15.44; $P > 0.05$; only one trial). There were no significant differences in the subgroup analysis of the composition, treatment duration, dosage form, and dose of Xiao Chai Hu Tang formula ($P > 0.05$). Xiao Chai Hu Tang formula showed a reduction of ALT serum level of participant (MD -18.94, 95% CI -29.42 to -8.45; $I^2 = 97\%$; seven trials; 633 participants; very low certainty). Xiao Chai Hu Tang formula showed a reduction of AST serum level of participant

(MD -22.75, 95% CI -44.64 to -0.85; $I^2 = 99\%$; six trials; 867 participants; very low certainty).

Two of the included 10 randomised clinical trials received academic funding from government or hospital. The remaining eight trials provided no information on funding. Additionally, 47 trials lacked the necessary methodological information needed to ensure the inclusion of these trials in our review.

2. Evidence-based evaluation of acupuncture therapies for chronic hepatitis B

We included eight randomised clinical trials with 552 participants randomised. All the included trials compared acupuncture (manual needle acupuncture, electroacupuncture, laser acupuncture, moxibustion, acupoint injection, herbal patching, and acupoint pressure) versus no intervention. The included trials assessed heterogeneous acupuncture interventions. All trials used heterogeneous co-interventions applied equally in the compared groups. Seven trials included participants with chronic hepatitis B and one trial included chronic hepatitis B with comorbid tuberculosis. The treatment duration ranged from 14 days to 12 months. All the trials were assessed at high risk of bias, and the certainty of evidence for all outcomes was very low.

None of the trials aimed at assessing all-cause mortality, serious adverse events, health-related quality of life, hepatitis B-related mortality, and hepatitis B-related morbidity. We are uncertain whether acupuncture has an effect regarding adverse events considered not to be serious (RR 0.67, 95% CI 0.43 to 1.06; $I^2 = 0\%$; three trials; 203 participants; very low certainty) and detectable HBeAg (RR 0.79, 95% CI 0.69 to 0.91, $I^2 = 98\%$; two trials; 158 participants; very low certainty). Manual needle acupuncture, compared with no intervention, showed a reduction in detectable HBV-DNA (a non-validated surrogate outcome; RR 0.45, 95% CI 0.27 to 0.74; one trial, 58 participants; very low certainty). For separately reported adverse events considered 'not to be serious', the intervention effect was not statistically significant in acupuncture therapies group compared with the no intervention group: proportion of participants with fatigue (RR 0.67, 95% CI 0.12 to 3.78; $P > 0.05$; only one trial), proportion of participants with loss of appetite (RR 0.60, 95% CI 0.15 to 2.34; $P > 0.05$; only one trial), proportion of participants with flank pain (RR 0.67, 95% CI 0.26 to 1.70; $P > 0.05$; only one

trial), proportion of participants with bradycardia (RR 2.81, 95% CI 0.12 to 66.77; $P > 0.05$; only one trial). The meta-analysis showed that acupuncture was associated with a higher proportion of participants with bloating (RR 4.50, 95% CI 1.04 to 19.54; $P < 0.05$; only one trial). There were no significant differences in the subgroup analysis of different acupuncture therapies, treatment duration, diagnostic criteria, and participants with complications ($P > 0.05$). Manual needle acupuncture, moxibustion, acupoint herbal patching showed a reduction of ALT serum level of participant (MD -21.35, 95% CI -33.60 to -9.09; $I^2 = 81\%$; four trials; 275 participants; very low certainty). Manual needle acupuncture, moxibustion, acupoint herbal patching showed a reduction of AST serum level of participant (MD -18.31, 95% CI -35.44.64 to -1.18; $I^2 = 94\%$; four trials; 275 participants; very low certainty).

Three of the eight included randomised clinical trials received academic funding from government or hospital. None of the remaining five trials reported information on funding. Additionally, 79 trials lacked the necessary methodological information that could have ensured their inclusion in this review.

3. Study of clinical research quality assessment and clinical research design improvement methods for Xiao Chai Hu Tang formula and acupuncture therapies separately intervening chronic hepatitis B

Taking Xiao Chai Hu Tang formula and acupuncture therapies (manual needle acupuncture, electroacupuncture, laser acupuncture, moxibustion, acupoint injection, herbal patching, and acupoint pressure) separately intervening chronic hepatitis B as examples, we constructed evidence matrix to evaluate the quality of retrieved evidence body.

There were two meta-analyses and 52 randomized controlled trials to investigate the effect and safety of Xiao Chai Hu Tang formula for chronic hepatitis B. Through the visual error matrix, we found out that due to high risk of systematic errors and design errors, and the small risk random errors, the two meta-analyses are rated as low-quality research, and none of them is systematic review.

17% (9/54) of the evidence showed high risk of random errors. 100% of the evidence showed high-risk of systematic errors. About design errors, no evidence reported all-cause mortality, 2% (1/54) of the evidence reported serious adverse events, and 20% (11/54) of the evidence

reported non-serious adverse events, 72% (39/54) of the evidence only reported surrogate outcomes, 43% (23/54) of the evidence clearly defined the diagnostic criteria, inclusion criteria, and exclusion criteria, 57% (31/54) of the evidence clearly reported the implementation of the combined intervention, 19% (10/54) of the evidence reports the composition of Xiaochaihutang, 0% of the evidence can be assessed as a practical study, and 4% (2/54) of the evidence was designed as multicenter trials. Evidence at the very northwest corner of the evidence matrix represents the highest quality evidence, which is a low-quality meta-analysis.

Two meta-analyses, 77 clinical randomized controlled trials, and two case series report studies explored the effect and safety of acupuncture therapies for chronic hepatitis B.

Through the visual error matrix, we found out that due to high risk of systematic errors and design errors, and the small risk random errors, the two meta-analyses are rated as low-quality research, and none of them is systematic review.

22% (18/81) of the evidence showed high risk of random errors. 100% of the evidence showed high-risk systematic errors. About design errors, no studies reported all-cause mortality, 1% (1/81) of the evidence reported both all-cause mortality and hepatitis B-related complications, 19% (15/81) of the evidence reported serious adverse events, 36% (29/81) of the evidence reported non-serious adverse events, 12% (10/81) of the evidence reported health-related quality of life, 7% (6/81) reported hepatitis B-related complications, 58% (47/81) of the evidence clearly defined the diagnostic criteria, inclusion criteria, and exclusion criteria, 44% (36/81) of the evidence clearly reported the implementation of the joint intervention, 9% (3/33) reported the acupuncture Individualized treatment protocol, 0% of the evidence reported eligibility criteria for medicine care practitioners, 0% of studies can be assessed as superiority studies or practical studies, and 4% (3/81) of evidence was designed as multicenter trials. Evidence at the very northwest corner of the evidence matrix represents the highest quality evidence, which is a low-quality meta-analysis.

Conclusion:

1. The evidence for Xiao Chai Hu Tang formula for chronic hepatitis B lacks assessment and

report of patient-centred outcomes (i.e. all-cause mortality, health-related quality of life, serious adverse events, hepatitis B-related mortality, and hepatitis B-related morbidity). By evaluating low-quality evidence, we found that Xiao Chai Hu Tang formula compared with no intervention may reduce the proportion of participants with detectable HBV-DNA, ALT serum level of participant, AST serum level of participant. We are uncertain whether Xiao Chai Hu Tang formula compared with no intervention has an effect on non-serious adverse events and HBeAg. Because of the high risk bias of the retrieved studies, the proportion of HBV-DNA positive participants, ALT serum level of participant, AST serum level of participant were unvalidated surrogate outcomes, and the certainty of retrieved evidence is very low, the conclusions were not confirmed.

2. The evidence for acupuncture therapies for chronic hepatitis B lacks assessment and report of patient-centred outcomes (i.e. all-cause mortality, health-related quality of life, serious adverse events, hepatitis B-related mortality, and hepatitis B-related morbidity). By evaluating low-quality evidence, we found that manual needle acupuncture compared with no intervention may reduce the proportion of participants with detectable HBV-DNA, and manual needle acupuncture, moxibustion, acupoint herbal patching may reduce ALT serum level of participant, AST serum level of participant, and moxibustion may induce the proportion of participants with bloating. We are uncertain whether acupuncture therapies compared with no intervention has an effect on non-serious adverse events and HBeAg. Because of the high risk bias of the retrieved studies, the proportion of HBV-DNA positive participants, ALT serum level of participant, AST serum level of participant were unvalidated surrogate outcomes, the certainty of retrieved evidence is very low, and the limited number of included studies for each analysis, the conclusions was not confirmed.

3. The included studies are all assessed as high risk bias, and the quality of the studies is very low. Future studies should define diagnostic criteria of participants and concomitant diseases more strictly and clearly, and should be stratified and randomized for different concomitant diseases. We should design RCT with longer follow-up and patient-centered outcomes (e.g. mortality, serious adverse events, and health-related quality of life). We should conduct more clinical randomized controlled trials to validate the relationship between surrogate outcomes

and patient-centered outcomes. For RCT on Xiao Chai Hu Tang formula for chronic hepatitis B: we should carry out more trials for different composition, doses, and different treatment duration of Xiao Chai Hu Tang formula. We should design trials comparing Xiao Chai Hu Tang formula to placebo for chronic hepatitis B. For RCT on acupuncture therapies for chronic hepatitis B: we should carry out more trials on different acupuncture therapies such as laser acupuncture, moxibustion, or the combination of different acupuncture therapies. We should design trials comparing acupuncture therapies to sham acupuncture for chronic hepatitis B.

4. We recommend that in clinical practice, before implementing any interventions, to obtain evidence from a high-quality systematic review to minimize errors and to use Trial Sequential Analysis and GRADE to assist to evaluate the quality of evidence body.

5. Using the visualised 'three-dimensional evidence matrix' established by our research, we found that clinical evidence of Xiao Chai Hu Tang formula and acupuncture therapies separately intervening hepatitis B has serious problems in system errors, random errors and design errors. We proposed relevant recommendations for improving the quality of clinical evidence in TCM: we should strictly design, perform, and report important factors of systematic errors (eg, randomized, blinded, randomized allocation, incomplete outcomes, selective reports). We should design adequate sample size and diversity participants to minimise random errors. We should clearly define and report the diagnostic criteria, inclusion criteria, and exclusion criteria of participants. We should design more multi-center clinical trials, superiority studies, and pragmatic studies. For clinical research of TCM therapy, we should clearly define and report the details of individualization and the eligibility criteria for the implementers of medical interventions.

6. The 'Three-dimensional evidence matrix' can be applied to the visualization of evidence quality and evidence evaluation in clinical research of Chinese medicine. Evidence matrix is consistent with commonly used evidence quality assessments methods (i.e. 'Risk of bias tool' and GRADE assessment), and the evidence matrix may work as a good supplement to put enough attention on random errors and design errors.

Keyword: Xiao Chai Hu Tang formula; acupuncture therapies; chronic hepatitis B; evidence-based evaluation; randomised clinical trial; evidence matrix

英文缩略词表

英文缩写	英文全称	中文全称
A/G	Albumin/globulin ratio	白蛋白/球蛋白
ALT	Alanine aminotransfease	丙氨酸氨基转移酶
AIDS	Acquired immunodeficiency syndrome	获得性免疫缺陷综合征
AST	Aspartate transaminase	天门冬氨酸氨基转移酶
cccDNA	covalently closed circular DNA	共价闭合环状 DNA
CD4	Cluster of differentiation 4	分化抗原簇 4
CD8	Cluster of differentiation 8	分化抗原簇 8
CI	Confidence interval	置信区间
DARIS	Diversity-adjusted required information size	多样性校正后的所需样本量
EBM	Evidence-based medicine	循证医学
ECG	Electrocardiography	心电图
GRADE	The Grading of Recommendations Assessment, Development and Evaluation	GRADE 分级
GGT	Gamma-glutamyltransferase	γ -谷氨酰
HA	hyaluronic acid	透明质酸
HBeAg	Hepatitis B e-antigen	乙型肝炎 E 抗原
HBV-DNA	Hepatitis B virus DNA	乙型肝炎病毒 DNA
HBcAb	Hepatitis B core-antibody	乙型肝炎病毒核心抗体
HBeAb	Hepatitis B virus e-antibody	乙型肝炎 E 抗体

HBsAg	Hepatitis B surface antigen	乙型肝炎病毒表面抗原
HBV	Hepatitis B virus	乙型肝炎病毒
HCV	Hepatitis C virus	丙型肝炎病毒
HIV	Human Immunodeficiency Virus	人类免疫缺陷病毒
IL-1 β	Interleukin 1 beta	白细胞介素 1
IS	Information size	样本量
IV-C	Type IV Collagen	IV前胶原
LN	Laminin	层粘连蛋白
MD	Mean difference	均差
NK	Natural killer cell	自然杀伤细胞
PTA	Plasma thromboplastin antecedent	血浆促凝血酶原激酶前因
PCIII	Type III procollagen ubulointerstitial injury.	III型前胶原
RCT	Random clinical trial	随机对照试验
RIS	Required information size	所需样本量
RR	Risk ratio	风险比
SE	Standard error	标准误
SOF	Summary of finding	结果摘要表
TSA	Trial sequential analysis	试验序贯分析
TBIL	Total bilirubin	总胆红素
TC	Total cholesterol	总胆固醇
TNF- α	Tumor Necrosis Factor- α	肿瘤坏死因子- α

前 言

随着国家‘一带一路’政策的推行，中医药科学研究作为其中重要的组成部分越来越受到关注。慢性乙型肝炎是全球范围内的常见传染病，其患病人群已超过2.57亿^[1]，每年约有65万人死于肝纤维化、肝功能衰竭、肝癌等慢性乙型肝炎并发症^[2]。中国超过7.18%的人群患有慢性乙型肝炎，其疾病负担日趋加重^[3]。

中医药治疗慢性乙型肝炎具有悠久的历史，大量临床研究评价了多种中医药疗法治疗慢性乙型肝炎的有效性及安全性，因此，对这些证据的深入评价，对于传播中医药成果，使更多患者受益具有重要意义。开展中医药循证医学研究的目的是为患者找到最佳的中医药治疗方案，因此，临床决策者必须运用最佳质量的临床证据为患者选择治疗措施^[4]。如何判断现有的临床试验证据是否是最佳证据是学术界研究的热点。目前，采用Meta分析方法进行的系统综述联合GRADE分级所提供的临床证据是业界公认的最高等级证据^[5]。

本研究基于循证医学理念，以中药复方小柴胡汤和中医针灸疗法（手针、艾灸疗法、穴位注射、中药贴敷等）为切入点，采用Meta分析的系统综述联合GRADE分级方法探讨其治疗慢性乙型肝炎的有效性和安全性，为临床应用中医疗法治疗慢性乙型肝炎提供最高等级证据；运用序贯分析（TSA）法对目前该领域内临床证据的充分性、显著统计性差异和无效性进行判断，为中医药未来临床研究指明方向；构建“三维误差证据矩阵”（error matrix）这一可视化评估工具，通过系统误差、随机误差、设计误差三个维度全面评估小柴胡汤和针灸疗法干预慢性乙型肝炎临床证据体的质量，补充常用的证据质量评估方法对随机误差和设计误差重视不足的缺陷，从系统误差、随机误差、设计误差三个维度提出中医临床研究设计改进方法，提高中医临床研究质量，促进中医疗法的临床实践及应用。

循证医学（EBM）概念于1992年由国际循证医学工作组提出^[6]，并随着以证据为基础的临床实践指南的实施和Cochrane图书馆的发展而在全球范围内广泛应用^[7]。对临床RCT（随机对照试验）进行的系统综述研究所得到的临床证据被公认为评价干预措施疗效的最高等级证据，并应用于临床实践指南的制定，指导临床决策者治疗方案的选择^[5,8,9]。2000年后行业内学者逐渐引入系统综述方法研究中医药疗法治疗疾病的有效性和安全性^[10]。由于中医药临床RCT的设计和实施尚不完善等原因，其提供的证据质量尚不充分，

阻碍了中医药疗法的临床实践与应用^[11]。

2000年 GRADE 分级理念由 GRADE 工作组提出,其建议临床干预措施的实施应基于对干预措施获益和风险的权衡^[12]。GRADE 分级从研究的偏倚风险、证据的间接性、结果的不一致性和异质性、结果的不精确性、研究的发表偏倚五个方面对临床证据进行等级评估^[13,14]。时至今日,采用 Meta 分析方法的系统综述联合 GRADE 分级被逐渐引入中医药领域研究,并被应用于中医疗法治疗疾病的临床诊疗指南制定工作^[15]。

由于系统综述是将目前检索到的相关临床研究结果进行数据累积性分析,因此该过程中采用的重复性统计检验会导致 Meta 分析得到假阳性或假阴性结果,从而影响系统综述的可信性^[16,17]。丹麦哥本哈根临床试验中心提出将序贯分析 TSA 理念引入 Meta 分析中以最小化 Meta 分析的随机误差,并开发了 TSA 软件以实现这一操作^[18,19]。由于该方法较新,在中医药领域的应用尚处于起步阶段。

无论是系统综述还是 GRADE 分级,其研究和评估的主体均是由临床研究提供的证据,由此可见,临床研究的质量是决定临床证据质量和实施的关键。临床研究的真实性和可信度主要反映在其内部真实性和外部真实性两方面,而这两种属性对于证据的质量影响和临床实施至关重要^[20,21]。研究的内部真实性和外部真实性均受到系统误差、随机误差和设计误差的影响^[5,22]。不同的临床研究设计可提供不同强度的证据,行业内将不同的临床证据进行了证据水平分级,即“证据等级金字塔”^[23],可用于评估临床证据的系统误差^[22]。而临床证据的随机误差通常是由于零散分布的小样本数据造成的,可通过计算标准误(Standard error, SE)进行评估,因为标准误可表明样本均值与真实均值的接近程度^[22]。有研究发现,在临床试验及其文献报告中“重视内部真实性、忽视外部真实性”的倾向严重^[24]。吴宇侠等提出可通过外部真实性的影响因素(即临床研究的设计)对临床研究的外部真实性进行评估^[25]。由于结局指标及其测量是临床研究的核心要素,可以将“结局指标的选择”为代表对临床研究的设计误差进行评估^[22]。Keus 等人通过构建系统误差、随机误差和设计误差三个纬度的误差矩阵对临床证据进行评估^[22]。但国际上尚无统一的综合评价临床研究内、外部真实性的方法和工具^[25]。

国际上已有‘SPIRIT’声明即‘规范临床研究方案内容’声明,为临床研究方案所应包含的最基本内容提供指引。其主要用于促进研究的透明度、描述整个研究计划而非指导如何设计一个高质量临床研究^[26,27]。本研究通过现有临床证据,评估小柴胡汤及针灸疗法治疗慢性乙型肝炎的有效性、安全性以及现有证据的充分性,结合中医药自身特

点，构建“三维误差证据矩阵”（error matrix），通过系统误差、随机误差、设计误差三个方面探讨如何提高中医药临床研究的质量，补充常用的证据质量评估方法对随机误差和设计误差重视不足的缺陷，从系统误差、随机误差、设计误差三个维度提出中医临床研究设计改进方法，提高中医药临床证据质量，促进中医药疗法的临床实施应用。

论文一 中药复方小柴胡汤干预慢性乙型肝炎的循证评价

全球约 3.5%的人口感染乙肝病毒^[28]，其中西太平洋地区(6.2%)和非洲地区(6.2%)的感染率最高^[28]。乙肝病毒感染可以是急性或慢性的，从无症状到有症状的不同严重程度^[29]的进行性疾病^[29]。慢性乙型肝炎定义为 HBsAg 阳性超过 6 个月，血清 HBV-DNA 阳性超过 2000 IU / mL (即 104 copies/ mL)，天冬氨酸氨基转移酶 (AST) 或丙氨酸氨基转移酶 (ALT) 持续或间歇性升高，以及肝脏活检结果显示乙型肝炎病毒感染伴有中度或重度坏死性炎症^[30,31]。慢性乙型肝炎极大地加重了乙肝患者及其家庭的经济、心理及生活负担^[32-35]。

乙肝病毒通常经血液、体液、母婴、性接触进行传播^[28]。乙肝病毒 DNA(HBV-DNA) 是乙肝病毒颗粒的核心，同时也是乙肝病毒复制的最敏感的生化指标。而共价闭合环状 DNA (cccDNA) 可作为病毒 RNA 复制的模板^[36,37]，同时也可导致乙肝病毒的持续感染和激活^[38-41]。慢性乙型肝炎患者的初始评估包括对疾病史，体格检查，肝炎的严重程度以及乙肝病毒标志物的评估^[30,31]。降低肝脏细胞的炎性坏死，肝纤维化，失代偿期肝硬化，肝脏衰竭，以及肝癌的发生风险，提高患者生存质量，延长生存期限已成为治疗慢性乙型肝炎的长期目标^[29,31]。

目前较公认的治疗慢性乙型病毒性肝炎的方法为抗病毒治疗，较为广泛应用的药物有两类：干扰素和核苷（酸）类似物。然而患者可因干扰素或核苷（酸）类似物治疗出现不良反应如骨髓抑制^[42]，自身免疫性症状^[43-45]，神经精神症状^[46]以及耐药现象^[47-52]，有些患者则难以支付此类药物治疗的高昂费用^[53,54]。

中药复方小柴胡汤（日文名 sho-sai-ko-to, 或英文名 minor Bupleurum decoction）在古代时被用于治疗具有“往来寒热，胸胁苦满，默默不欲饮食，心烦喜呕，口苦，咽干，目眩，舌苔薄白，脉弦”症状的患者，其症状与现代慢性乙型病毒性肝炎症状十分相似。现今小柴胡汤及其加减方已在中国^[55-57]和日本^[58,59]被广泛应用于慢性乙型肝炎患者。小柴胡汤首次记载于东汉张仲景所著《伤寒论》，其主要成分为柴胡、半夏、生姜、大枣、人参、黄芩、甘草^[60,61]，具有和解少阳，和胃降逆，扶正祛邪的功效。有研究发现，小柴胡汤可能通过抑制肝炎病毒复制^[62]，提高机体免疫功能^[63,64]，抑制肝脏炎症反应并改善肝纤维化^[64-67]，保护肝细胞^[68]，以及抗肿瘤作用而发挥疗效^[69-72]。目前小柴胡汤已具备水煎剂、片剂、胶囊、颗粒剂，以及注射剂等不同剂型^[55,57,73]。

然而，亦有研究报道，不良事件如肺炎^[74-76]、假性醛固酮增多症^[77]、急性肝损伤^[78, 79]、急性肝炎^[80]、急性血小板减少性紫癜^[81]和急性呼吸窘迫综合征^[82]与应用小柴胡汤有关。目前，尚未有关于经过严格方法学设计的循证医学证据对小柴胡汤干预慢性乙型肝炎的获益和风险进行评估。本研究通过严格方法学设计，对小柴胡汤干预慢性乙型肝炎的获益和风险进行循证评价。

资料与方法

1 纳排标准

1.1 研究类型

对于小柴胡汤干预慢性乙型肝炎的获益评估，纳入临床随机临床对照试验，对于试验的盲法、语言、发表年份、发表形式和是否发表无限制。对于小柴胡汤干预慢性乙型肝炎的风险评估，在纳入临床随机临床对照试验的同时也纳入符合标准的半随机对照试验、对照临床研究以及其他观察性研究。

1.2 参与者

1.2.1 纳入标准

根据指南或者由研究者诊断为慢性乙型肝炎的参与者：HBsAg 阳性超过 6 个月，血清 HBV-DNA 阳性超过 2000 IU / mL（即 104 copies/ mL）；AST 或 ALT 持续或间歇性升高；肝脏活检结果显示乙型肝炎病毒感染伴有中度或重度坏死性炎症^[30, 31]，对于患者的性别、年龄、诊断依据指南和诊断标准无限制。参与者除患有慢性乙型肝炎外，亦可患有肝硬化、肝癌、人类免疫缺陷病毒感染(HIV)感染或获得性免疫缺陷综合症(AIDS)、丙肝、丁肝，或其他伴随疾病。

1.2.2 排除标准

无。

1.2.3 干预措施的类型

试验组应用小柴胡汤或小柴胡汤加减方（至少包含四味组成药物：柴胡、半夏、人参、黄芩），对剂量、剂型、给药方案、方剂组成、药物产地均不作限制。对照组应用安慰剂或无任何干预措施。试验组和对照组可应用具有相同的联合治疗干预措施。

1.2.4 结局指标的类型

（1）主要结局指标：

- ①全因死亡率；
- ②发生一种或多种严重不良事件的参与者比例（严重不良事件即任何非预期的医疗事件，可导致死亡、危及生命、需要住院或延长住院、持续或重大残疾或无行为能力、先天性异常或先天性缺陷^[83, 84]。）；
- ③健康相关生存质量：研究者用于评估试验参与者生存质量报告的任何国际公认的量表。

(2) 次要结局指标:

①乙肝相关死亡率;

②乙肝相关并发症(发生一种或多种如下事件的参与者的比例:肝硬化,腹水,静脉曲张出血,肝肾综合征,肝细胞癌,肝性脑病,或肝移植,但并没有发生死亡者);

③发生一种或多种非严重不良事件的参与者比例:非严重不良事件即不满足严重不良事件定义的不良事件^[83,84]。

(3) 探索性结局指标:

①血清、血浆中 HBV-DNA 阳性或者 HBV-DNA 病毒载量阳性的参与者比例;

②血清、血浆中 HBeAg 阳性的参与者比例;

③分开报告的严重不良事件;

④分开报告的非严重不良事件;

⑤分开报告的乙肝相关并发症。

我们评估最长随访期时的结局指标数据。

2. 检索策略

2.1 电子检索

我们检索英文数据库: Cochrane 肝胆病组对照试验注册库(The Cochrane Hepato-Biliary Group Controlled Trials Register)、Cochrane 图书馆中的 Cochrane 对照试验注册中心(Cochrane Central Register of Controlled Trials (CENTRAL) in The Cochrane Library)、MEDLINE Ovid、Embase Ovid、LILACS Bireme (Latin American and Caribbean Health Science Information database)、Science Citation Index Expanded (Web of Science)、Conference Proceedings Citation Index-Science (Web of Science)。我们还检索了中文数据库: 中国知网数据库(China National Knowledge Infrastructure Database, CNKI)、万方数据库、重庆维普中文科技期刊数据库(Chongqing VIP Chinese Science and Technology Periodical Database, VIP)、中国生物医学文献数据库(Chinese Biomedical Literature Database, CBM)。

附表 1 提供了不同数据库的检索策略及时间跨度。

2.2 其他检索

我们检索了相关系统综述、meta 分析以及符合纳入标准的文章的参考文献列表。我们还检索了世界卫生组织国际临床试验注册平台 World Health Organization

International Clinical Trials Registry Platform (www.who.int/ictrp)、ClinicalTrial.gov (www.clinicaltrials.gov/) 以及中国临床试验注册中心 (ChiCTR) 以检索正在进行中的研究和未发表研究。

3. 数据收集与分析

根据 Cochrane 干预措施系统综述手册^[85] 以及 Cochrane 肝胆病协作组模板^[86] 设计实施本研究。运用 Review Manager 5.3 软件^[87] 及试验序贯分析软件 Trial Sequential Analysis version 0.9.5.10 Beta Software^[88,89] 进行数据分析。

3.1 研究的选择

研究者两人为一组独立提取可能符合纳入标准的研究的题目和摘要。将同一研究的多个报告列入该研究参考文献下。如遇研究筛选不一致的情况可以通过讨论或由第三位研究者协助解决及确定。根据 PRISMA 流程图^[90] 记录研究筛选过程。

3.2 数据提取与处理

研究者两人一组运用电子数据收集表格独立提取数据。如有分歧，则重新检查所提取的数据。如遇数据提取不一致的情况可以通过讨论或由第三位研究者协助解决及确定。提取内容包括：发表相关数据（即年份，国家，作者）、研究特征与设计、试验参与者特征、试验纳入和排除标准、干预措施、结局指标、随访、数据分析类型（即意向性治疗分析、校正意向性治疗分析、符合分析方案集分析）。提取最长随访时间节点数据。联系研究人员获得研究报告缺失信息。

3.3 纳入研究的偏倚风险评估

依据 Cochrane 偏倚风险评估工具^[85]，Cochrane 肝胆病组模板 the Cochrane Hepato-Biliary Group Module^[86]，以及相关方法学研究^[91-94]，研究者两人一组独立评估纳入研究的偏倚风险，主要包括以下方面：随机分配产生、分配方案隐藏、对参与者及主要研究者的施盲、对研究结果评估者的施盲、结局指标数据的不完整性、选择性报告、营利偏倚、其他偏倚。对纳入的 RCT 进行质量评价，针对每一项研究结果，对以上方面做出“低偏倚”、“高偏倚”以及对于缺乏相关信息或偏倚情况不确定情况下的“不清楚”的判断，评价者为两名或两名以上，判断意见不一致时征求第三方意见。具体评估标准如下：

(1) 随机分配产生 (Allocation sequence generation) :

①低偏倚风险：研究者运用电脑或随机数字表产生随机序列，如果由没有参与这项

研究的独立人员进行如抽签，掷硬币，洗牌以及掷骰子也可。

②不清楚的偏倚风险：研究者未明确随机序列产生的方法。

③高偏倚风险：随机序列产生的方法是非随机的。对于这部分研究，我们仅纳入其评估干预措施的风险。

(2) 分配方案隐藏 (Allocation concealment) :

①低偏倚风险：在入组前、试验期间或入组阶段，均无法提前预知参与者的分配。可由中央和独立的随机单元控制随机分配。研究人员不知道随机分配顺序（例如，分配顺序被隐藏在由顺序编号的、不透明和密封的信封中）。

②不清楚的偏倚风险：研究作者没有描述用于隐藏分配的方法，因此可能在入组前或入组期间可预见到干预分配。

③高偏倚风险：分配参与者的研究人员很可能知道分配顺序。对于这部分研究，我们仅纳入其评估干预措施的风险。

(3) 对参与者及主要研究者的施盲 (Blinding of participants and personnel) :

①低偏倚风险：符合以下任何一项：确保参与者和关键研究人员的施盲，并且盲法不太可能被打破；虽然没有施盲或不完全施盲，但评价者认为结果不太可能受到缺乏施盲的影响，例如死亡率。

②不清楚的偏倚风险：符合以下任何一项：信息不足以判断低风险或高风险；或者研究没有处理这个问题。

③高偏倚风险：符合以下任何一种：没有施盲或不完全施盲，结果很可能受到缺乏施盲的影响；或研究参与者和关键人员尝试施盲，但可能是施盲已被打破，结果可能受到缺乏施盲的影响。

(4) 对研究结果评估者的施盲 (Blinding of outcome assessment) :

①低偏倚风险：符合以下任何一项：确保结果评估的施盲，并且不太可能破坏施盲；没有结果评估的施盲，但评价者认为结果测量不太可能受到缺乏施盲的影响，例如死亡率。

②不清楚的偏倚风险：符合以下任何一项：信息不足以判断低风险或高风险；或者研究没有处理这个问题。

③高偏倚风险：符合以下任何一种：结果评估无施盲，结果测量可能受到缺乏施盲的影响；或结果评估施盲，但可能已经破坏了施盲，并且结果测量可能受到缺乏施盲的

影响。

(5) 结局指标数据的不完整性 (Incomplete outcome data) :

①低偏倚风险: 缺失的数据不太可能使治疗效果偏离合理的效应值。该研究使用了足够的方法, 例如多重填补, 来处理缺失的数据。

②不清楚的偏倚风险: 没有足够的信息来评估缺失数据以及与用于处理缺失数据的方法可能导致结果产生偏倚。

③高偏倚风险: 结果可能会由于数据缺失而产生偏倚。

(6) 选择性报告 (Selective reporting) :

①低偏倚风险: 试验报告了以下预定结局指标: 全因死亡率、严重不良事件和健康相关的生活质量。如果可获得原始试验方案, 则结果应该是该方案中要求的结局指标。如果试验方案是从试验登记平台 (例如 www.clinicaltrials.gov) 获得, 则如若试验方案是在试验开始之前或试验开始时登记的, 则方案中列举的结局指标可被认为是试验最初预定要求的结局指标。如果试验方案在试验开始后才进行登记, 那么这些结局指标就不被认为是可靠的。

②不清楚的偏倚风险: 并非所有预定的结局指标都被完全报告, 或者不清楚是否记录了这些结局指标的数据。

③高偏倚风险: 一项或多项预定的结局指标未被报告。

(7) 其他偏倚 (Other bias) :

①低偏倚风险: 该研究似乎没有其他可能使其面临偏见风险的因素。

②高偏倚风险: 如果上述偏倚源的任何风险分类为不明确的偏倚风险或高偏倚风险, 则结果结果将被归类为高偏倚风险。

3.4 效应测量

运用相对危险度 risk ratio (RR) 测量二分类变量结局指标, 运用均差 mean difference (MD) 测量连续变量, 两者均以 95%置信区间 (95%CI) 表示。当运用不同仪器测量同一连续变量时, 运用标准均差及 95%CI 测量。

3.5 分析单位

根据干预措施系统综述 Cochrane 手册^[85], 拟定数据分析单位为随机分配入试验的参与者个体。对于具有多个试验组的临床试验, 我们纳入符合本研究内容的组别。如果不同的试验组共用同一对照组, 我们将对照组分为两组以避免重复计算。对于整群随机

对照试验，如果试验合理考虑到了组群设计，我们将从分析中直接提取数据。如果试验没有对组群进行控制，我们运用膨胀标准误及 Review Manager 5 软件中的倒方差法进行处理。对于交叉试验，我们只对第一阶段试验进行数据提取以避免残留治疗影响^[85]。

3.6 缺失数据处理

联系研究者以获得纳入的研究中报告不详细的数据。我们运用意向性治疗方法进行数据分析，当无法运用此方法时，我们使用可以运用的数据进行分析。对于所有的结局指标，我们纳入不完整或有缺失数据的参与者，并在敏感性分析中按如下方法进行数据填补：

(1) 二分类变量

①最好-最坏情况：假设试验组所有失访参与者全部存活、临床症状全部改善、全部没有严重不良事件，全部没有并发症发生；对照组中所有失访参与者全部死亡，临床症状全未改善，全部发生严重不良事件，全部发生并发症。

②最坏-最好情况：假设试验组所有失访参与者全部死亡，临床症状全未改善，全部发生严重不良事件，全部发生并发症；对照组所有失访参与者全部存活、临床症状全部改善、全部没有严重不良事件，全部没有并发症发生。

(2) 连续变量

对于获益性结局指标在组别均值基础上加 2 个标准差，而对于风险性结局指标在组别均值上减去 2 个标准差^[98]。如果试验未报道标准差值，则联系作者以获得相关信息或者运用试验所报告数据进行计算。

3.7 异质性评估

通过仔细分析试验参与者特征和纳入研究的设计以评估纳入研究的临床异质性和方法学异质性。如果可获得相关数据，则通过对比纳入研究中不同的参与者诊断标准、参与者的合并疾病、小柴胡汤方类型、方剂剂型、不同疗程、小柴胡汤不同剂量、联合干预措施、不同试验的对照组以及不同随访期的干预效应对比，对纳入研究的临床异质性进行分析。

不同的研究设计类型和偏倚风险可导致方法学异质性。我们通过比较固定效应模型和随机效应模型结果评估统计学异质性。首先我们通过观察森林图寻找统计学异质性征象；其次，我们运用 Chi^2 检验观察 I^2 值及 P 值判断异质性程度^[86, 95, 96]。通常， I^2 在 0%–40% 表明异质性可能不重要；30%–60% 代表中等程度异质性；50%–90% 代表大量异质性；

75%–100%表明重大异质性^[85]。

在 TSA 分析中，我们运用 D^2 进行多样性校正后的信息量 (DARIS) 分析^[97]。

3.8 报告偏倚评估

如果每项比较包含至少 10 项研究，我们通过评估倒漏斗图的对称性评估报告偏倚。对于二分类变量，我们运用 Harbord 检验评估对称性；对于连续变量，我们运用不对称回归检验评估对称性。

3.9 数据合成

3.9.1 Meta-分析

我们依据干预措施系统综述的 Cochrane 手册以及 Cochrane 肝胆病组模板实施 meta-分析^[85]。我们运用 Review Manager 5 软件进行数据分析^[87]。我们既运用固定效应模型也运用随机效应模型评估我们的干预措施，当结果不一致时（例如其中一个模型得出显著性干预效应，而另一个模型得出无显著性干预效应），我们报告两个模型的分析结果，但是我们将更多的权重赋予更接近无干预效应的结果（即最高 P 值结果）^[98]。

为了使多重比较谬误低于 0.05，我们评估三个主要结局指标和三个次要结局指标时，以 P 值 ≤ 0.025 为显著统计学差异^[98]。对于探索性指标，我们认为 P 值 < 0.05 为显著统计学差异。我们通过评估纳入研究的统计学异质性和临床异质性来决定是否进行 Meta 分析或进行描述性分析。

我们在敏感性分析中进行缺失数据填补（参见敏感性分析）^[98]。当仅可获得一项研究的数据时，我们运用 Fisher 检验检测二分类变量^[99]，运用 Student's t 检验检测连续变量^[100]。

3.9.2 试验序贯分析

由于小样本数据和重复测量，累积 meta 分析可能存在随机误差风险，因此我们运用试验序贯分析。为了减少由随机误差导致的对于干预措施效应的过高或过低估计，我们计算“所需样本量”（required information size, RIS）（即接受或拒绝干预措施效果在 Meta 分析中所需要的参与者数量）^[89, 101, 102]。RIS 计算应该同时考虑 meta 分析中的多样性^[101, 103, 104]。

我们对二分类变量和连续变量控制一类错误和二类错误^[101, 103-107]。对于二分类变量，我们依据对照组的事件发生比例估计根据多样性校正的所需样本量 (diversity-adjusted required information size, DARIS)，相关风险降低率为 15%，主要结局指标

和次要结局指标的 α 值为 2.5%，探索性指标结局指标的 α 值 5.0%， β 值为 10%^[108]，多样性为 meta 分析中纳入研究的异质性^[98, 103]。对于连续性变量，我们根据对照组观察到的标准差 (SD) 估计 DARIS，最低标准差差异为 50%SD 值，主要结局指标和次要结局指标的 α 值为 2.5%，探索性指标结局指标 α 值 5.0%， β 值为 10%^[108]，多样性同样为 meta 分析中纳入研究的异质性^[98, 103]。

我们运用统计检测边界检测有效性和有害性的统计显著性差异，并运用无效性边界检测无效性^[109]。如果累积检验统计量 (Z 曲线) 在到达 DARIS 前穿过统计检测边界，则可认为干预措施的效应优于或者差于对照组干预措施。如果 Z 曲线在到达 DARIS 前穿过无效性检验，则意味着干预措施不具备假设效应，并且进一步的随机对照试验可能是无意义的。

如果试验序贯分析的统计检测边界和无效边界都没有被穿过，则仍有必要进一步行临床试验以探究是否应接受或者拒绝干预措施效应^[101, 102]。

4. 亚组分析和异质性检测

如果能够提取相关数据，则进行如下亚组分析：

总体低偏倚风险的研究与总体高偏倚风险的研究；结局指标盲法被评估为低偏倚风险的研究与其被评估为高偏倚风险的研究；不完整结局指标数据被评估为低偏倚风险的研究与其被评估为高偏倚风险的研究；选择性结局指标报告被评估为低偏倚风险的研究与其被评估为高偏倚风险的研究；对照组运用不同干预措施的研究：即对照组无干预措施与对照组使用安慰剂干预措施的研究；经典小柴胡汤与小柴胡汤加减方的研究；小柴胡汤不同剂型的研究；干预措施不同疗程和不同剂量的研究（根据观测到的均值进行分层）；参与者诊断标准不同的研究；只患有慢性乙型肝炎参与者与患有其他合并疾病参与者（肝硬化，干细胞癌，HIV 感染，AIDS，丙肝，丁肝或多个合并疾病者）的研究（我们将单独分析每个合并疾病）。

5. 敏感性分析

除了缺失数据处理部分的敏感性分析，我们还比较 GRADE 分级中的不精确性评估与试验序贯分析中的不精确性评估^[98]。必要时可进行其他敏感性分析^[98, 110, 111]。

6. GRADE 分级

构建 SOF 表以展示主要结局指标和次要结局指标发现的证据结果。运用 GRADEpro GDT 软件 (community.cochrane.org/help/tools-and-software/grade-pro-gdt) 评估证

据质量的五个方面：研究间偏倚风险评估、证据的间接性（参与者，干预措施，对照，结局）、结局不能解释的异质性和非一致性（包含亚组分析的问题）、结局的不精确性、发表偏倚风险^[112, 113, 114]。

证据质量定义如下^[112, 113, 114]：高确定性（High certainty）：我们非常确定真实效应与估计效应非常接近；中等确定性（Moderate certainty）：我们对估计效应中等程度确定，即真实效应很可能与估计效应非常接近，但是可能存在大量差异；低确定性（Low certainty）：我们对于估计效应的确定程度有限：真实效应可能与估计效应存在较大差异。极低确定性（Very low certainty）：我们对于估计效应非常不确定：真实效应非常可能与估计效应存在较大差异。

研究结果

1. 纳入研究描述

1.1 文献检索结果

通过电子检索出 311 条记录（见图 1-1）。我们通过检索纳入文献的参考文献，我们检索出 65 条参考文献。在排除 235 篇重复发表或明显不相关的参考文献，我们阅读了 141 篇发表研究的全文。84 篇研究未符合纳入标准。在联系研究人员后，43 篇研究没有提供明确的随机序列产生方法，并且 4 项研究由于缺少足够信息以判断是否是随机对照试验，这 47 项研究被排除我们的 meta 分析。未检索到正在进行的或者未发表的研究。

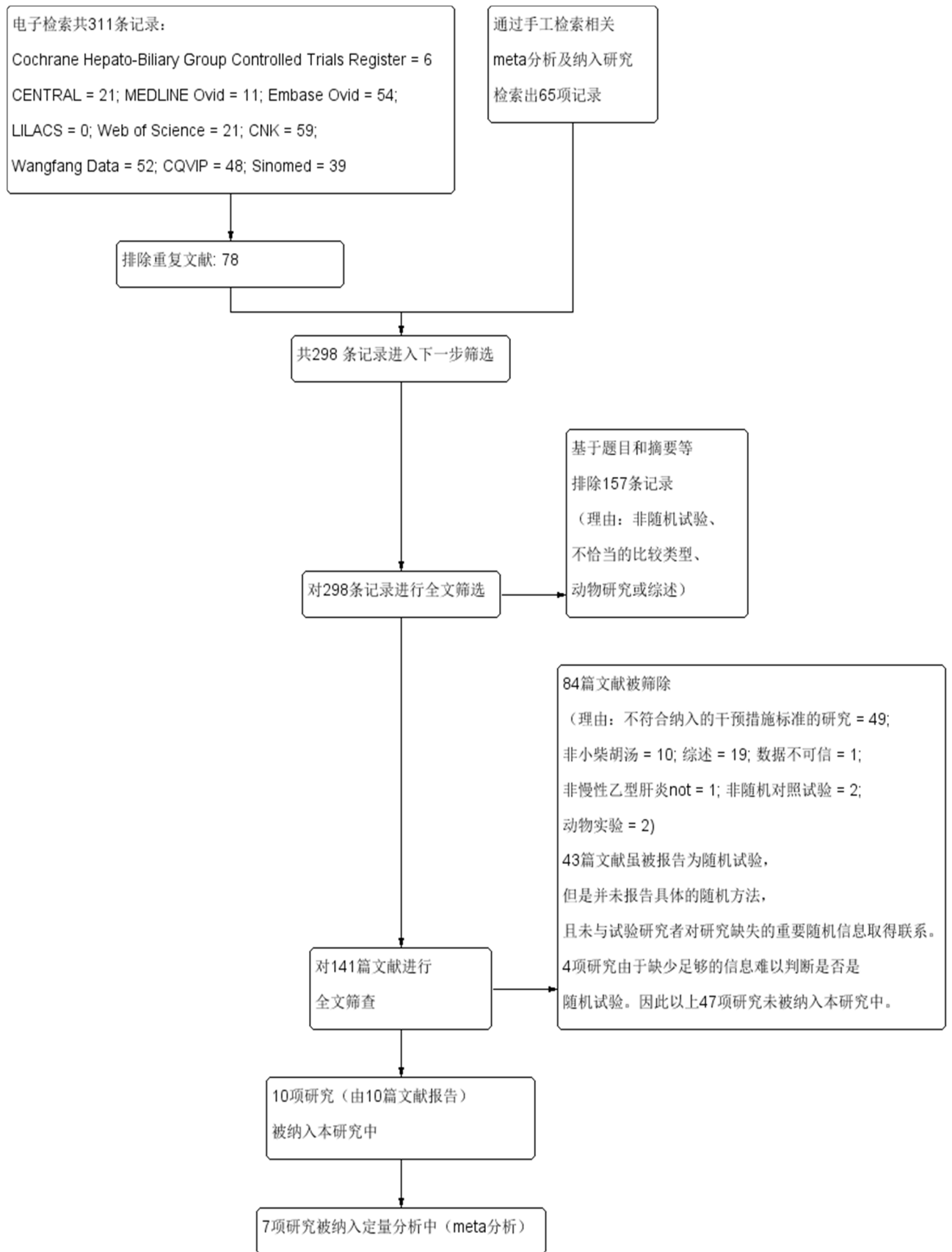


图 1-1 小柴胡汤治疗慢性乙型肝炎纳入研究流程图

1.2 纳入研究的一般特征

本研究共纳入 10 项临床研究共涉及 926 名参与者。仅成功联系到 1 项研究的研究者^[115]。其中 7 项研究为 meta 分析提供数据。其余 3 项研究并未对本研究的目标结局指标进行报告，因此对此 3 项研究进行描述性分析。所有研究地点均为中国大陆，报告语言为普通话。纳入研究的样本含量 79-160 例，平均年龄为 33-64 岁，男女比例 512: 335，1 项研究的受试者合并肺结核^[122]，1 项研究的受试者合并肝硬化^[123]，2 项研究纳入肝郁脾虚证患者^[126, 131]，1 项研究纳入肝郁脾虚血瘀证^[129]。纳入研究的疾病类型为慢性乙型肝炎。纳入研究均为小柴胡汤联合常规疗法对经常规疗法。5 项研究运用小柴胡汤原方，3 项研究运用小柴胡汤加减方，2 项研究未报告小柴胡汤组成。1 项研究运用小柴胡汤颗粒剂，9 项研究运用小柴胡汤水煎剂。小柴胡汤中君药小柴胡剂量 6-25g 不等，疗程 3-8 个月不等。纳入研究的随访期均等同于干预期。纳入研究具体特点参见表 1-1。

表 1-1 小柴胡汤治疗慢性乙型肝炎随机对照试验临床特征

研究 ID	疗程 (月)	试验组例数	对照组例数	男: 女	平均年龄 (岁)	诊断标准	试验组 干预措施	对照组 干预措施	结局指标
Chen 2014 ^[122]	3	60	60	73:47	34.32	《慢性乙型肝炎防治指南》《肺结核诊断和治疗指南》(无参考文献)	小柴胡汤原方+ 常规治疗	常规治疗	胸部 X 线, HBV-DNA 水平, 肝功能指标变化 (ALT, AST, TBIL)
Chen 2017 ^[123]	3	32	32	39:25	试验组 32.5 对照组 35.2	西医: [116][117] 中医: [118]	小柴胡汤加减方+ 常规治疗	常规治疗	肝功能指标 (ALT, AST, TBIL), 肝纤维化指标 (HA, LN, PCIII, IV-C)
Kang 2016 ^[124]	8	40	41	52:29	试验组 41.89 对照组 42.33	未报告	小柴胡汤原方+ 常规治疗	常规治疗	复合指标: 临床有效率, ALT 复常率, HBeAg 阴性率和应答率, HBV-DNA 阴性率和应答率

Liu 2017 ^[125]	3	40	40	47:33	试验组 63.5 对照组 63.6	未报告	小柴胡汤原方+ 常规治疗	常规治疗	总体有效率, 非严重不良事件
Mao 2014 ^[126]	3	47	21	47:21	试验组 49.89 对照组 50.1	西医:[119] 中医:[118]	小柴胡汤方(具体 不详)+常规治疗	常规治疗	中医证候积分表
Sun 2004 ^[117]	3	60	40	66:34	试验组 43.2 对照组 42.1	[120]	小柴胡汤原方+ 常规治疗	常规治疗	肝功能指标改善(ALT, AST, ALB, TBiL), 肝脏病理变化指标(HA, PC III, LN, IV-C)
Wang 2013 ^[128]	4	80	80	78:82	试验组 40.2 对照组 42.1	未报告	小柴胡汤原方+ 常规治疗	常规治疗	非严重不良事件, 生化指标水平(ALT, AST, ALB, GLB), 肝纤维化改善(具体不详), HBV-DNA 阴转率(但无具体数据)
Wang 2016 ^[129]	3	41	41	55:27	试验组 35.1 对照组 35.8	[117][119]	小柴胡汤加减方+ 常规治疗	常规治疗	肝功能指标改善(ALT, AST, TBiL), HBV-DNA 拷贝数变化, 复合指标: 应答率
Wu 2015 ^[130]	6	46	46	55:37	试验组 44.1 对照组 44.5	未报告	小柴胡汤方(具体 不详)+常规治疗	常规治疗	肝功能指标改善(ALT, AST, ALB, TBiL), 门静脉动力学(门静脉宽度及门静脉流速)
Zhao 2014 ^[132]	12	40	39	未报告	试验组 38.1 对照组 40	[121]	小柴胡汤加减方+ 常规治疗	常规治疗	肝功能指标改善(ALT, AST, TBiL), 乙肝病毒标志物, HBsAg 阴转率, HBV-DNA 阴转率, 临床症状改善

2. 纳入研究的偏倚风险评估

我们通过联系作者及文献提供信息, 对纳入研究进行风险偏倚评估(见图 1-2, 图 1-3)。

2.1 随机分配(Random sequence generation and allocation concealment 选择偏倚)

纳入研究均使用电脑或随机数字表产生随机序列，但没有研究报告如何进行随机分配。因此，所有研究均存在选择偏倚风险。

2.2 盲法 (Blinding of participants and personnel and Blinding of outcome assessment 实施偏倚及检测偏倚)

纳入研究均未报告参与者和/或研究者盲法。我们评估所有纳入研究为实施偏倚高风险。其中一项纳入研究明确报告没有实施结局指标评估施盲 (Sun 2004)，因此被评为高偏倚风险，其余研究未报告结局指标评估施盲，因此被评为不确定偏倚风险。

2.3 不完整的结局指标数据 (Incomplete outcome data 磨损偏倚)

9 项研究报告无缺失数据，并在数据分析中纳入所有参与者 (Sun 2004; Wang 2013; Chen 2014; Mao 2014; Wu 2015; Kang 2016; Wang 2016; Chen 2017; Liu 2017)，因此被评为低风险偏倚。Zhao 2014 在数据分析时排除了 8 名参与者。我们将其评估为高风险偏倚。

2.4 选择性报告 (Selective reporting 报告偏倚)

由于缺少提前发表的研究方案以及无死亡率、严重不良事件、健康相关生存质量相关数据报告，所有纳入的研究均被视为高风险偏倚。

2.5 其他潜在偏倚 (Other bias)

所有纳入研究不存在其他潜在偏倚风险，因此被评估为低风险偏倚。

2.6 总体偏倚风险

我们将纳入的研究均评估为总体高风险偏倚。

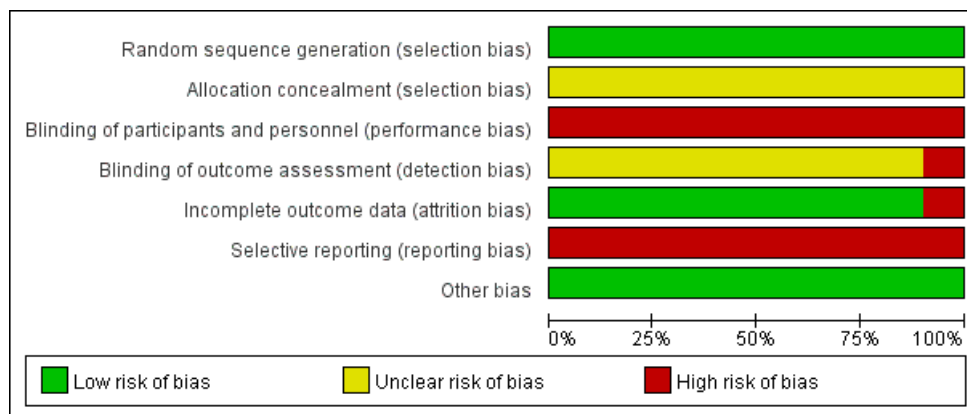


图 1-2 小柴胡汤治疗慢性乙型肝炎随机对照试验总体偏倚风险评估情况

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Chen 2014	+	?	-	?	+	-	+
Chen 2017	+	?	-	?	+	-	+
Kang 2016	+	?	-	?	+	-	+
Liu 2017	+	?	-	?	+	-	+
Mao 2014	+	?	-	?	+	-	+
Sun 2004	+	?	-	-	+	-	+
Wang 2013	+	?	-	?	+	-	+
Wang 2016	+	?	-	?	+	-	+
Wu 2015	+	?	-	?	+	-	+
Zhao 2014	+	?	-	?	-	-	+

图 1-3 小柴胡汤治疗慢性乙型肝炎各随机对照试验偏倚风险评估详情

3. 干预措施效应评估

3.1 主要结局指标

纳入研究均未报告本系统综述主要结局指标：全因死亡率、一项或多项不良事件的发生率、健康相关生存质量。

3.2 次要结局指标

纳入研究均未报告本系统综述次要结局指标：乙肝相关死亡率、乙肝相关发病率。

3.2.1 一项或多项非严重不良事件发生率

两项研究涉及 240 例参与者报告该指标数据（Wang 2013; Liu 2017）。我们不确定小柴胡汤是否对一项或多项非严重不良事件发生率有效应（RR=0.43，95% CI [0.02,11.98]; P = 0.62; I² = 69%; 图 1-4）。两项研究的异质性较大，通过分析研究的临床特点发现，Wang 2013 报告患者除乙型肝炎外患有其他疾病，Liu 2017 未报告是否有合并疾病，可能是异质性的来源。

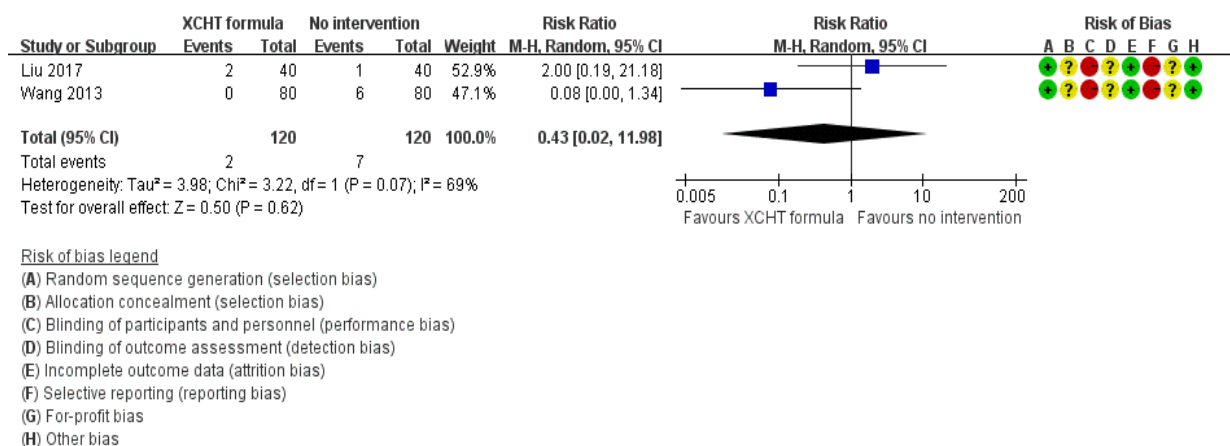


图 1-4 小柴胡汤治疗慢性乙型肝炎一项或多项非严重不良事件发生率森林图

试验序贯分析

对于该结局指标，通过试验序贯分析获得的多样性校正后的所需样本量（DARIS）为 110,984 名参与者。我们运用如下参数计算所需样本量：对照组事件发生比例 15%，α 值为 2.5%，把握度为 90%，多样性为 70%。由于只有 0.22%（240/110984）名患者提供信息，统计差异检测边界未能获得（Wang 2013; Liu 2017）。因此，对于该指标，试验序贯分析显示无充分证据支持或者拒绝小柴胡汤联合常规疗法对比单用常规疗法可降低 15% 风险。

亚组分析

我们仅能针对不同合并疾病患者进行亚组分析。我们未发现亚组间显著统计学差异（亚组差异检测： $\text{Chi}^2 = 2.97$ ； $P = 0.09$ ； $I^2 = 66.3\%$ ；图 1-5）。

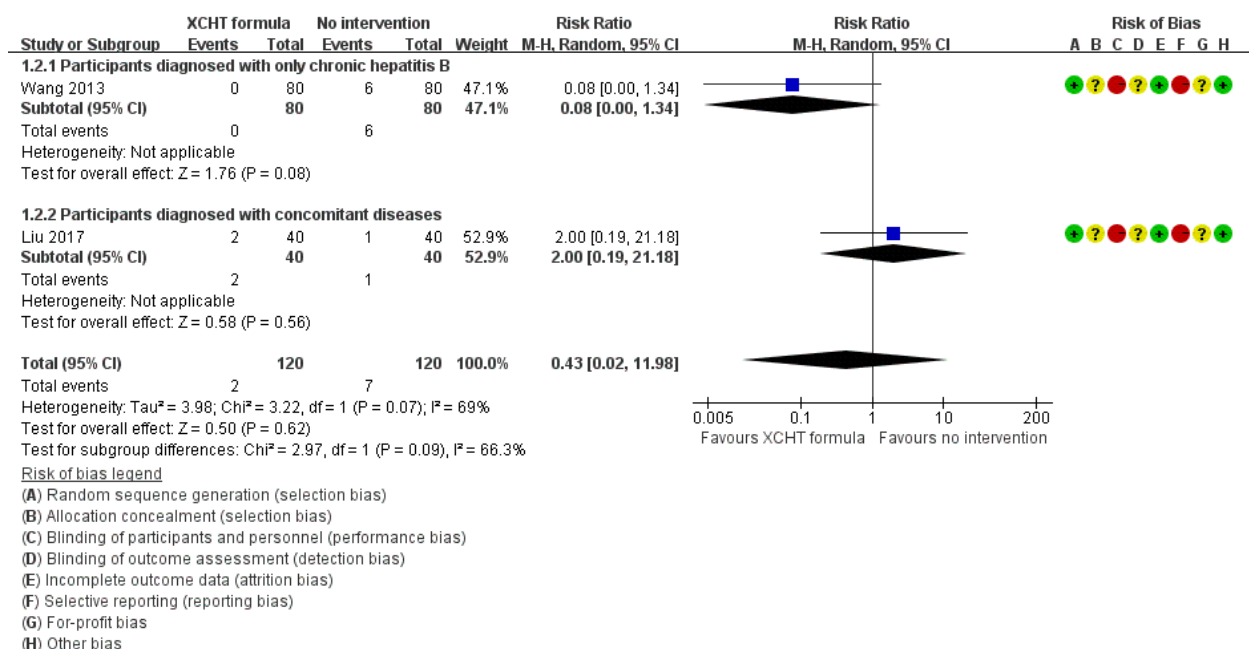


图 1-5 小柴胡汤治疗慢性乙型肝炎随机对照试验一项或多项非严重不良事件发生率-不同合并疾病森林图

敏感性分析

两项报告该指标的研究将所有随机分组的参与者的数据纳入数据分析中。因此，不涉及‘最好-最坏’情况或‘最坏-最好’情况分析。

运用 GRADE 分析和 TSA 分析评估关于该指标证据的不精确性，得到一致结果：运用 GRADE 分析对证据质量不精确性评估时，由于事件发生数小于 300，并且可信区间 CI 包含了无效线以及显著疗效 ($\text{RR} < 0.75$) 或显著伤害 ($\text{RR} > 1.25$)，对证据质量降两级^[131]；运用 TSA 分析对证据质量不精确性评估时，由于统计差异边界（有益、有害及无效）均未被穿过^[98]，对证据质量降两级。

3.3 探索性结局指标

3.3.1 HBV-DNA 阳性的参与者比例

三项研究涉及 222 例参与者报告该指标数据（Wang 2013；Zhao 2014；Kang 2016）。

小柴胡汤可降低 HBV-DNA 阳性检测率 (RR=0.62, 95% CI [0.45, 0.85]; I²=0%; P=0.003 (图 1-6)。

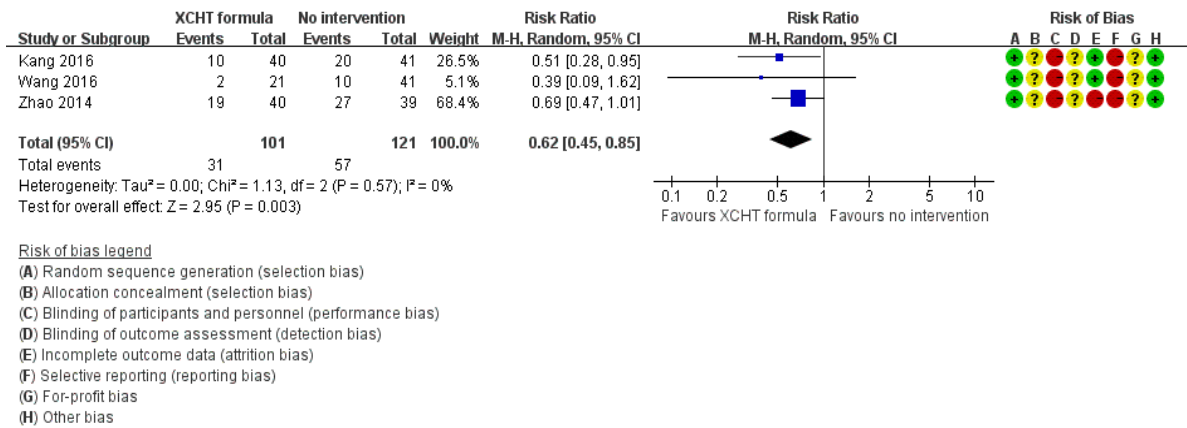


图 1-6 小柴胡汤治疗慢性乙型肝炎 HBV-DNA 阳性的参与者比例森林图

试验序贯分析

对于该结局指标，通过试验序贯分析获得的 DARIS 为 2068 名参与者。我们运用如下参数进行分析：对照组事件发生比例 47.1%，相对危险度减低 15%， α 值为 5%，把握度为 90%，多样性为 0%，结果显示 Z 曲线未到达 DARIS（2068 名参与者），也未穿过统计差异边界（有益、有害及无效）。试验序贯分析显示无充分证据支持或者拒绝小柴胡汤联合常规疗法对比单用常规疗法可降低 HBV-DNA 阳性患者比例 15% 风险（TSA 校正的 RR=0.62；95% CI [0.17, 2.80]；图 1-7）。

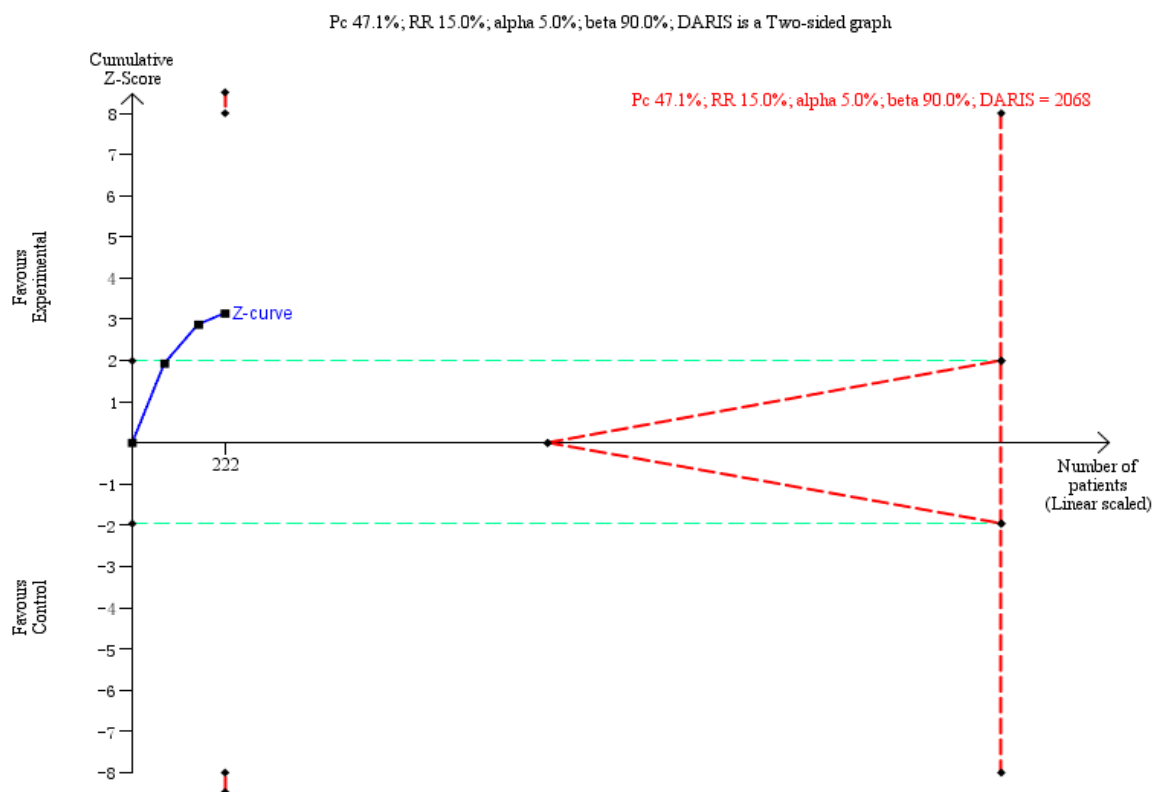


图 1-7 小柴胡汤治疗慢性乙型肝炎 HBV-DNA 阳性的参与者比例试验序贯分析图

亚组分析

由于无法获得全部预定结局指标数据，因此无法进行全部预定亚组分析。在如下亚组分析中我们未发现亚组间显著统计学差异：比较具有高风险和低风险不完整结局指标数据研究（ $\text{Chi}^2=0.91$ ； $P=0.34$ ； $I^2=0\%$ ；图 1-8）；比较经典小柴胡汤和小柴胡汤加减方的研究（ $\text{Chi}^2=0.01$ ； $P=0.92$ ； $I^2=0\%$ ；图 1-9）；比较小柴胡汤水煎剂和小柴胡汤颗粒剂研究（ $\text{Chi}^2=0.91$ ； $P=0.34$ ； $I^2=0\%$ ；图 1-10）；比较君药小柴胡汤剂量高于 15g 和小柴胡汤剂量低于 15g（ $\text{Chi}^2=0.47$ ； $P=0.49$ ； $I^2=0\%$ ；图 1-11）；比较疗程小于 6 个月和疗程大于 6 个月的研究（ $\text{Chi}^2=0.42$ ， $P=0.52$ ， $I^2=0\%$ ；图 1-12）；比较依据研究者诊断为慢性乙型肝炎的研究和根据指南诊断为慢性乙型肝炎的研究（ $\text{Chi}^2=0.04$ ； $P=0.84$ ； $I^2=0\%$ ；图 1-13）。

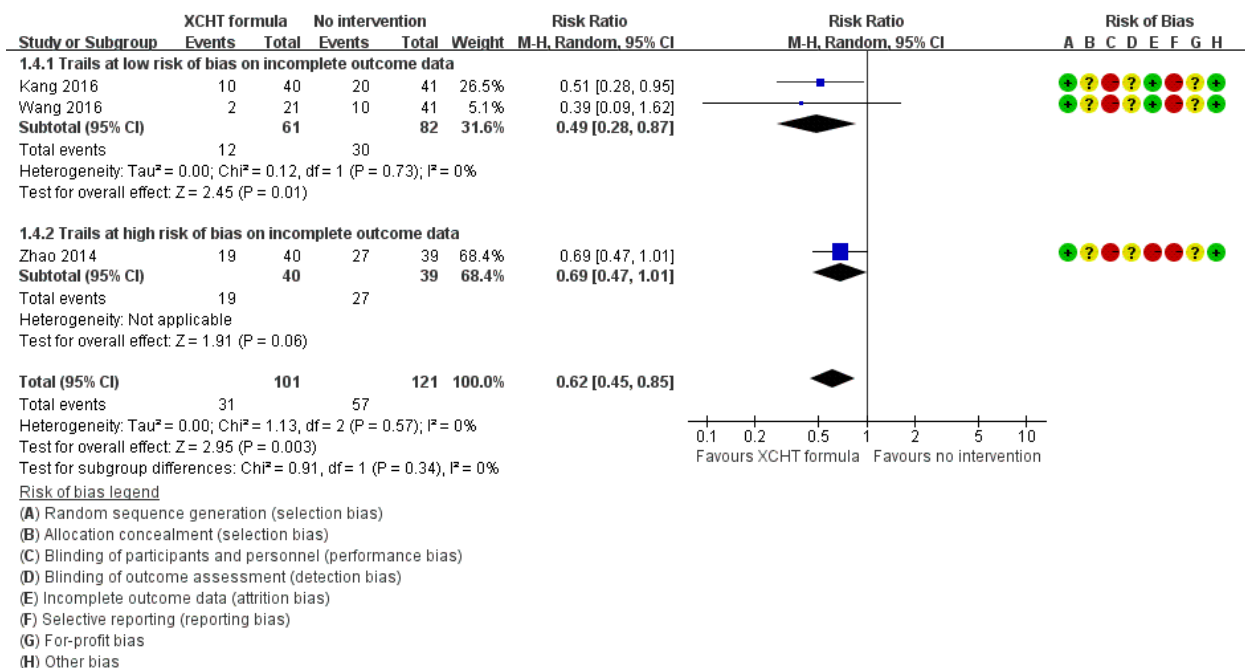


图 1-8 小柴胡汤治疗慢性乙型肝炎 HBV-DNA 阳性的参与者比例-
不完整结局指标数据亚组分析森林图

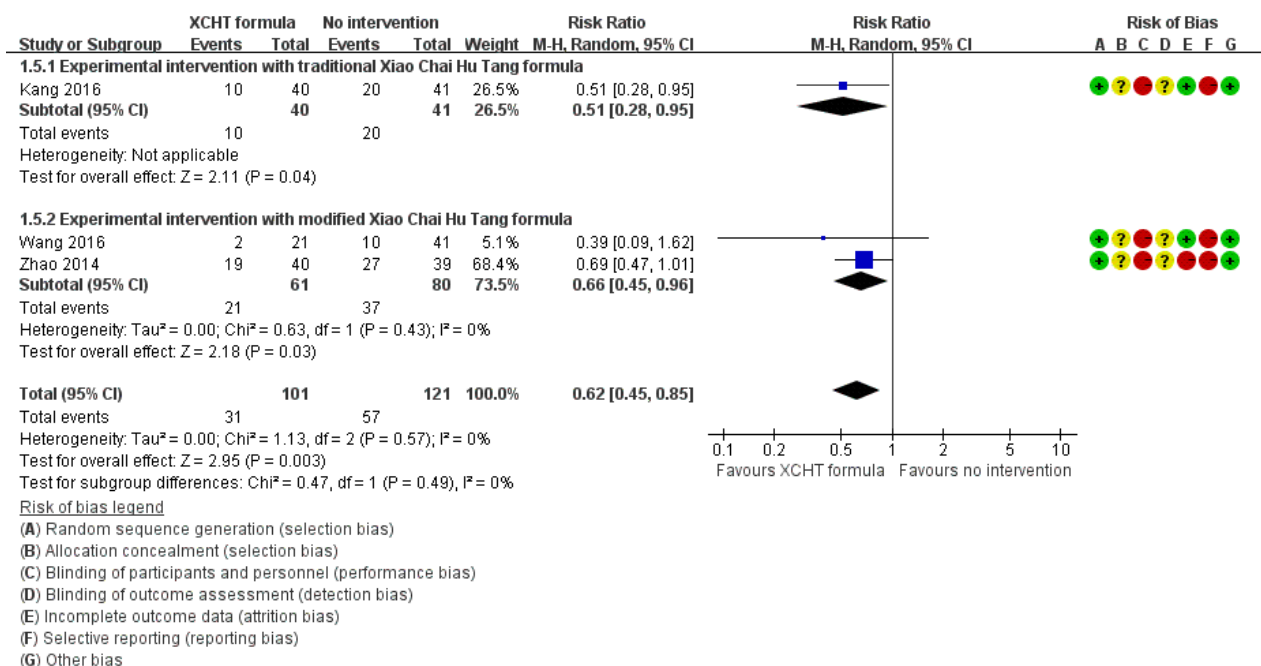


图 1-9 小柴胡汤治疗慢性乙型肝炎 HBV-DNA 阳性的参与者比例-
小柴胡汤类型亚组分析森林图

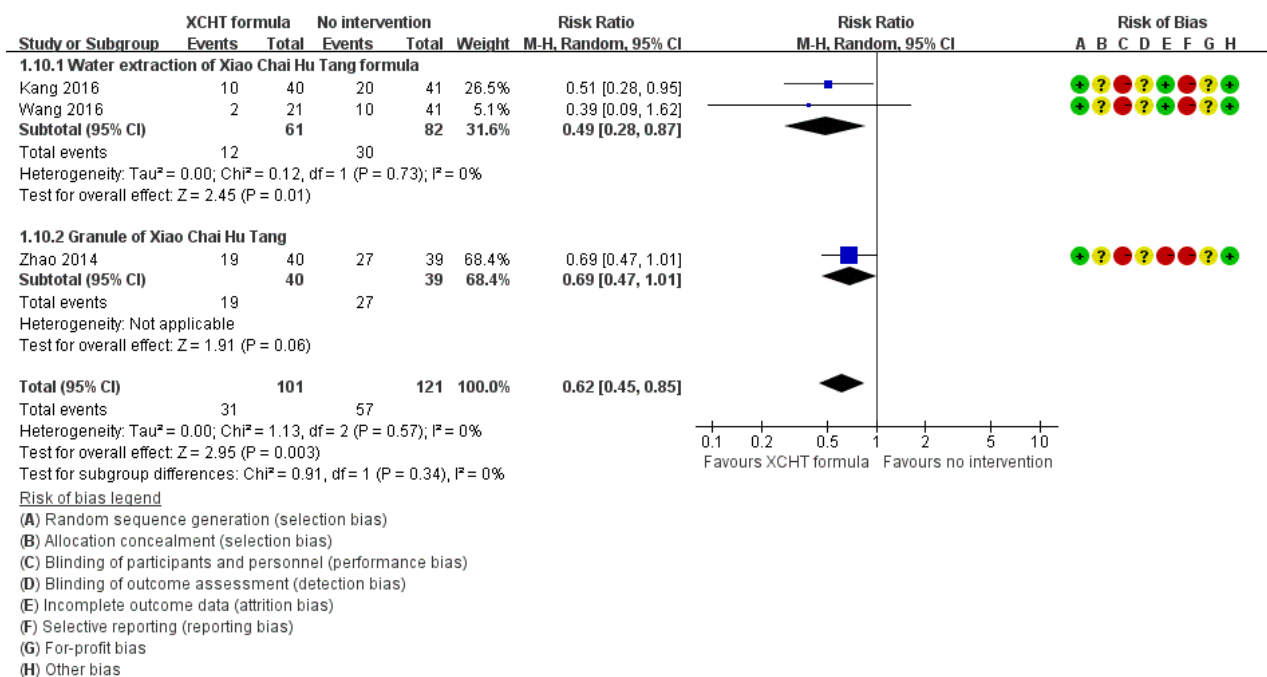


图 1-10 小柴胡汤治疗慢性乙型肝炎 HBV-DNA 阳性的参与者比例-小柴胡汤剂型亚组分析森林图

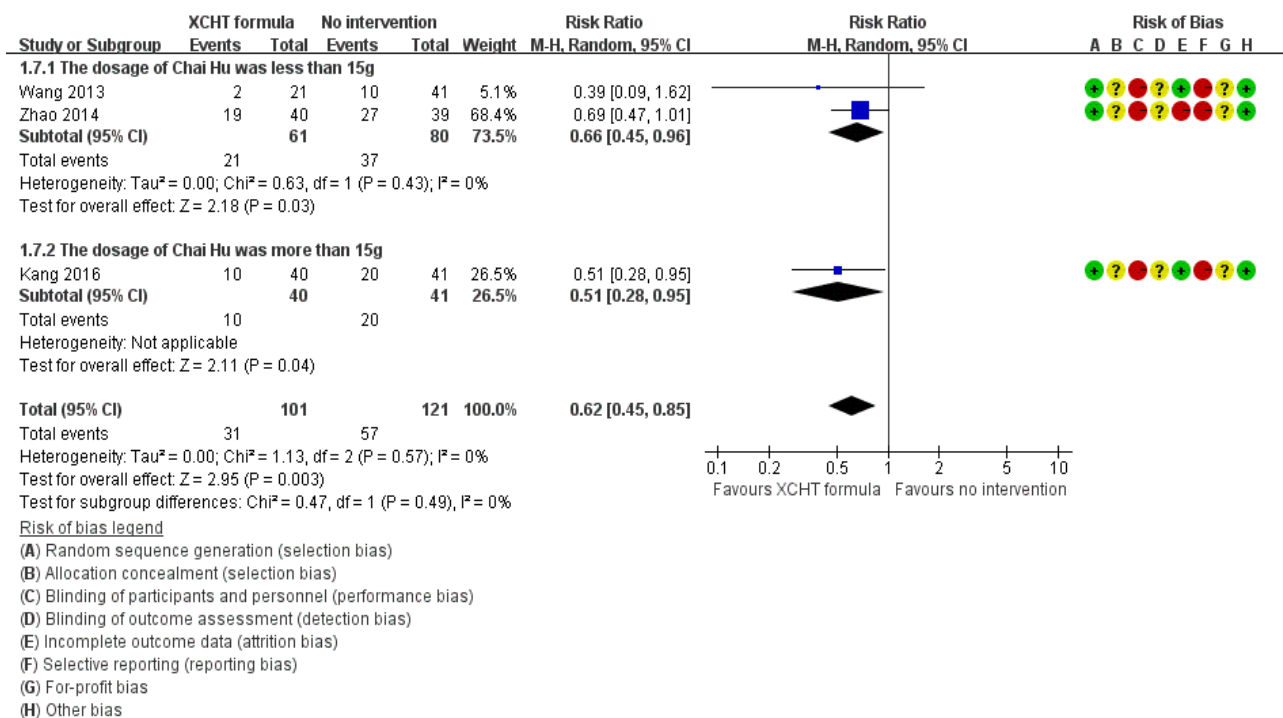


图 1-11 小柴胡汤治疗慢性乙型肝炎 HBV-DNA 阳性的参与者比例-小柴胡汤剂量亚组分析森林图

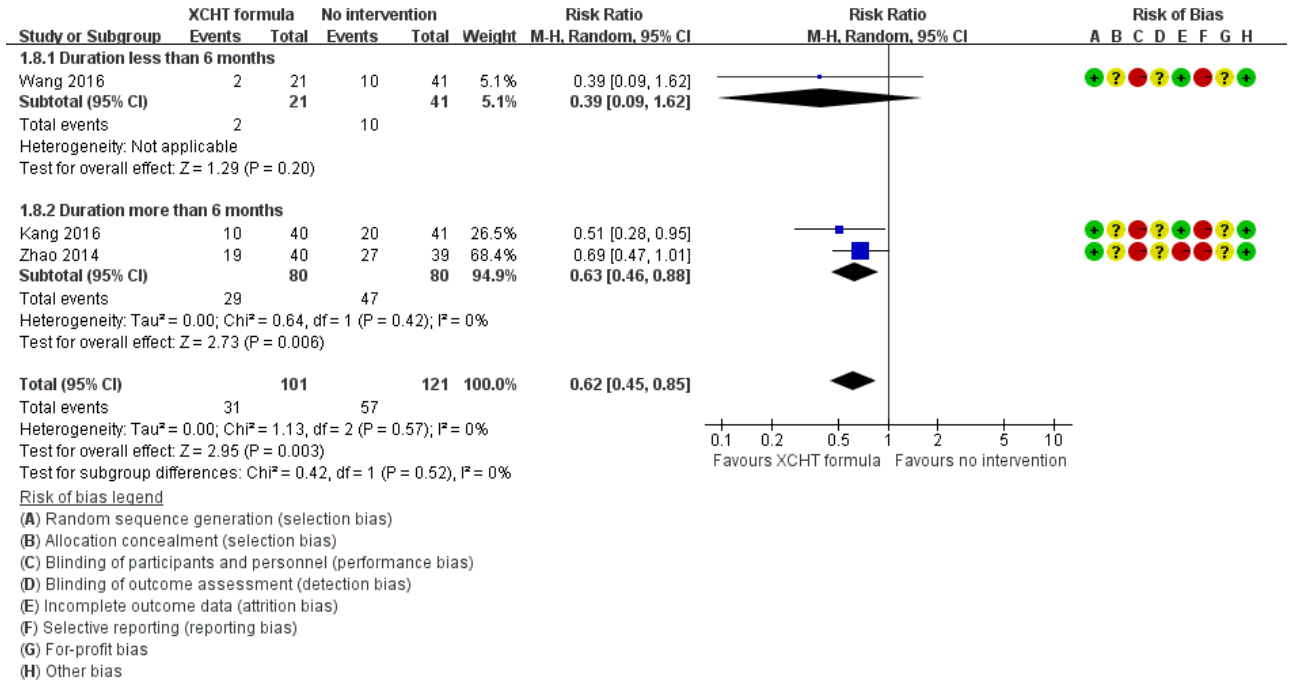


图 1-12 小柴胡汤治疗慢性乙型肝炎 HBV-DNA 阳性的参与者比例-小柴胡汤疗程亚组分析森林图

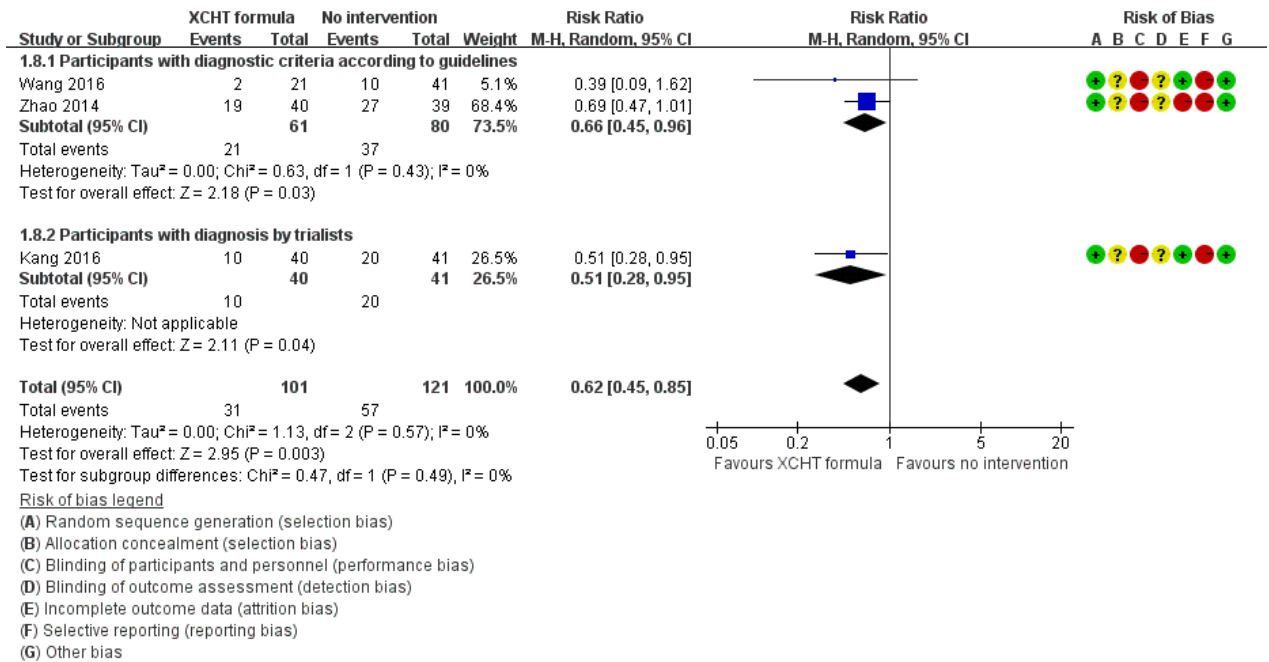


图 1-13 小柴胡汤治疗慢性乙型肝炎 HBV-DNA 阳性的参与者比例-不同诊断标准亚组分析森林图

敏感性分析

三项研究中均报告了所有参与者关于 HBV-DNA 阳性检出率的数据，我们没有进行敏感性分析 (Wang 2013; Zhao 2014; Kang 2016)。

运用 GRADE 分析和 TSA 分析评估关于该指标证据的不精确性，结果并不一致：运用 GRADE 分析对证据质量不精确性评估时，由于最优样本量未达到，样本量小于 4000，我们将证据质量降一级^[131, 132]；运用 TSA 分析对证据质量不精确性评估时，由于统计差异边界（有益、有害及无效）均未被穿过，对证据质量降两级^[98]。

3.3.2 HBeAg 阳性的参与者比例

两项研究涉及 160 例参与者报告该指标数据 (Zhao 2014; Kang 2016)。固定效应模型显示小柴胡汤可降低 HBeAg 阳性检测率 (RR=0.70, 95% CI [0.55, 0.91]; P=0.007; I²=38%; 图 1-14)。但是，随机效应模型显示小柴胡汤对 HBeAg 阳性检测率改变无显著统计学效应 (RR=0.72, 95%CI [0.50, 1.02]; P=0.06; I²=38%; 图 1-15)。随机效应模型结果是考虑了研究间异质性而分析得出的结果。我们对于更接近零效应的结果赋予更多的权重（即 RR=0.72, 95%CI [0.50, 1.02]; P=0.06），即小柴胡汤对 HBeAg 阳性检测率改变无显著统计学效应。两项研究具有一定的异质性，通过分析研究的临床特点发现，Zhao 2014 小柴胡汤柴胡剂量为 6g，Kang 2016 研究者柴胡剂量为 25g，可能是异质性的来源。

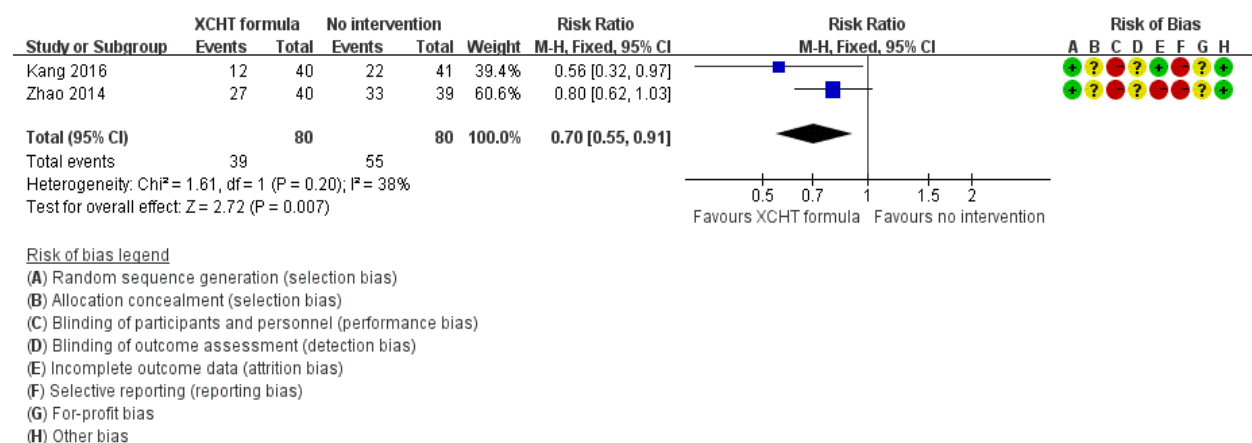
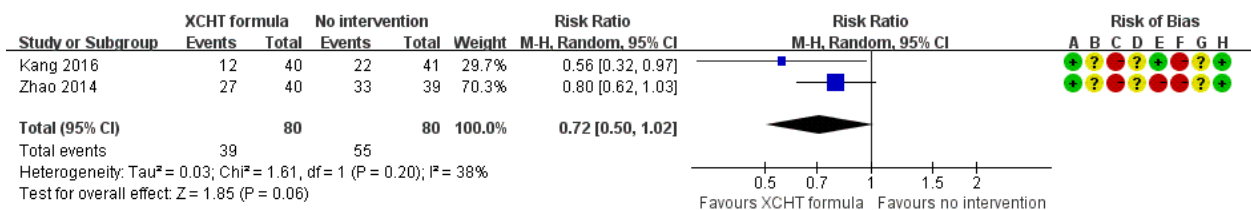


图 1-14 小柴胡汤治疗慢性乙型肝炎 HBeAg 阳性的参与者比例固定效应模型分析森林图



Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) For-profit bias
- (H) Other bias

图 1-15 小柴胡汤治疗慢性乙型肝炎 HBeAg 阳性的参与者比例随机效应模型分析森林图

试验序贯分析

我们运用如下参数进行分析：对照组事件发生比例 68.75%，相对危险度减低 15%， α 值为 5%，把握度为 90%，多样性为 57%，结果显示 Z 曲线未到达 DARIS（2143 名参与者），也未穿过统计差异边界（有益、有害及无效）。试验序贯分析显示无充分证据支持或者拒绝小柴胡汤联合常规疗法对比单用常规疗法可降低 HBeAg 阳性患者比例 15% 风险（TSA 校正的 RR=0.72；95% CI[0.17, 3.02]；图 1-16）。

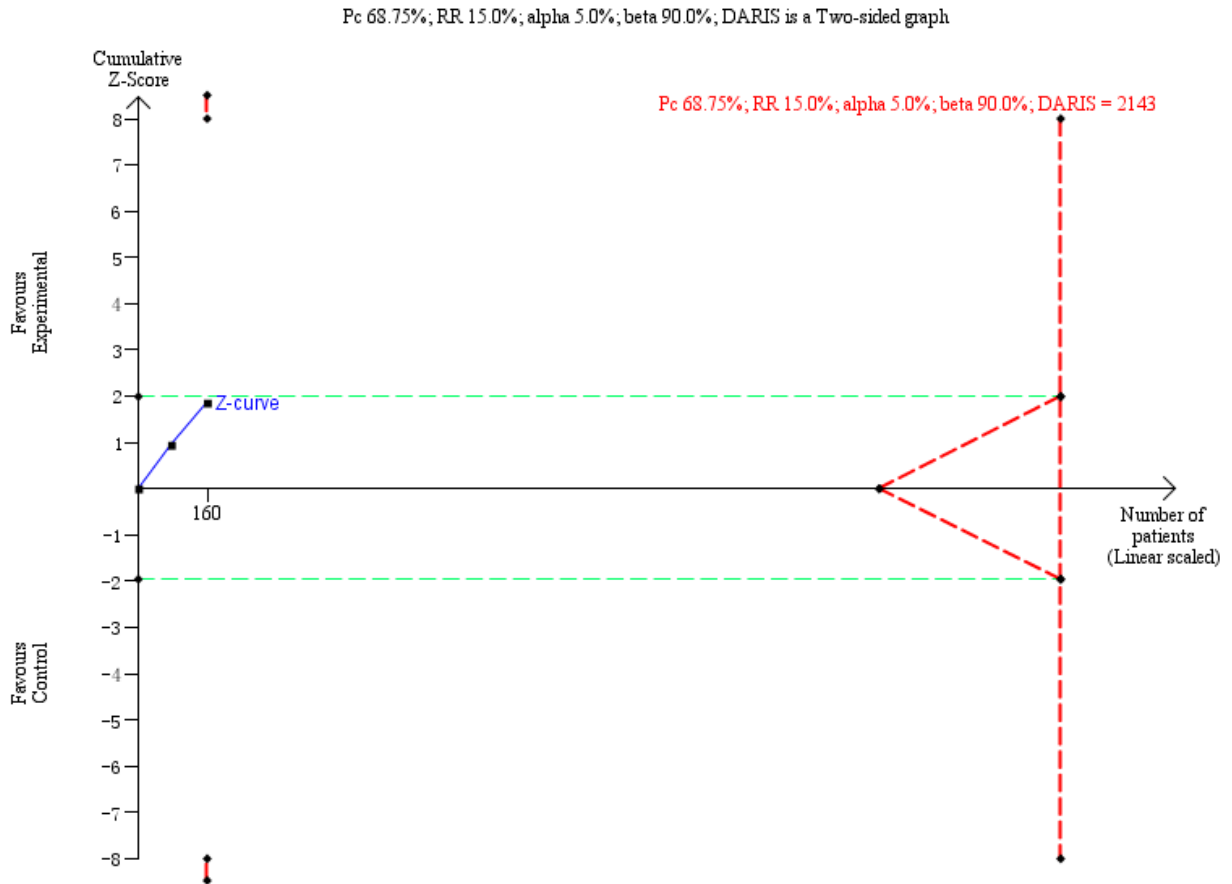


图 16 小柴胡汤治疗慢性乙型肝炎 HBeAg 阳性检测率试验序贯分析图

亚组分析

由于无法获得全部预定结局指标数据，因此无法进行全部预定亚组分析。如下亚组分析中，我们未发现亚组间显著统计学差异在：比较高风险和低风险不完整结局指标研究 ($\text{Chi}^2=1.31$; $P=0.25$; $I^2=23.9\%$; 图 1-17)；比较经典小柴胡汤和小柴胡汤加减方的研究 ($\text{Chi}^2=1.31$; $P=0.25$; $I^2=23.9\%$; 图 1-18)；比较小柴胡汤水煎剂和小柴胡汤颗粒剂研究 ($\text{Chi}^2=1.31$; $P=0.25$; $I^2=23.9\%$; 图 1-19)；比较君药小柴胡汤剂量高于 15g 和小柴胡汤剂量低于 15g 的研究 ($\text{Chi}^2=1.31$; $P=0.25$; $I^2=23.9\%$; 图 1-20)；比较疗程小于 6 个月和疗程大于 6 个月的研究 ($\text{Chi}^2=1.31$; $P=0.25$; $I^2=23.9\%$; 图 1-21)。

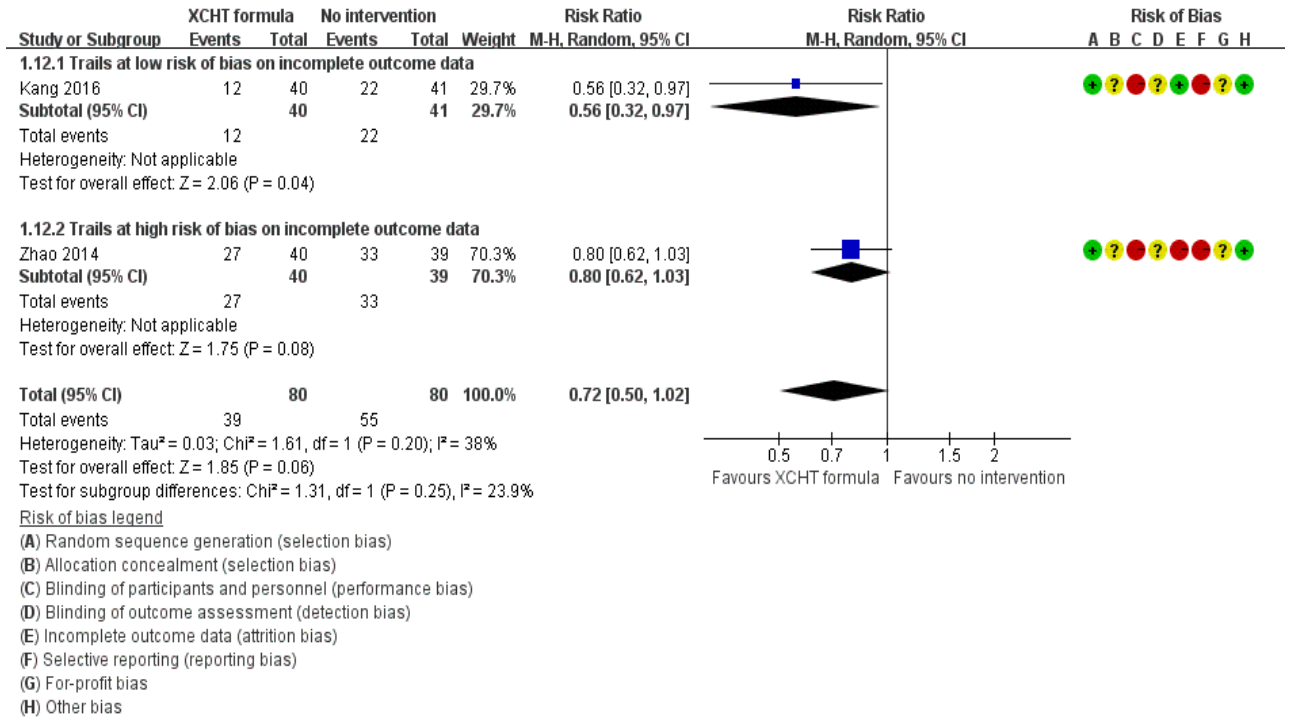


图 1-17 小柴胡汤治疗慢性乙型肝炎 HBeAg 阳性检测率-
不完整结局指标数据亚组森林图

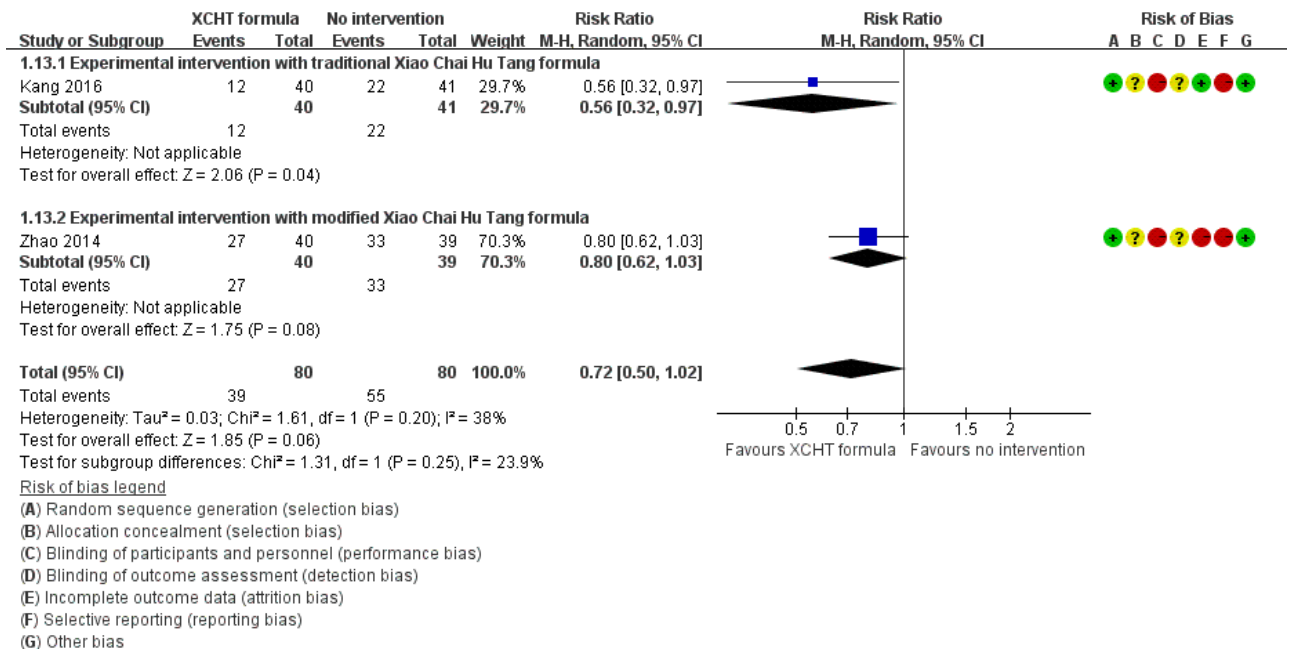


图 1-18 小柴胡汤治疗慢性乙型肝炎 HBeAg 阳性检测率-
不同小柴胡汤类型亚组森林图

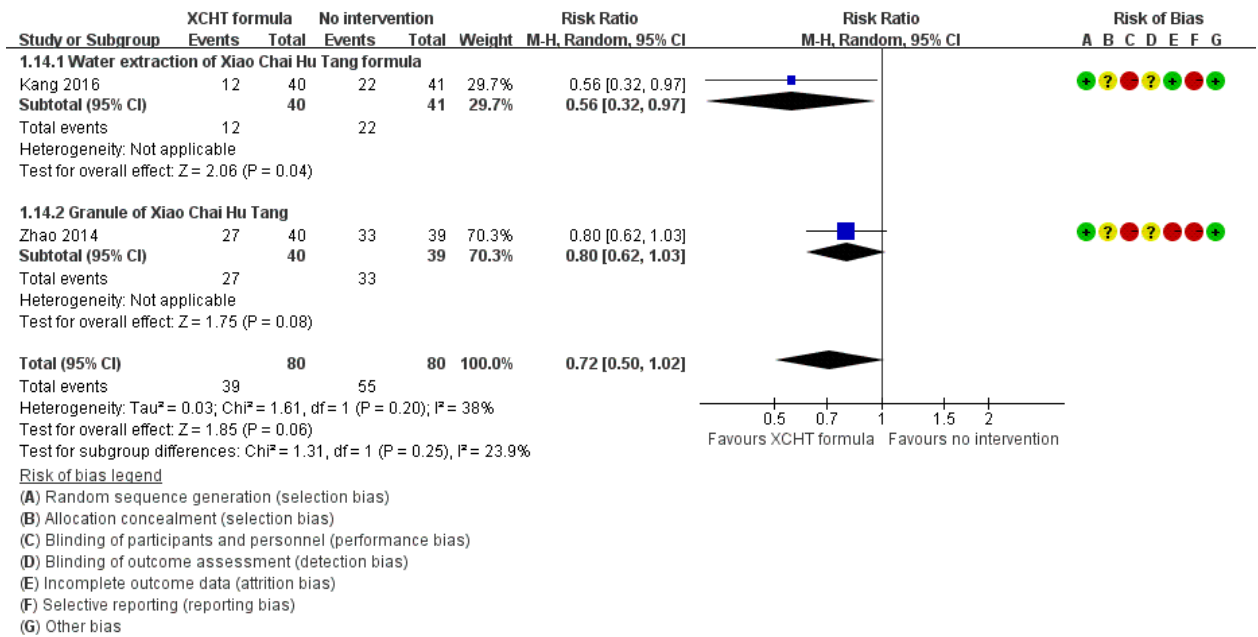


图 1-19 小柴胡汤治疗慢性乙型肝炎 HBeAg 阳性检测率-
不同小柴胡汤剂型亚组森林图

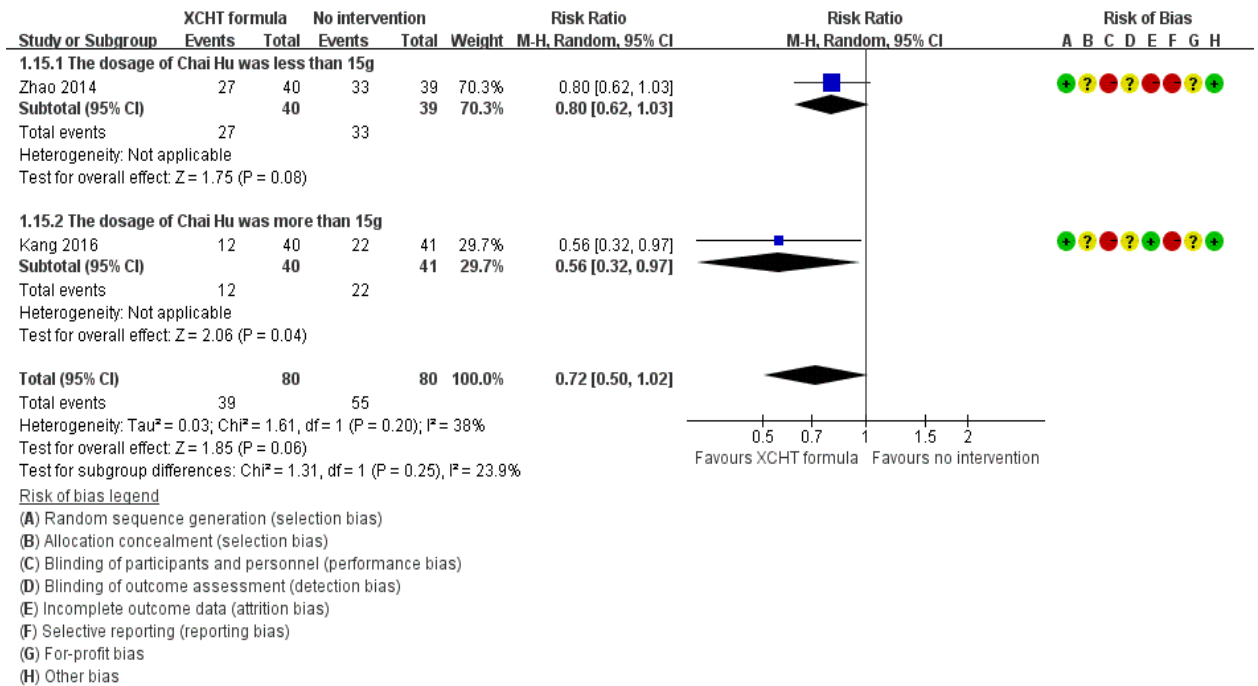


图 1-20 小柴胡汤治疗慢性乙型肝炎 HBeAg 阳性的参与者比例-
不同小柴胡汤剂量亚组森林图

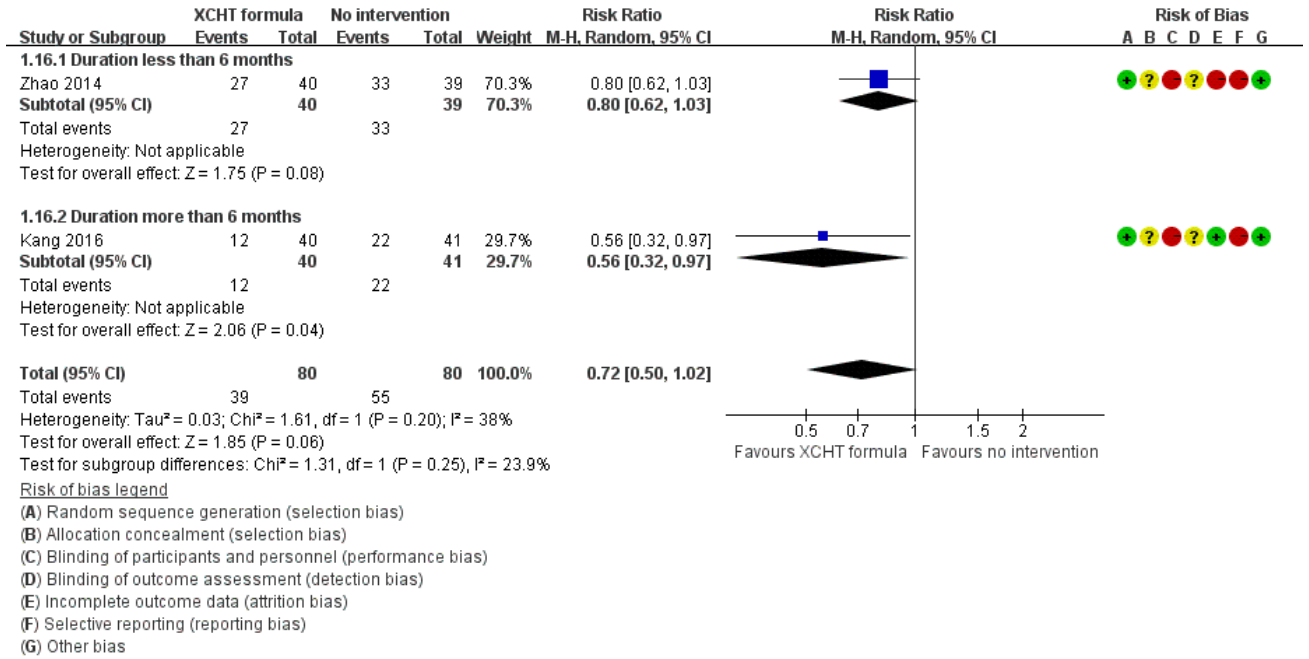


图 1-21 小柴胡汤治疗慢性乙型肝炎 HBeAg 阳性的参与者比例-不同小柴胡汤疗程亚组森林图

敏感性分析

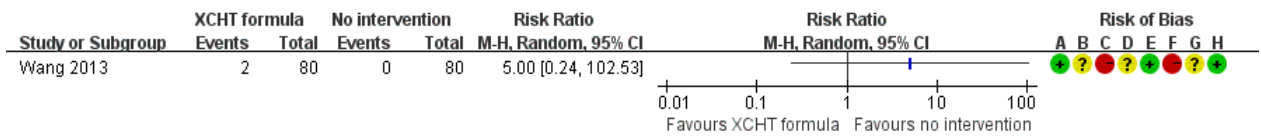
由于两项研究均报告了所有参与者关于 HBeAg 阳性比例的数据，我们没有进行敏感性分析（Zhao 2014；Kang 2016）。

运用 GRADE 分析和 TSA 分析评估关于该指标证据的不精确性，结果并不一致：运用 GRADE 分析对证据质量不精确性评估时，由于最优样本量未达到，样本量小于 4000，我们将证据质量降一级^[131, 132]；运用 TSA 分析对证据质量不精确性评估时，由于统计差异边界（有益、有害及无效）均未被穿过，对证据质量降两级^[98]。

3.3.3 分别报道的非严重不良事件

根据研究方案，对于只纳入一项研究并且数据为二分类变量的分析，我们运用 Fisher 精确检验。由于 Fisher 精确检验与 RevMan 检验结果一致，我们运用 RevMan 检验结果展示相关结局指标的结果。

不良事件恶心的发生率：只有一项研究包含 160 例参与者报告了发生恶心事件的参与者比例（Wang 2013）。我们不确定小柴胡汤是否对不良事件恶心的发生率有效应（RR=5.00；95%CI[0.24, 102.53]；图 1-22）。

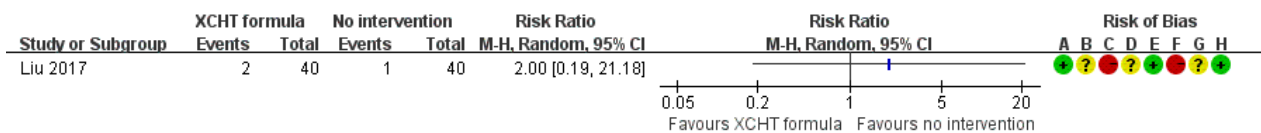


Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) For-profit bias
- (H) Other bias

图 1-22 小柴胡汤治疗慢性乙型肝炎不良事件恶心的发生率森林图

不良事件恶心及呕吐的发生率：只有一项研究包含 80 例参与者报告了不良事件恶心及呕吐的发生率（Liu 2017）。我们不确定小柴胡汤是否对不良事件恶心及呕吐的发生率有效应（RR=2.00, 95%CI [0.19, 21.18]；图 1-23）。

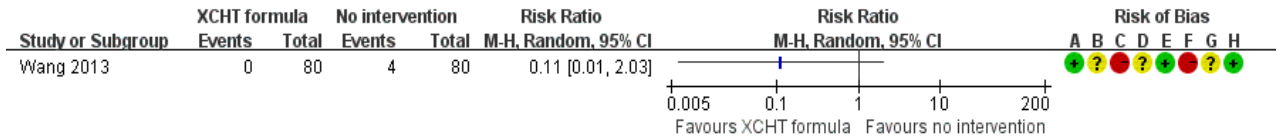


Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) For-profit bias
- (H) Other bias

图 1-23 小柴胡汤治疗慢性乙型肝炎不良事件恶心及呕吐的发生率森林图

不良事件头晕及失眠的发生率：只有一项研究包含 160 例参与者报告了不良事件头晕及失眠的发生率（Wang 2013）。我们不确定小柴胡汤是否对不良事件头晕及失眠的发生率有效应（RR=0.11, 95%CI [0.01, 2.03]；图 1-24）。

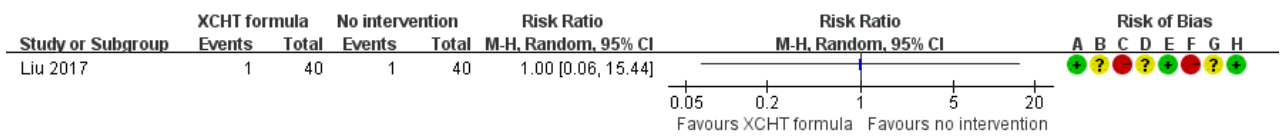


Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) For-profit bias
- (H) Other bias

图 1-24 小柴胡汤治疗慢性乙型肝炎不良事件头晕及失眠的发生率森林图

不良事件头晕及乏力的发生率：只有一项研究包含 80 例参与者报告了不良事件头晕及乏力的发生率（Liu 2017）。我们不确定小柴胡汤是否对不良事件头晕及乏力的发生率有效应（RR=1.00，95%CI [0.06, 15.44]；图 1-25）。

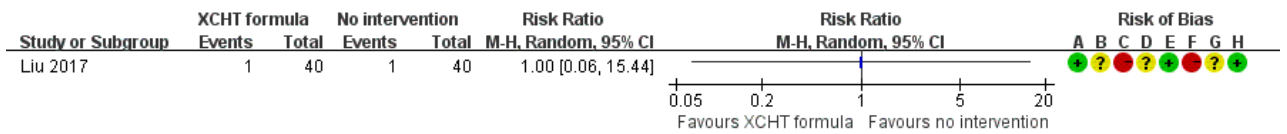


Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) For-profit bias
- (H) Other bias

图 1-25 小柴胡汤治疗慢性乙型肝炎不良事件头晕及乏力的发生率森林图

不良事件口干口苦的发生率：只有一项研究包含 80 例参与者报告了不良事件口干口苦的发生率（Liu 2017）。我们不确定小柴胡汤是否对不良事件口干口苦的发生率有效应（RR=1.00，95%CI [0.06, 15.44]；图 1-26）。

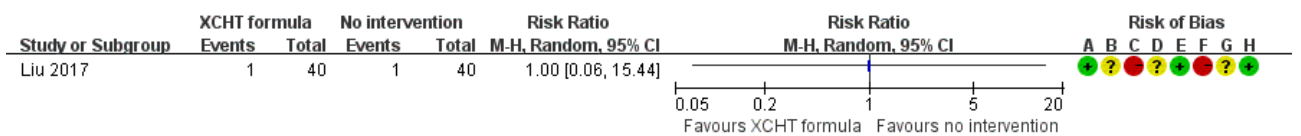


Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) For-profit bias
- (H) Other bias

图 1-26 小柴胡汤治疗慢性乙型肝炎不良事件口干口苦的发生率比例森林图

不良事件腹胀暖气的发生率：只有一项研究包含 80 例参与者报告了不良事件腹胀暖气的发生率（Liu 2017）。我们不确定小柴胡汤是否对不良事件腹胀暖气的发生率有效应（RR=1.00，95%CI [0.06, 15.44]；图 1-27）。

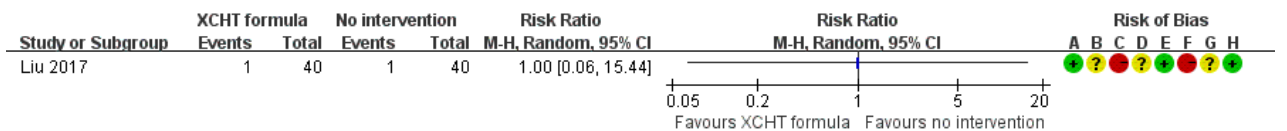


Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) For-profit bias
- (H) Other bias

图 1-27 小柴胡汤治疗慢性乙型肝炎不良事件腹胀暖气的发生率森林图

不良事件食欲不振的发生率：只有一项研究包含 80 例参与者报告了不良事件食欲不振的发生率（Liu 2017）。我们不确定小柴胡汤是否对不良事件食欲不振的发生率有效应（RR=1.00，95%CI [0.06, 15.44]；图 1-28）。



Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) For-profit bias
- (H) Other bias

图 1-28 小柴胡汤治疗慢性乙型肝炎不良事件食欲不振的发生率森林图

4. GRADE 分级

我们构建了结果总结表（SOF 表, Summary of finding table）以展示小柴胡汤联合常规疗法对比单用常规疗法的结局指标结果。然而，由于缺少试验数据，我们只能展示对于一项或多项非严重不良事件发生率，阳性 HBV-DNA 参与者比例以及阳性 HBeAg 参与者比例相关证据的 GRADE 评估结果。我们针对偏倚风险、无法解释的异质性或者结果的不一致性、不精确性以及发表偏倚三个因素对证据等级进行降级，最终认为所获得的证据确定程度极低（表 1-2）。

表 1-2 小柴胡汤治疗慢性乙型肝炎结果总结表(SOF 表)

结果总结:

小柴胡汤联合常规疗法对比单用常规疗法干预慢性乙型肝炎人群

患者或人群: 慢性乙型肝炎
 环境: 门诊或住院患者
 干预: 小柴胡汤联合常规治疗
 对照: 常规治疗

结局	预期绝对效应量* (95% CI)		相对效果(95% CI)	参与者数量 (研究数量)	证据的确定性 (GRADE)	评论
	单用常规治疗 的风险	小柴胡汤联合 常规治疗的 风险				
全因死亡率 - 未报告	-	-	-	-	-	
发生一项或多项严重不良事件参与者比例 - 未报告	-	-	-	-	-	
健康相关生存质量 - 未报告	-	-	-	-	-	
乙肝相关死亡率 - 未报告	-	-	-	-	-	
乙肝相关并发症 - 未报告	-	-	-	-	-	
发生一项或多项非严重不良事件患者比例 随访:范围 3 个月到 4 个月	58 每 1,000	25 每 1,000 (1 到 699)	相对危险度 0.43 (0.02 到 11.98)	240 (2 随机对照试验 (RCT))	⊕ VERY LOW ^{a,b,c}	c. 事件发生数 < 300, 置信区间包含重大获益值 (RR < 0.75) 及重大风险值 (RR > 1.25)
HBV-DNA 阳性检测率 随访:范围 4 月到 12 月	471 每 1,000	292 每 1,000 (212 到 400)	相对危险度 0.62 (0.45 到 0.85)	222 (3 随机对照试验 (RCT))	⊕ VERY LOW ^{c,d}	c.最优样本量未达到, 样本量不够大 (小于 4000 参与者)
HBeAg 阳性检测率 随访:范围 8 个月到 12 个月	688 每 1,000	495 每 1,000 (344 到 701)	相对危险度 0.72 (0.50 到 1.02)	160 (2 随机对照试验 (RCT))	⊕ VERY LOW ^{c,d}	c.最优样本量未达到, 样本量不够大 (小于 4000 参与者)

*干预组风险 (及其 95%可信区间) 以对照组假设风险以及干预的相对效果 (及其 95%可信区间) 为基础。
 CI: Confidence interval; RR: Risk ratio

GRADE 工作组证据等级标准:

高确定性: 我们非常确定真实效应与估计效应非常接近

中等确定性: 我们对估计效应中等程度确定, 即真实效应很可能与估计效应非常接近, 但是可能存在大量差异

低确定性: 我们对于估计效应的确定程度有限: 真实效应可能与估计效应存在较大差异

非常低确定性: 我们对于估计效应非常不确定: 真实效应非常可能与估计效应存在较大差异

表注:

- a. 由于存在分配隐藏、施盲、选择结局指标报告偏倚而降一级
- b. 由于不能解释的异质性或结果的不精确性: 相反的效应结果, 异质性 P 值 < 0.1, I² > 50%而降一级
- c. 由于发表偏倚降一级: 所有纳入的研究都是小样本量研究, 强烈怀疑存在发表偏倚
- d. 由于事件发生数过低导致的不精确性而降级, 对于一项或多项非严重不良事件发生率这一结局指标降两级, 对于探索性指标降一级

5. 后验性指标 meta 分析

临床上常常通过血清 ALT 及 AST 水平的变化来判断患者肝功能情况, 以进一步判断慢性乙型肝炎的轻重程度及患者的预后。参考临床医生的意见, 我们后验性的对于血清 ALT 及 AST 水平这两个指标进行后验性的 meta 分析, 具体如下:

参与者血清 ALT 水平:七项研究包含 633 例参与者报告了参与者血清 ALT 水平(Chen 2014, Chen 2017, Sun 2004, Wang 2013, Wang 2016, Wu 2015, Zhao 2014)。小柴胡汤可降低参与者血清 ALT 水平 (MD=-18.94, 95%CI[-29.42, -8.45]; 图 1-29)。纳入研究的异质性较大 ($I^2=97%$)，通过分析研究的临床特点发现, Chen 2017, Wang 2016, Zhao 2014 运用小柴胡汤加减方, Chen 2014, Sun 2004, Wang 2013 运用小柴胡汤原方, 且 Wang 2013 纳入患者的诊断标准并非国际较公认的指南, 可能是异质性的来源。

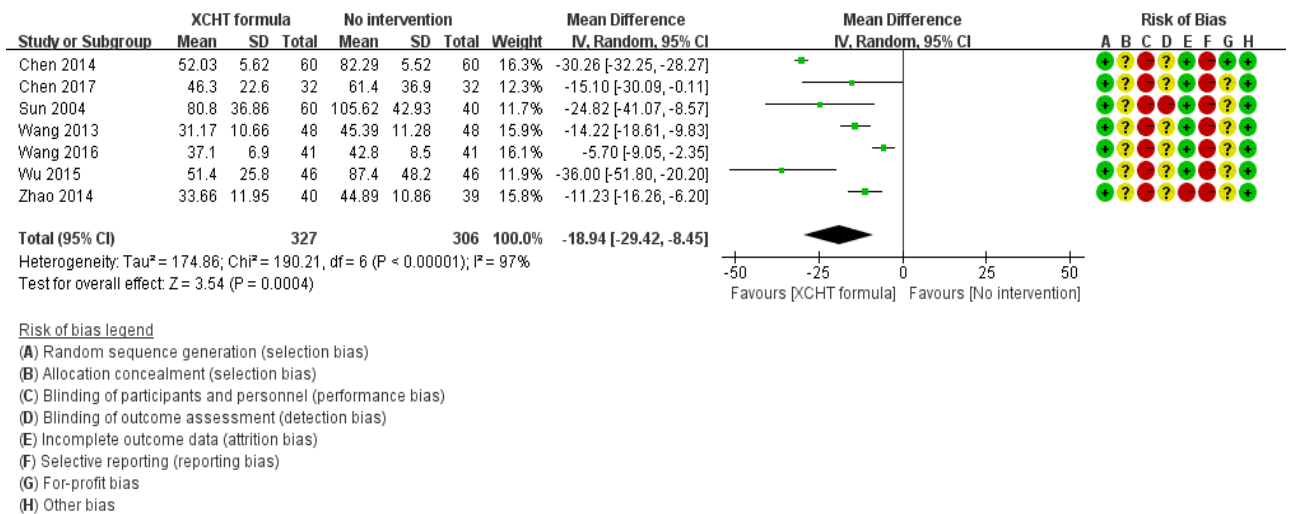


图 1-29 小柴胡汤治疗慢性乙型肝炎参与者血清 ALT 水平森林图

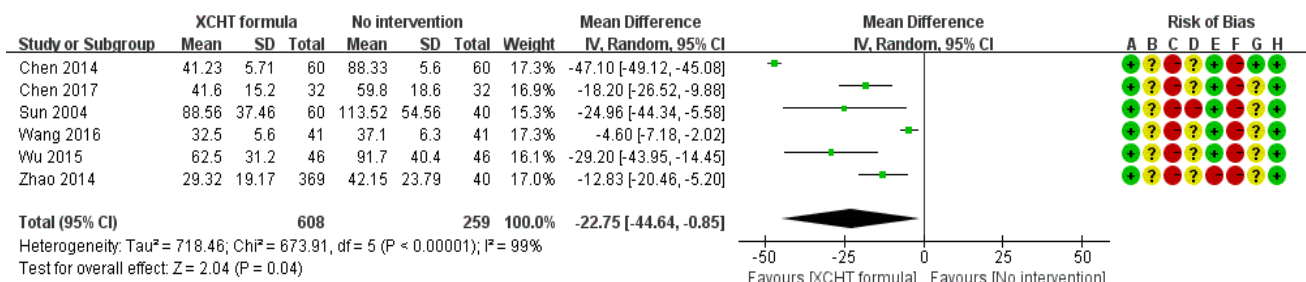
试验序贯分析

对于该结局指标, 通过试验序贯分析获得的 DARIS 为 8256 名参与者。我们运用如下参数进行分析: 对照组 SD 为 36.09, 最低标准差差异为 (MIREDF) -18.05, 对照组方差为 1302.5, α 值为 5%, 把握度为 90%, 多样性为 97.95%, 结果显示 Z 曲线未到达 DARIS (8256 名参与者), 也未穿过统计差异边界 (有益、有害及无效)。试验序贯分析显示无充分证据支持或者拒绝小柴胡汤联合常规疗法对比单用常规疗法可降低参与者血清 ALT 水平 (TSA 校正的 MD=-18.94; 95% CI [-61.74, 23.86]; 图 1-30)。



图 1-30 小柴胡汤治疗慢性乙型肝炎参与者血清 ALT 水平试验序贯分析图

参与者血清 AST 水平:六项研究包含 633 例参与者报告了参与者血清 ALT 水平(Chen 2014, Chen 2017, Sun 2004, Wang 2016, Wu 2015, Zhao 2014)。小柴胡汤可降低参与者血清 AST 水平 (MD=-22.75, 95%CI[-44.64, -0.85]; 图 1-31)。纳入研究的异质性较大 (I²=99%), 通过分析研究的临床特点发现, Chen 2017, Wang 2016, Zhao 2014 运用小柴胡汤加减方, Chen 2014, Sun 2004, Wang 2013 运用小柴胡汤原方, 且 Wang 2013 纳入患者的诊断标准并非国际较公认的指南, 可能是异质性的来源。



Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) For-profit bias
- (H) Other bias

图 1-31 小柴胡汤治疗慢性乙型肝炎参与者血清 AST 水平森林图

对于该结局指标，通过试验序贯分析获得的多样性校正后的所需样本量（DARIS）为 34,696 名参与者。我们运用如下参数进行分析：对照组 SD 为 40.43，最低标准差差异为（MIREDF）-20.21，对照组方差为 1634.6，α 值为 5%，把握度为 90%，多样性为 99.51%，由于只有 1.55%（867/110984）名患者提供信息，统计差异检测边界未能获得。试验序贯分析显示无充分证据支持或者拒绝小柴胡汤联合常规疗法对比单用常规疗法可降低参与者血清 AST 水平（TSA 校正的 MD=-28；95% CI [-44.76, -2.27]）。

讨 论

1. 效应分析

本研究纳入了 10 项随机对照试验共包含 926 例参与者。10 项研究均被评估为高偏倚风险研究。最终只能从 7 项共包含 670 例参与者的研究中提取数据进行定量分析。纳入研究均未评估本研究的主要结局指标及两项次要结局指标（乙肝相关并发症及乙肝相关死亡率）。所有纳入研究均比较小柴胡汤联合常规治疗与联合治疗干预慢性乙型肝炎的疗效。

对于发生一项或多项不良事件参与者比例这一结局指标，基于两项低质量随机对照试验的证据未显示小柴胡汤联合常规疗法对比单用常规疗法具有显著差异性效应。亚组分析显示一致结果。试验序贯分析表明需要更多信息以得到进一步结论。敏感性分析表明运用 GRADE 分级及试验序贯分析评估证据的不精确性，结果一致。

对于 HBV-DNA 阳性检测率这一替代结局指标，我们不确定小柴胡汤联合常规疗法对比单用常规疗法是否具有效应。试验序贯分析表明无充分证据肯定或者否定小柴胡汤对于这一结局有效应。亚组分析显示一致结果。

对于 HBeAg 阳性检测率这一替代结局指标，我们不确定小柴胡汤是否具有效应。试验序贯分析表明无充分证据支持或者拒绝小柴胡汤对于这一结局有效应。亚组分析显示一致结果。

对于分开报告的不良事件，即发生恶心、恶心及呕吐、头晕及乏力、头晕及失眠、口干口苦、腹胀暖气、食欲不振等事件结局，我们不确定小柴胡汤联合常规疗法对比单用常规疗法是否有效应。

对于参与者血清 ALT、AST 水平，我们不确定小柴胡汤联合常规疗法对比单用常规疗法是否有效应。

2. 证据的整体完整性和适用性

我们没有发现未发表或正在进行的试验。我们一共联系了 57 项研究的作者（包括 47 项由于不能提供必要的方法学信息而最终排除本研究的临床试验）以获得有关试验设计和方法的更多信息，只与一位作者取得联系并获得了相关信息。其中 43 项研究未能描述随机序列产生方法，4 项研究报告的信息不足以判断其与本研究的相关性。这 47 项研究未纳入我们的评估分析。证据体缺乏以患者为中心的结局指标的信息。

10 项试验纳入的参与者能够反映一般慢性乙型肝炎人群特征。所有试验均未纳入儿童。纳入试验包括小柴胡汤颗粒剂和水提取物、改良和经典小柴胡汤组成常用小柴胡汤剂量、小柴胡汤不同疗程的研究。对于以患者为中心的结局指标，我们只能对非严重不良事件进行系统综述，并只能获得两项研究数据，对于其他以患者为中心的结局指标如全因死亡率、健康相关的生活质量、乙型肝炎相关死亡率和乙型肝炎相关的发病率，我们未能获得相关报告。相反，我们发现了替代结局指标数据，例如阳性 HBV-DNA 参与者比例和阳性 HBeAg 参与者比例，参与者血清 ALT、AST 水平。我们还进行了亚组分析和敏感分析，试图根据相关临床因素探索治疗效果的差异。

3. 证据质量

3.1 研究间偏倚风险

纳入的 10 项研究中，10 项（100%）报告了充分的随机序列生成方法，0 项（0%）报告了随机分配的隐藏，10 项（100%）未予参与者施盲，因此被评价为结局指标高偏倚风险，8 项（80%）表明不受不完整结局数据影响，10 项（100%）均具有选择性报告偏倚偏倚风险，2 项（20%）表明不受营利偏倚影响，10 项（100%）表明无其他偏倚风险。所有研究均被评价为高风险偏倚。

3.2 证据的间接性

本研究中所有的纳入研究均评价小柴胡汤治疗慢性乙型肝炎患者的疗效。我们发现研究参与者与临床人群十分接近。因此，我们未发现证据间接性方面存在严重问题。

3.3 结果的异质性和不一致性

对于发生一项或多项非严重不良事件结局，研究间具有较大异质性 ($I^2 = 69\%$)。只有两项研究报告这一结局 (Wang 2013; Liu 2017)。通过分析研究特点寻找异质性原因发现，Wang 2013 纳入参与者患有其他合并疾病，而 Liu 2017 并未报告研究纳入的参与者是否合并其他疾病，这可能是导致异质性的原因。

对于阳性 HBeAg 参与者比例这一结局，研究间具有中等程度的异质性 ($I^2 = 38\%$)。我们分析了报告这一结局的两项研究 (Zhao 2014; Kang 2016) 的临床特点，发现 Zhao 2014 中小柴胡汤柴胡剂量为 6g，而 Kang 2016 中小柴胡汤柴胡剂量为 25g，这可能是导致异质性的原因。

对于阳性 HBV-DNA 参与者比例这一结局，无显著研究间异质性 ($I^2 = 0\%$)。

对于参与者血清 ALT、AST 水平结局，纳入研究的异质性较大 ($I^2 > 95\%$)，通过

分析研究的临床特点发现, Chen 2017, Wang 2016, Zhao 2014 运用小柴胡汤加减方, Chen 2014, Sun 2004, Wang 2013 运用小柴胡汤原方, 且 Wang 2013 纳入患者的诊断标准并非国际较公认的指南, 可能是导致异质性的原因。

3.4 研究结果的不精确性

根据 GRADE 标准, 对于一项或多项不良事件发生率这一结局, 由于事件发生数小于 300, 并且可信区间 CI 包含了无效线以及显著疗效($RR < 0.75$)或显著伤害($RR > 1.25$), 对证据质量降两级^[131, 132]; 对于阳性 HBV-DNA 参与者比例结局及阳性 HBeAg 参与者比例结局, 由于最优样本量未达到, 样本量小于 4000, 我们将证据质量均降一级^[131, 132]。

运用 TSA 分析对证据质量不精确性评估时, 由于统计差异边界(有益、有害及无效)均未被穿过, 对发生一项或多项不良事件患者比例、阳性 HBV-DNA 参与者比例以及阳性 HBeAg 参与者比例结局, 我们将证据质量降两级^[98]。

3.5 发表偏倚风险

对于本研究探索的结局指标, 均未能纳入超过 3 项以上的研究, 因此我们无法构建倒漏斗图。由于纳入研究均为小样本研究, 并且大多数研究表明小柴胡汤具有特异的阳性效应, 我们怀疑纳入的研究具有发表偏倚^[113, 114]。

4. 本研究中可能存在的偏倚

未发表的具有阴性结果的文章可能是潜在的偏倚来源。我们试图联系纳入本研究的 10 项研究及潜在可能为随机对照试验的 47 项研究的作者以获得方法学或者结局指标相关缺失数据, 但只与一篇研究作者成功取得联系。严重不良事件报告的缺失可能会导致对以患者为中心的结局指标结论的缺失。

对于非严重不良事件, 我们选择试验组或者对照组出现的最高事件数量以计算发生一个或多个非严重不良事件参与者比例。这一解决措施可能导致低估了发生非严重不良事件的参与者比例而存在问题。

我们进行了 7 项亚组分析并纳入了大量的替代结局指标。因此每个结局指标的 7 项亚组分析可能存在重复分析问题^[133]。

由于本研究进行的序贯分析中, 所有分析的所需样本量以及统计差异边界均未被穿过, 因此不能排除本研究涉及结局指标结果的随机误差风险。

小 结

1. 以患者为中心的结局指标如全因死亡率, 严重不良事件, 健康相关生存质量, 乙肝相

关死亡率以及乙肝相关并发症的相关数据并未被报告及评估。通过评价低质量的证据表明，我们发现小柴胡汤联合常规疗法对比单用常规疗法可降低 HBV-DNA 阳性检测率，降低参与者血清 AST、ALT 水平，但对于非严重不良事件和 HBeAg 阳性检测率改变这两个结局指标，目前尚不能表明小柴胡汤联合常规疗法对比单用常规疗法具有显著有效性。由于检索到的研究具有高风险偏倚，HBV-DNA 阳性检测率是未经证实的替代结局指标，纳入的研究未达到所需样本量并且纳入的研究其确定性较低，因此结论尚不能确证。

2. 纳入的小柴胡汤干预慢性乙型肝炎的研究均具有高风险偏倚，研究质量均很低。在今后的临床随机对照试验中，对于随机序列生成方法、随机分配的隐藏以及盲法应进行严格设计、实施和详尽报告，应尽量避免结局指标报告不完整或者存在缺失数据，未来研究报告的透明度和报告水平亟需提高。此外，还应针对不同的小柴胡汤方剂组成、不同的剂量以及不同的疗程进行试验设计；更多设计对比小柴胡汤与安慰剂干预慢性乙型肝炎的临床试验；应该对于研究人群的诊断标准以及合并疾病进行更严格清晰的限定，并且应针对不同的合并疾病进行分层随机；应设计更长随访期和更大样本量的临床随机对照试验；应设计更多针对以患者为中心的结局指标（如死亡率、严重不良事件、健康相关生存质量等）的临床试验，并应设计临床随机对照试验以验证替代结局指标与以患者为中心的结局指标的关系。通过提高相关中医临床研究质量，以进一步获得关于小柴胡汤干预慢性乙型肝炎的高质量临床证据，从而指导小柴胡汤的临床应用实施。

论文二 针灸疗法治疗慢性乙型肝炎的循证评价

针灸这一具有完整体系的干预措施，首次记载于《黄帝内经》。在古代，针灸被用于治疗具有食欲不振、恶心、轻度右侧季肋部不适症状的患者。现今，针灸，包括手工针灸、电针、激光针灸、穴位按压、穴位注射、艾灸以及穴位贴敷治疗，被广泛应用于中国、韩国、日本^[134, 135]、美国^[136]、英国^[137]以及澳大利亚。针灸治疗慢性乙型肝炎已被纳入中医药医疗体系^[134]。西方针灸应用主要涉及扳机点、分段点、公式点^[138, 139]。

中医理论认为健康是体内能量保持平衡的状态^[140, 141]。当器官或生理系统功能亢进或不足时，体内能量循环会发生异常。传统针灸是基于通过特定穴位刺激能量通路以使体内气机运行平衡，调节血液循环，以影响生理系统功能^[140, 142, 143]。刺激可由针刺、热量、药物、激光以及电流引起^[140, 142, 143]。不同疾病的针灸处方在穴位和操作手法方面不同^[140, 142]。慢性乙型肝炎常用穴位为足三里、三阴交、太冲、肝腧、内关、阳陵泉、期门、脾腧及肾腧^[144-148]。临床试验表明针灸可通过增强体内白细胞水平^[149, 150]、NK 细胞水平和活性^[151-153]、以及提高红细胞免疫功能而提高机体免疫功能^[154]。亦有临床试验表明针灸可通过提高血清中 CD3 水平、CD4 水平、CD4+/CD8+ 比例^[155, 156]以及降低 TNF- α 以及 IL-1 β 水平以调控机体免疫因子水平^[157]。

有报告指出应用针灸与 HBV 病毒感染^[158-161]及 HCV 病毒感染有关^[162]。这些感染可由反复应用消毒不彻底的针灸针、缺少本地监管机构检测、以及缺少对于血液传播疾病的清晰理解导致^[161, 162]。现今，在针灸实施过程中必须使用彻底消毒的针灸针或者一次性针灸针^[163]。其他不良事件如气胸、中枢神经系统损害、感染、困倦易睡（严重到可导致车祸）、以及局部银质沉着病已被报告可能由针灸引起^[164, 165]。尽管，在标准实践中由于针灸引起的严重不良事件风险并不常见（据估计发生率为每 10000 人 0.55 例）^[164, 165]。针灸师的训练和进一步非物理侵入性针灸方法的研发显得尤为重要。

目前，尚无经过严格方法学设计的循证医学证据对针灸疗法干预慢性乙型肝炎的获益和风险进行评估。本研究通过严格方法学设计，对针灸疗法（手工针灸、电针、激光针灸、穴位按压、穴位注射治疗、艾灸、穴位贴敷治疗）干预慢性乙型肝炎的获益和风险进行循证评价。

资料与方法

1 纳排标准

1.1 研究类型

对于针灸疗法干预慢性乙型肝炎的获益评估，纳入临床随机临床对照试验，对于试验的盲法、语言、发表年份、发表形式和是否发表无限制。对于针灸疗法干预慢性乙型肝炎的风险评估，在纳入临床随机临床对照试验的同时也纳入符合标准的半随机对照试验、对照临床研究以及其他观察性研究。

1.2 参与者

1.2.1 纳入标准

根据指南或者由研究者诊断为慢性乙型肝炎的患者：HBsAg 阳性超过 6 个月，血清 HBV-DNA 阳性超过 2000 IU / mL（即 10⁴ copies/ mL）；AST 或 ALT 持续或间歇性升高；肝脏活检结果显示乙型肝炎病毒感染伴有中度或重度坏死性炎症^[30,31]，对于患者的性别和年龄以及具体诊断指南和诊断标准无限制。参与者除患有慢性乙型肝炎外，亦可患有肝硬化、肝癌、人类免疫缺陷病毒感染（HIV）或获得性免疫缺陷综合症（AIDS）、丙肝、丁肝，或其他伴随疾病。

1.2.2 排除标准

无。

1.3 干预措施的类型

试验组

应用以下技术实施的针灸疗法，对针灸的施用风格（如中医式、日式、韩式、西式）以及实施方案（如穴位的数量和名称、针刺的深度、刺激停留时间、每日针刺次数）无限制：

手针：将针灸针直接刺入人体穴位中。以垂直于皮肤表面或与皮肤呈 15° -45° 角斜刺入皮肤中，伴或不伴有提插和捻转。留针时间通常为最后一根针刺入后 20-30 分钟。通常一日 1-2 次^[166, 167]。

电针：刺入的针被连入电针设备上（如 G6850 型电针仪）。该设备可通过小电极片产生持续电冲动以调节刺激的频率和强度。通常，刺激在持续波（如 2Hz）刺激下持续 20-30 分钟。

激光针灸：该技术运用低水平的激光以刺激能量及机体血液循环。持续的波长为

685nm 的激光可用作激光针。刺激强度可由 20mw-90mw 每针，刺激时间可通常为 10-40 分钟^[168, 169]。

穴位按压：穴位按压原理与针刺原理相似。通常运用手、肘、或其他设备对穴位实施物理按压，通常持续 3-20 分钟，每天 1-2 次^[170, 171]。

穴位注射：针灸的延展疗法。该疗法将药物注入穴位中。穴位注射通常注入抗病毒注射液（如聚乙二醇干扰素 α -2b 180 μ g 每周一次）^[172, 173]或中医药注射剂（如香丹注射液 4ml 每日一次）^[174, 175]。

艾灸：该技术燃烧干艾绒或将干艾绒或其他具有治疗作用的药材直接或间接的放置于身体的穴位上。对于间接灸，将大约 2 厘米长的艾柱与穴位皮肤上方点燃。以局部获得温热感为宜^[176, 177]。对于直接灸，将锥形艾柱直接于穴位皮肤上点燃^[178, 179]。

穴位贴敷：通过将草药制成贴剂，或者研磨后的草药粉末与水、醋、酒、蛋白、蜂蜜、及植物油或固体油（如凡士林）等混合，或者加入黄酒制成贴剂、膏剂、或小圆饼。将穴位贴敷直接覆盖于穴位上或阿是穴 3-24 小时，每日一次^[180-182]。

对照组：

应用假针灸或无任何干预措施。试验组和对照组可应用具有相同的联合干预措施。

1.4 结局指标的类型

(1) 主要结局指标：

①全因死亡率；

②一种或多种严重不良事件的发生率（严重不良事件即任何非预期的医疗事件，可导致死亡、危及生命、需要住院或延长住院、持续或重大残疾或无行为能力、先天性异常或先天性缺陷^[83, 84]。）；

③健康相关生存质量：研究者用于评估试验参与者生存质量报告的任何国际公认的量表。

(2) 次要结局指标：

①乙肝相关死亡率；

②乙肝相关并发症（发生一种或多种如下事件的参与者的比例：肝硬化，腹水，静脉曲张出血，肝肾综合征，肝细胞癌，肝性脑病，或肝移植，但并没有发生死亡者）；

③一种或多种非严重不良事件的发生率：非严重不良事件即不满足严重不良事件定义的不良事件^[83, 84]。

(3) 探索性结局指标：

- ①血清、血浆中 HBV-DNA 阳性或者 HBV-DNA 病毒载量阳性的参与者比例；
- ②血清、血浆中 HBeAg 阳性的参与者比例；
- ③分开报告的严重不良事件发生率；
- ④分开报告的非严重不良事件发生率；
- ⑤分开报告的乙肝相关并发症发生率。

我们评估最长随访期时的结局指标数据。

2. 检索策略

2.1 电子检索

我们检索英文数据库：Cochrane 肝胆病组对照试验注册库（The Cochrane Hepato-Biliary Group Controlled Trials Register (Glud 2018)）、Cochrane 图书馆中的 Cochrane 对照试验注册中心（Cochrane Central Register of Controlled Trials (CENTRAL) in The Cochrane Library）、MEDLINE Ovid、Embase Ovid、LILACS Bireme (Latin American and Caribbean Health Science Information database)、Science Citation Index Expanded (Web of Science)、Conference Proceedings Citation Index-Science (Web of Science) (Royle 2003)。我们还检索了中文数据库：中国知网数据库 (China National Knowledge Infrastructure Database, CNKI)、万方数据库、重庆维普中文科技期刊数据库 (Chongqing VIP Chinese Science and Technology Periodical Database, VIP)、中国生物医学文献数据库 (Chinese Biomedical Literature Database, CBM)。

附表 2 提供了不同数据库的检索策略及时间跨度。

2.2 其他检索

我们检索了相关系统综述、meta 分析以及符合纳入标准的文章的参考文献列表。我们还检索了世界卫生组织国际临床试验注册平台 World Health Organization International Clinical Trials Registry Platform (www.who.int/ictrp)、ClinicalTrial.gov (www.clinicaltrials.gov/) 以及中国临床试验注册中心 (ChiCTR) 以检索正在进行中的研究和未发表研究。

3. 数据收集与分析

根据 Cochrane 干预措施系统综述手册^[85]以及 Cochrane 肝胆病协作组模板 (Glud 2018) 以设计设施本系统综述研究。运用 Review Manager 5.3 (RevMan 2014) 软件及试

验序贯分析软件 Trial Sequential Analysis version 0.9.5.10 Beta Software^[88,89] 进行数据分析。

3.1 研究的选择

研究者两人为一组独立提取可能符合纳入标准的研究的题目和摘要。将同一研究的多个报告列入该研究参考文献下。如遇研究筛选不一致的情况可以通过讨论或由第三位研究者协助解决及确定。根据 PRISMA 流程图 (PRISMA 2009) 记录研究筛选过程。

3.2 数据提取与处理

研究者两人一组运用电子数据收集表格独立提取数据。如有分歧,则重新检查所提取的数据。如遇数据提取不一致的情况可以通过讨论或由第三位研究者协助解决及确定。提取内容包括:发表相关数据(即年份、国家、作者)、研究特征与设计、试验参与者特征、试验纳入和排除标准、干预措施、结局指标、随访、数据分析类型(即意向性治疗分析、校正意向性治疗分析、符合分析方案集分析)。提取最长随访时间节点数据。联系研究人员获得研究报告缺失信息。

3.3 纳入研究的偏倚风险评估

依据 Cochrane 偏倚风险评估工具^[85], Cochrane 肝胆病组模板 the Cochrane Hepato-Biliary Group Module^[86], 以及相关方法学研究^[91-94], 研究者两人一组独立评估纳入研究的偏倚风险, 主要包括以下方面: 随机分配产生、分配方案隐藏、对参与者及主要研究者的施盲、对研究结果评估者的施盲、结局指标数据的不完整性、选择性报告、营利偏倚、其他偏倚。对纳入的 RCT 进行质量评价, 针对每一项研究结果, 对以上方面做出“低偏倚”、“高偏倚”以及对于缺乏相关信息或偏倚情况不确定情况下的“不清楚”的判断, 评价者为两名或两名以上, 判断意见不一致时征求第三方意见。具体评估标准参见论文一中的偏倚风险评估标准。

3.4 效应测量

运用相对危险度 risk ratio (RR) 测量二分类变量结局指标, 运用均差 mean difference (MD) 测量连续变量, 两者均以 95%CI 表示。当运用不仪器测量同一连续变量时, 运用标准均差及 95%CI 测量。

3.5 分析单位

根据干预措施系统综述 Cochrane 手册^[85], 拟定数据分析单位为随机分配入研究的参与者个体。对于具有多个试验组的临床试验, 我们纳入符合本综述研究内容的组别。

如果不同的试验组共用同一对照组，我们将对照组分为两组以避免重复计算。对于整群随机对照试验，如果试验合理考虑到了组群设计，我们将从分析中直接提取数据。如果试验没有对组群进行控制，我们运用膨胀标准误及 Review Manager 5 软件中的倒方差法进行处理。对于交叉试验，我们只对第一阶段试验进行数据提取以避免残留治疗影响^[85]。

3.6 缺失数据处理

联系研究者以获得纳入的研究中报告不详细的数据。我们运用意向性治疗方法进行数据分析，当无法运用此方法时，我们使用可以运用的数据进行分析。对于所有的结局指标，我们纳入不完整或有缺失数据的参与者，并在敏感性分析中按如下方法进行数据填补：

(1) 二分类变量

①最好-最坏情况：假设试验组所有失访参与者全部存活、临床症状全部改善、全部没有严重不良事件，全部没有并发症发生；对照组中所有失访参与者全部死亡，临床症状全未改善，全部发生严重不良事件，全部发生并发症。

②最坏-最好情况：假设试验组所有失访参与者全部死亡，临床症状全未改善，全部发生严重不良事件，全部发生并发症；对照组所有失访参与者全部存活、临床症状全部改善、全部没有严重不良事件，全部没有并发症发生。

(2) 连续变量

对于获益性结局指标在组别均值基础上加 2 个标准差，而对于风险性结局指标在组别均值上减去 2 个标准差^[98]。如果试验未报道标准差值，则联系作者以获得相关信息或者运用试验所报告数据进行计算。

3.7 异质性评估

通过仔细分析试验参与者特征和纳入研究的设计以评估纳入研究的临床异质性和方法学异质性。如果可获得相关数据，则通过对比纳入研究中不同的参与者诊断标准、参与者的合并疾病、报告参与者是否得气、针灸疗法的方式、不同疗程、不同联合干预措施、不同对照组干预措施、以及不同随访期的疗效对比，对纳入研究的临床异质性进行分析。

不同的研究设计类型和偏倚风险可导致方法学异质性。我们通过比较固定效应模型和随机效应模型结果评估统计学异质性。首先我们通过观察森林图寻找统计学异质性征

象；其次，我们运用 Chi^2 检验观察 I^2 值及 P 值判断异质性程度^[86, 95, 96]。通常， I^2 在 0%–40% 表明异质性可能不重要；30%–60% 代表中等程度异质性；50%–90% 代表大量异质性；75%–100% 表明重大异质性^[85]。

在 TSA 分析中，我们运用 D^2 进行多样性校正后的所需样本量 (DARIS) 分析^[97]。

3.8 报告偏倚评估

如果每项比较包含至少 10 项研究，我们通过评估倒漏斗图的对称性评估报告偏倚。对于二分类变量，我们运用 Harbord 检验评估对称性；对于连续变量，我们运用不对称回归检验评估对称性。

3.9 数据合成

3.9.1 Meta-分析

我们依据干预措施系统综述的 Cochrane 手册以及 Cochrane 肝胆病组模板实施 meta 分析^[85]。我们运用 Review Manager 5 软件进行数据分析^[87]。我们既运用固定效应模型也运用随机效应模型评估我们的干预措施，当结果不一致时（例如其中一个模型得出显著性干预效应，而另一个模型得出无显著性干预效应），我们报告两个模型的分析结果，但是我们将更多的权重赋予更接近无干预效应的结果（即最高 P 值结果）^[98]。

为了使多重比较谬误低于 0.05，我们评估三个主要结局指标和三个次要结局指标时，以 P 值 ≤ 0.025 为显著统计学差异^[98]。对于探索性指标，我们认为 P 值 < 0.05 为显著统计学差异。我们通过评估纳入研究的统计学异质性和临床异质性来决定是否进行 Meta 分析或进行描述性分析。

我们在敏感性分析中进行缺失数据填补 (参见敏感性分析)^[98]。当仅可获得一项研究的数据时，我们运用 Fisher 检验检测二分类变量^[99]，运用 Student's t 检验检测连续变量^[100]。

3.9.2 试验序贯分析

由于小样本数据和重复测量，累积 meta 分析可能存在随机误差风险，因此我们运用试验序贯分析。为了减小随机误差，我们计算所需样本量 (RIS)（即接受或拒绝干预措施效果在 Meta 分析中所需要的参与者数量）^[89, 101, 102]。所需样本量计算应该同时考虑 meta 分析中的多样性^[101, 103, 104]。

我们对二分类变量和连续变量控制一类错误和二类错误^[101, 103-107]。对于二分类变量，我们依据对照组的事件发生比例估计根据多样性校正的所需样本量 (DARIS)，相关风

险降低率为 15%，主要结局指标和次要结局指标 α 值为 2.5%，探索性指标结局指标 α 值 5.0%， β 值为 10%^[108]，多样性为 meta 分析中纳入研究的异质性^[98, 103]。对于连续性变量，我们根据对照组观察到的 SD 估计 DARIS，最低相关差异为 50%SD 值，主要结局指标和次要结局指标 α 值为 2.5%，探索性指标结局指标 α 值 5.0%， β 值为 10%^[108]，多样性为 meta 分析中纳入研究的异质性^[98, 103]。

我们运用统计检测边界检测有效性和有害性的统计显著性差异，并运用无效性边界检测无效性^[109]。如果累积检验统计量（Z 曲线）在到达 DARIS 前穿过统计检测边界，则可认为干预措施的效应优于或者差于对照组干预措施。如果 Z 曲线在到达 DARIS 前穿过无效性检验，则意味着干预措施不具备假设效应，并且进一步的随机对照试验可能是无效的。

如果试验序贯分析的统计检测检测边界和无效边界都没有被穿过，则仍有必要进一步进行临床试验以接受或者拒绝干预措施效应^[101, 102]。

4. 亚组分析和异质性检测

如果能够提取相关数据，则如下进行亚组分析：

总体低偏倚风险的研究与总体高偏倚风险的研究；结局指标盲法被评估为低偏倚风险的研究与其被评估为高偏倚风险的研究；不完整结局指标数据被评估为低偏倚风险的研究与其被评估为高偏倚风险的研究；选择性结局指标报告被评估为低偏倚风险的研究与其被评估为高偏倚风险的研究；对照组运用不同干预措施的研究（即对照组无干预措施与对照组运用假针灸进行干预）；不同的针灸干预措施（即手针、电针、激光针灸、穴位按压、穴位注射、艾灸、穴位贴敷）的研究；干预措施不同疗程的研究（根据观测到的均值进行分层）；报告针灸反应（得气或肌肉抽动）的参与者的研究与未报告针灸反应的参与者的研究；具有参与者不同诊断标准的研究；只患有慢性乙型肝炎患者对比患有其他合并疾病患者（肝硬化，干细胞癌，HIV 感染，AIDS，丙肝，丁肝或多个合并疾病者）的研究（我们将单独分析每个合并疾病）。

5. 敏感性分析

除了缺失数据处理部分的敏感性分析，我们还比较 GRADE 分级中的不精确性评估与试验序贯分析中的不精确性评估^[98]。必要时可进行其他敏感性分析^[98, 110, 111]。

6. GRADE 分级

构建 SOF 表以展示主要结局指标和次要结局指标发现的证据结果。运用 GRADEpro

GDT 软件 (community.cochrane.org/help/tools-and-software/grade-pro-gdt) 评估证据质量的五个方面：研究间偏倚风险评估、证据的间接性（参与者，干预措施，对照，结局）、结局不能解释的异质性和非一致性（包含亚组分析的问题）、结局的不精确性、发表偏倚风险^[112, 113, 114]。

证据质量定义如下^[112, 113, 114]：高确定性（High certainty）：我们非常确定真实效应与估计效应非常接近；中等确定性（Moderate certainty）：我们对估计效应中等程度确定，即真实效应很可能与估计效应非常接近，但是可能存在大量差异；低确定性（Low certainty）：我们对于估计效应的确定程度有限：真实效应可能与估计效应存在较大差异。非常低确定性（Very low certainty）：我们对于估计效应非常不确定：真实效应非常可能与估计效应存在较大差异。

研究结果

1. 纳入研究描述

1.1 文献检索结果

通过电子检索出 1838 条记录（见图 2-1）。我们通过检索纳入文献的参考文献，我们检索出 4 条参考文献。在排除 1645 篇重复发表或明显不相关的参考文献，我们阅读了 197 篇发表研究的全文。110 项研究未符合纳入标准。在联系研究人员后，73 项研究没有提供明确的随机序列产生方法，并且 6 项研究由于缺少足够信息以判断是否是随机对照试验，这 79 项研究被排除我们的 meta 分析。我们检索到两项正在进行中的研究 (ChiCTR-IOR-17011957^[183]; ChiCTR-IOR-17011962^[184])，未检索到未发表的研究。

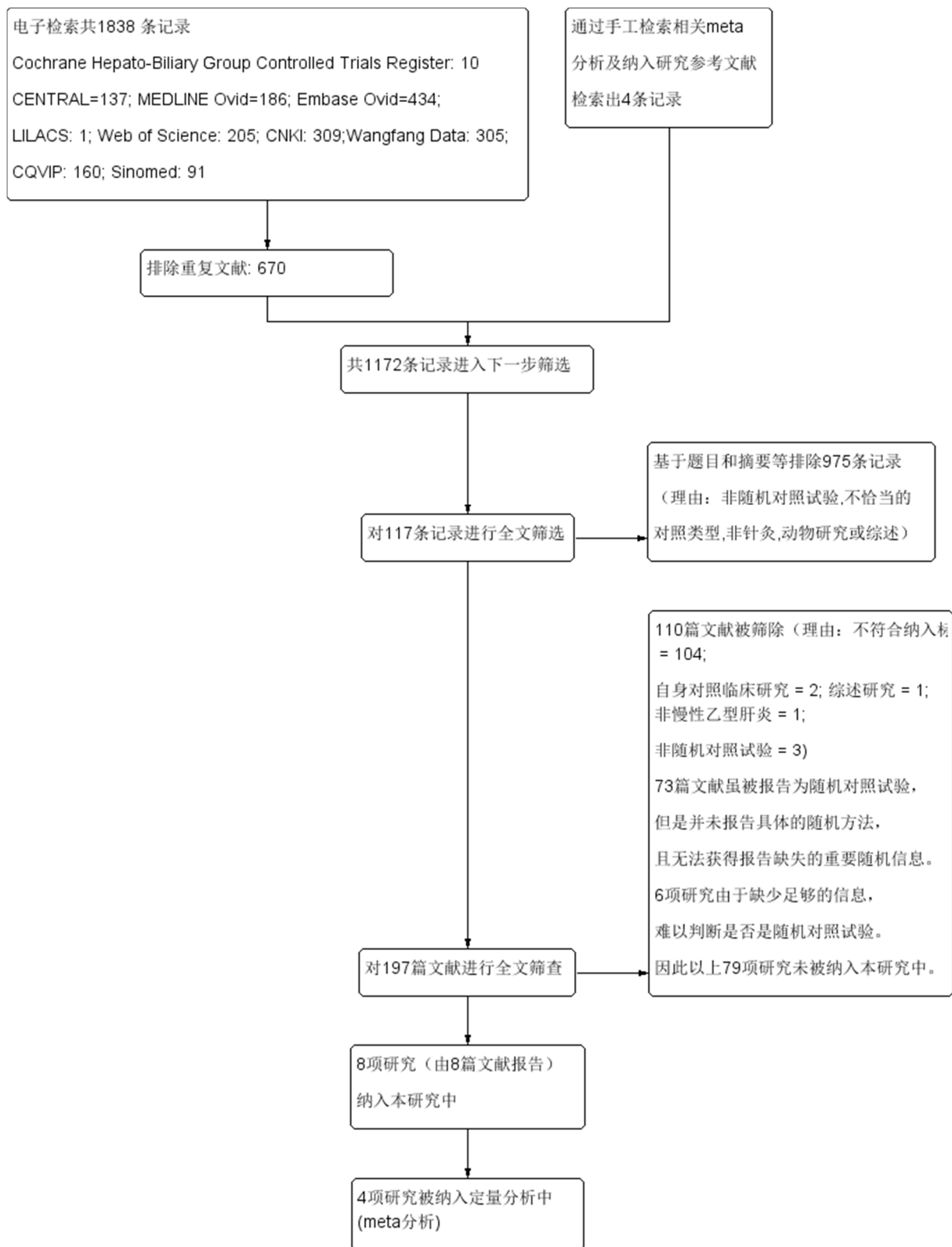


图 2-1 针灸治疗慢性乙型肝炎纳入研究流程图

1.2 纳入研究的一般特征

本研究共纳入 8 项临床研究共涉及 552 名参与者 (Wang 2002; Huang 2011; Liao 2011;

Zou 2011; Wang 2013; Zhang 2014; Jing 2016; Zhu 2016)。其中 4 项研究并未对本研究的目标结局指标进行报告，因此对这 4 项研究采用描述性分析。所有研究地点均为中国大陆，报告语言为普通话。纳入研究的样本含量 60-160 例，平均年龄为 26-43 岁，男女比例 263:131，1 项研究的部分受试者合并肺结核和腹水，肝郁脾虚证 1 项（Jing 2016），肝胆湿热证 1 项（Liao 2011）肝郁气滞证 1 项（Zhang 2014）。纳入研究均为针灸疗法联合常规疗法对照常规疗法。2 项研究手针：相同穴位：足三里，三阴交，太冲；留针时间 1h。

3 项研究

穴位贴敷：相同穴位：期门，日月；留贴时间 3-36h 不等。2 项研究穴位注射黄芪注射液；相同穴位：足三里；剂量 2-5ml 不等。

1 项研究艾灸；穴位：神阙；时长：30min。疗程为 14 天-12 个月不等。6 项研究随访期等同于干预期，2 项研究随访期持续疗程结束后 1-12 个月不等。3 项研究报告了非严重不良事件发生率，3 项研究报告了探索性结局指标，部分研究报告了血清肝纤维化指标、肝功能相关指标、复合指标（临床有效率）等。纳入研究特点具体参见表 2-1。

表 2-1 针灸治疗慢性乙型肝炎随机对照试验的一般特征

研究 ID	疗程	试验组 例数	对照组 例数	男：女	平均年龄 (岁)	诊断标准	试验组 干预措施	对照组 干预措施	结局指标
Huang 2011 ^[185]	6M	30	29	47:12	试验组 25 对照组 23	[190]	黄芪注射液穴位注射+ 常规治疗	常规治疗	乙肝病毒标志物水平：HbsAg, HbeAg, HbeAb, HbcAb, HBV-DNA；肝 功能指标变化（ALT, AST, TBIL）；血 清 TC 水平，复合指标；应答率
Jing 2016 ^[186]	12M	30	30	33:27	20-40	西医：[187] 中医：[188]	黄芪注射液穴位注射+复方 当归注射液+聚乙二醇干扰 素 α-2a 皮注	聚乙二醇干扰 素 α-2a 皮注	复合指标：应答率；生化指 标；CD4, CD8, CD4/CD8

Liao 2011 ^[189]	3W	30	30	48:12	试验组 36.5 对照组 35.4	[190]	茵陈蒿贴穴位贴敷+甘利欣 注射液静滴	甘利欣注射液 静滴	临床症状体征改善量化表, 肝功能指标变化 (ALT, AST, A/G, TBIL) 中医肝病 PRO 量表, 安全性指标
Wang 2002 ^[191]	6M	30	28	未报告	试验组 25.7 对照组 29.5	未报告	针刺+常规治疗	常规治疗	肝功能指标变化: ALT 等; 乙肝病毒标志物阴转率: HbsAg, HbeAg, HbeAb, HbcAb, HBV-DNA
Wang 2013 ^[192]	3M	40	40	58:22	40.3	西医:[193] 中医:[194]	脐部艾灸+常规治疗	常规治疗	生化指标: TBI1, ALT, AST, ALB, HA, LN, PCIII, and IV-C
Zhang 2014 ^[195]	3M	35	35	36:34	试验组 4.33 对照组 36.85	[196]	针刺+常规治疗	常规治疗	肝功能指标改善; 安全性指标
Zhu 2016 ^[197]	2W	33	32	41:24	试验组 39.3 对照组 42.3	[198-200]	穴位贴敷+常规治疗	常规治疗	临床体征; 彩超: 门静脉和脾静脉血流量、血流速, 腹水液性暗区深度
Zou 2011 ^[201]	3M	50	50	未报告	13-65	[202]	穴位贴敷+常规治疗	常规治疗	肝功能指标改善 (ALT, AST, TBI1), 乙肝病毒标志物阴转率; HBsAg, HBeAg, HBV-DNA; 安全性指标

2. 纳入研究的偏倚风险评估

我们通过联系作者及文献提供信息, 对纳入研究进行风险偏倚评估 (见图 2-2, 图 2-3)。

2.1 随机分配 (选择偏倚)

纳入研究均使用电脑或随机数字表产生随机序列, 只有 Liao 2011 报告在不透明信封中进行分配信息的隐藏, 因此被评为低风险, 而其余研究均存在选择偏倚风险。

2.2 盲法 (实施偏倚及检测偏倚)

纳入研究均未报告参与者和/或研究者施盲。我们评估所有纳入研究为实施偏倚高风险。所有研究均未报告结局指标评估施盲，因此被评为不确定偏倚风险。

2.3 不完整的结局指标数据（磨损偏倚）

6项研究报告无缺失数据，并在数据分析中纳入所有参与者(Liao 2011; Zou 2011; Wang 2013; Zhang 2014; Jing 2016; Zhu 2016)，因此被评为低风险偏倚。Huang 2011在数据分析时排除了1名参与者数据，Wang 2002在数据分析时排除了2名参与者数据，因此我们将其评估为高风险偏倚。

2.4 选择性报告（报告偏倚）

由于缺少提前发表的研究方案以及无死亡率、严重不良事件、健康相关生存质量相关数据报告，所有纳入的研究均被视为高风险偏倚。

2.5 其他偏倚

所有纳入研究不存在其他潜在偏倚风险，因此被评估为低风险偏倚。

2.6 总体偏倚风险

我们将纳入研究均评估为总体高风险偏倚。

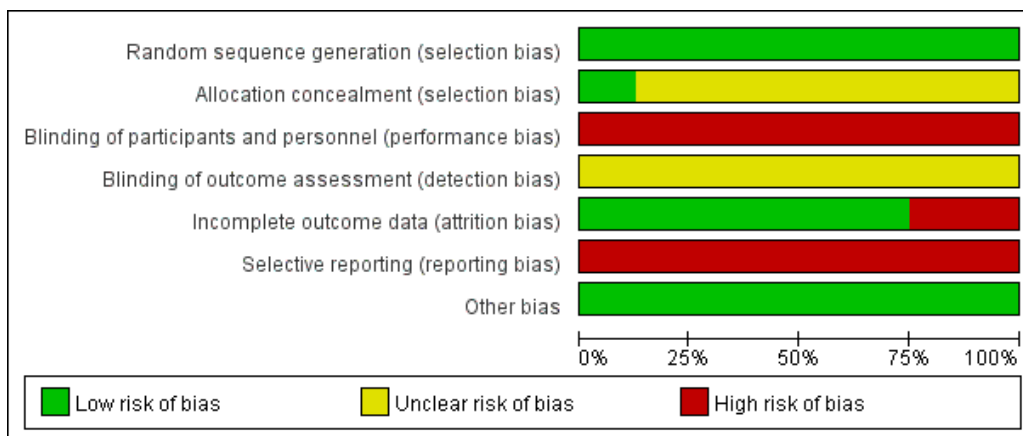


图 2-2 针灸治疗慢性乙型肝炎随机对照试验总体风险偏倚评估情况

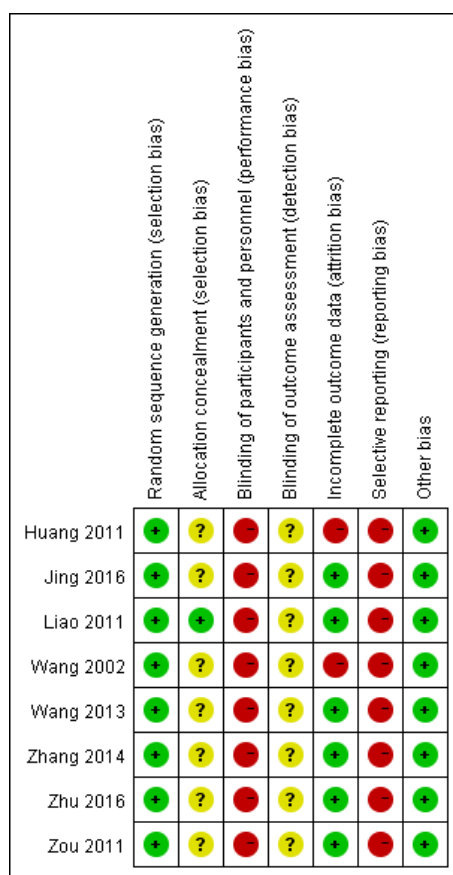


图 2-3 针灸治疗慢性乙型肝炎各随机对照试验偏倚风险评估详情

3. 干预措施效应评估

3.1 主要结局指标

纳入研究均未报告本系统综述主要结局指标：全因死亡率、一项或多项不良事件的发生率、健康相关生存质量。

3.2 次要结局指标

纳入研究均未报告本系统综述次要结局指标：乙肝相关死亡率、乙肝相关发病率。

3.2.1 一项或多项非严重不良事件发生率

三项研究涉及 203 例参与者报告该指标数据（Wang 2002；Wang 2013；Zhu 2016）。我们不确定针灸是否对一项或多项非严重不良事件发生率有效应（RR=0.67，95%CI[0.43, 1.06]；P=0.09；I²=0%；图 2-4）。

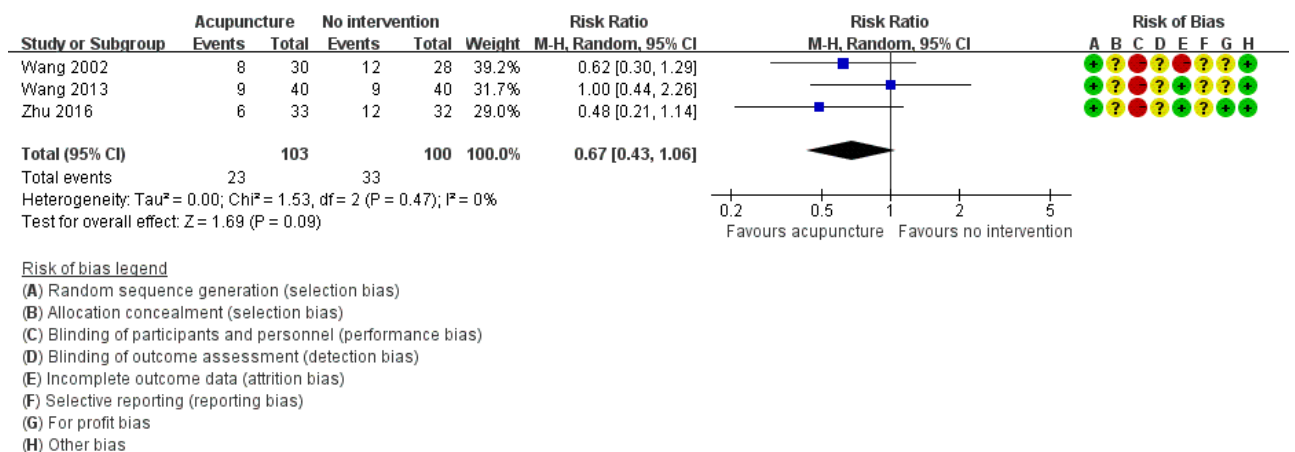


图 2-4 针灸治疗慢性乙型肝炎一项或多项非严重不良事件发生率森林图

试验序贯分析

对于该结局指标，通过试验序贯分析获得的多样性校正后的所需样本量（DARIS）为 4,297 名参与者。我们运用如下参数计算所需样本量：对照组事件发生比例 33%， α 值为 2.5%，把握度为 90%，多样性为 0%。由于只有 4.72%（203/4297）名患者提供信息，统计差异检测边界未能获得（Wang 2013；Liu 2017）。因此，对于该指标，试验序贯分析显示无充分证据支持或者拒绝针灸联合常规治疗对比常规治疗可降低 15% 风险。

亚组分析

由于无法获得全部预定结局指标数据，因此无法进行全部预定亚组分析。我们未发现亚组间显著统计学差异在如下亚组分析中：比较手针研究、穴位贴敷以及艾灸的研究（Chi²=1.53；P=0.47；I²=0%；图 2-5）；比较疗程大于 12 周和小于 12 周的研究（Chi²=0.19；P=0.66；I²=0%；图 2-6）；比较研究者诊断为慢性乙型肝炎的研究和根据指南诊断为慢性乙型肝炎的研究（Chi²=0.06；P=0.81；I²=0%；图 2-7）；比较纳入肝硬化患者与无明

确合并疾病信息的研究 ($\text{Chi}^2=0.80$; $P=0.37$; $I^2=0\%$; 图 2-8)。

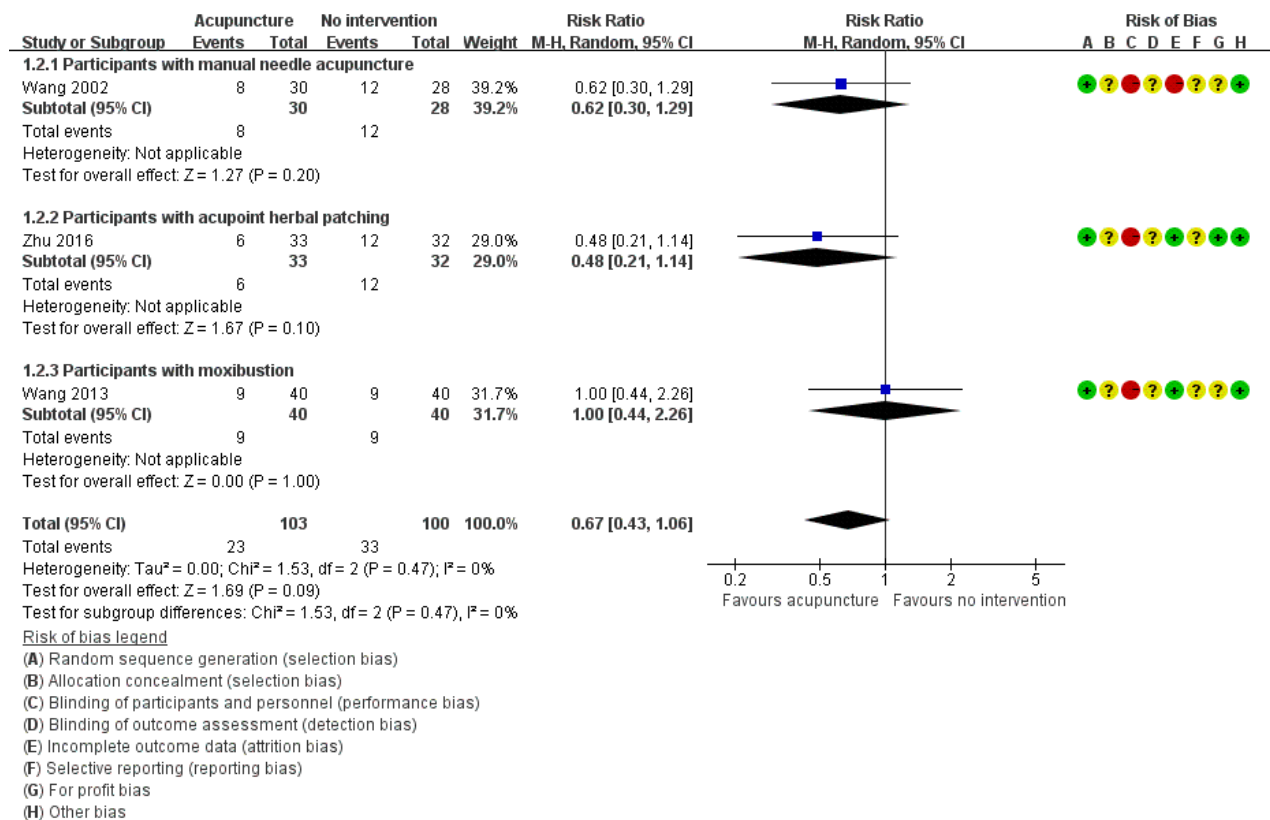


图 2-5 针灸治疗慢性乙型肝炎一项或多项非严重不良事件发生率-
不同针灸方式亚组分析森林图

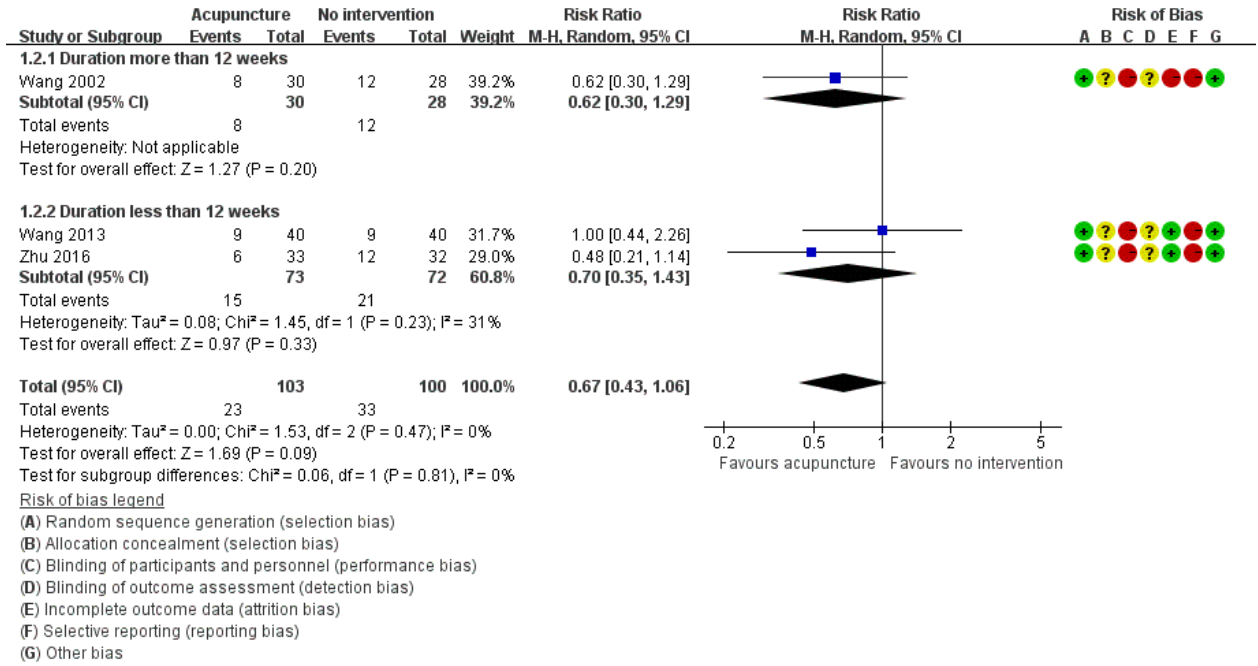


图 2-6 针灸治疗慢性乙型肝炎一项或多项非严重不良事件发生率-
不同针灸疗程亚组分析森林图

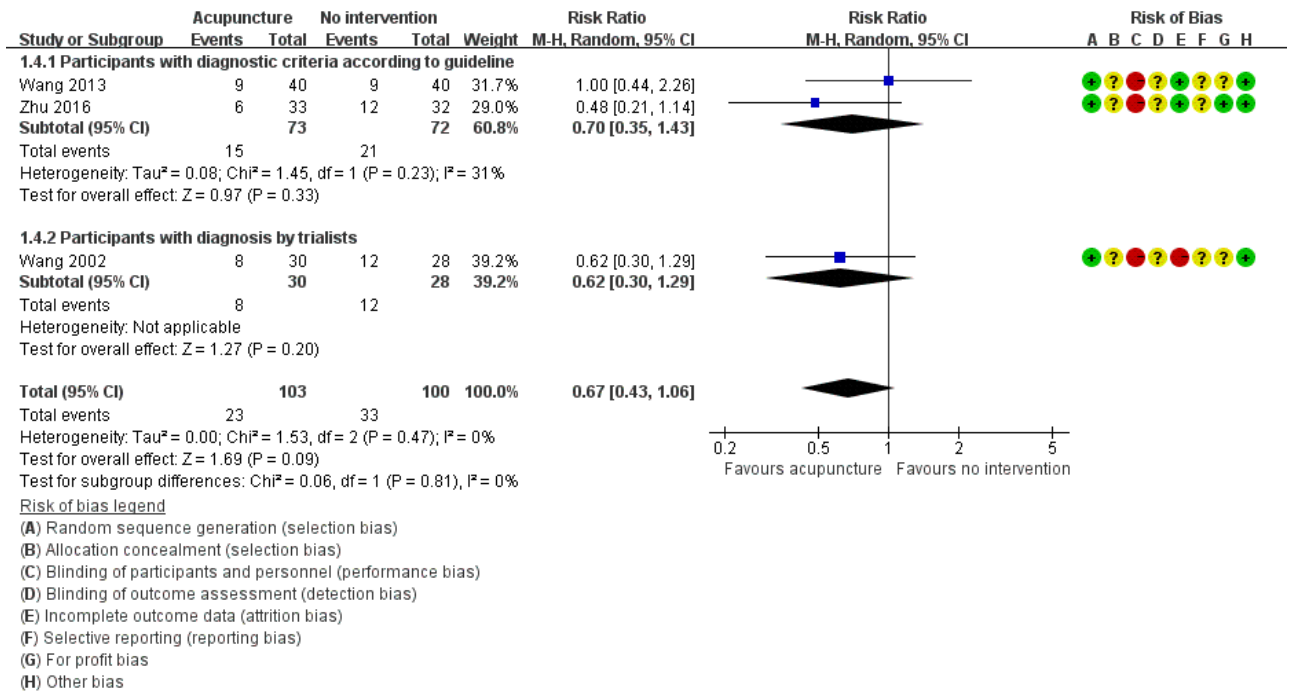


图 2-7 针灸治疗慢性乙型肝炎一项或多项非严重不良事件发生率-
不同诊断标准亚组分析森林图

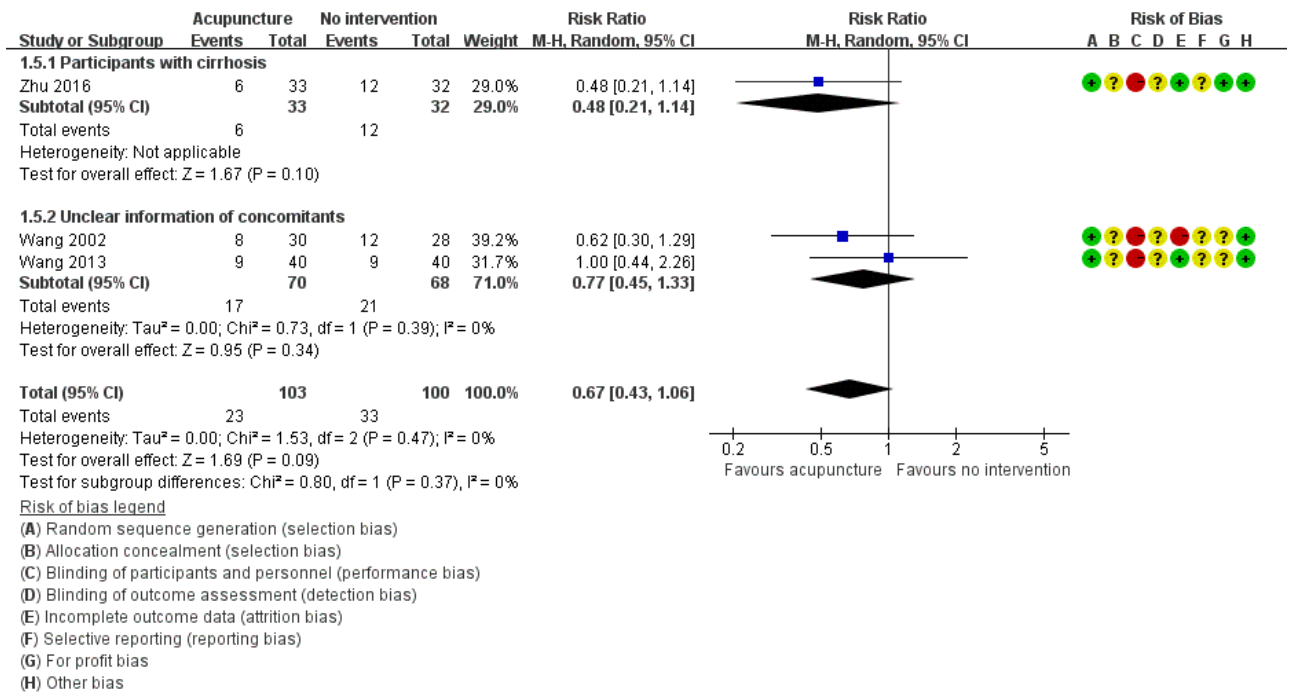


图 2-8 针灸治疗慢性乙型肝炎一项或多项非严重不良事件发生率-
不同合并疾病亚组分析森林图

敏感性分析

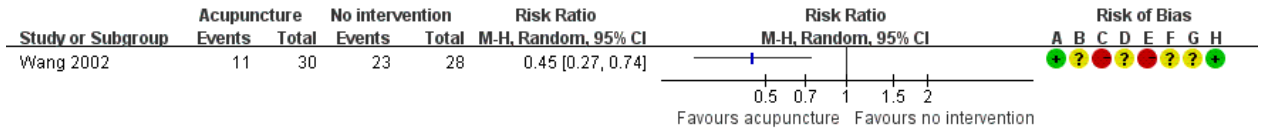
三项报告该指标的研究将所有随机分组的参与者的数据纳入数据分析中。因此，不涉及‘最好-最坏’情况或‘最坏-最好’情况分析。

运用 GRADE 分析和 TSA 分析评估关于该指标证据的不精确性，得到不一致结果：运用 GRADE 分析对证据质量不精确性评估时，由于最优样本量未达到，样本量小于 4000，我们将证据质量降一级^[131, 132]；运用 TSA 分析对证据质量不精确性评估时，由于统计差异边界（有益、有害及无效）均未被穿过，对证据质量降两级^[98]。

3.3 探索性结局指标

3.3.1 HBV-DNA 阳性的参与者比例

只有一项研究涉及 58 例参与者报告该指标数据（Wang 2002）。针灸与更低的 HBV-DNA 阳性检测率有关（RR=0.45，95%CI [0.27, 0.74]；图 2-9）。



Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) For profit bias
- (H) Other bias

图 2-9 针灸治疗慢性乙型肝炎 HBV-DNA 阳性的参与者比例森林图

试验序贯分析

我们运用如下参数进行分析：对照组事件发生比例 82.1%，风险度减低 15%， α 值为 5%，把握度为 90%，多样性为 0%，结果显示 Z 曲线未到达 DARIS（506 名参与者），也未穿过统计差异边界（有益、有害及无效）。试验序贯分析显示无充分证据支持或者拒绝针灸疗法联合常规疗法对比单用常规疗法可降低 HBV-DNA 阳性患者比例 15% 风险（TSA 校正后的 RR=0.45，95%CI [0.07, 2.95]；图 2-10）。

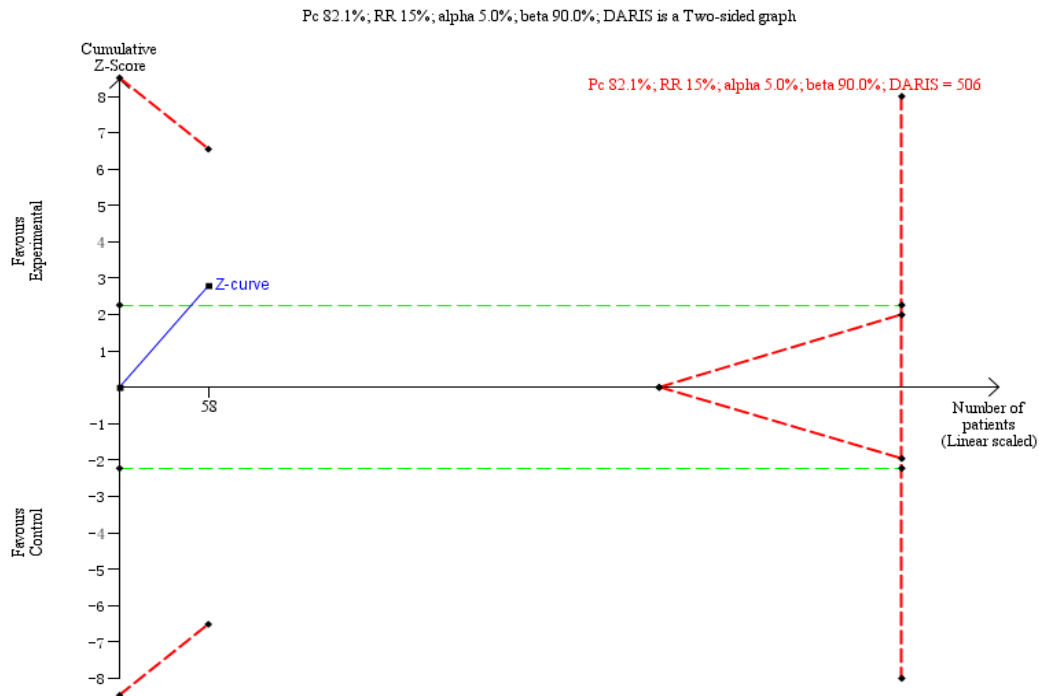


图 2-10 针灸治疗慢性乙型肝炎 HBV-DNA 阳性的参与者比例试验序贯分析图

亚组分析

由于只有一项研究提供了 HBV-DNA 阳性患者比例（Wang 2002），我们无法进行亚组分析。

敏感性分析

Wang 2002 研究中，对照组中的两例患者由于发生严重不良事件（具体未报告），因此在数据分析时排除了该两例患者。敏感性分析显示，排除该两例患者数据后，未影响 HBV-DNA 阳性检测率结局的 meta 分析结果。（‘最好-最坏’情况：RR=0.44，95%CI [0.27, 0.72]，P=0.001，一项研究，60 例参与者，图 38；‘最坏-最好’情况：RR=0.48，95%CI [0.29, 0.80]，P=0.005，一项研究，60 例参与者，图 2-11。）

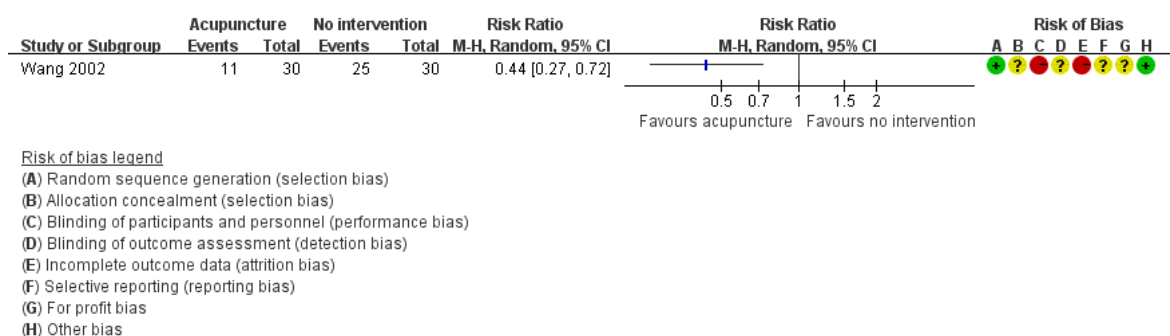


图 2-11 针灸治疗慢性乙型肝炎 HBV-DNA 阳性检测率最好-最坏情况森林图

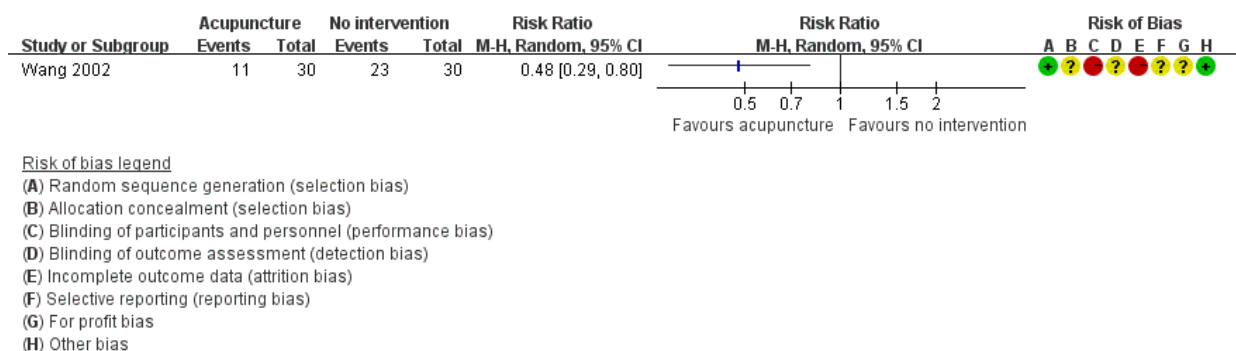


图 2-12 针灸治疗慢性乙型肝炎 HBV-DNA 阳性检测率最坏-最好情况森林图

运用 GRADE 分析和 TSA 分析评估关于该指标证据的不精确性，结果并不一致：运用 GRADE 分析对证据质量不精确性评估时，由于最优样本量未达到，样本量小于 4000，我们将证据质量降一级^[131, 132]；运用 TSA 分析对证据质量不精确性评估时，由于统计差异边

界（有益、有害及无效）均未被穿过，对证据质量降两级^[98]。

3.3.2 HBeAg 阳性的参与者比例

两项研究涉及 158 例参与者报告该指标数据（Wang 2002；Zou 2011）。固定效应模型显示针灸可降低 HBeAg 阳性检测率(RR=0.79, 95% CI [0.69, 0.91], P=0.0007; I²=98%; 图 2-13)。但是, 随机效应模型显示针灸对 HBeAg 阳性检测率无显著统计性效应(RR=0.64, 95%CI [0.11, 3.68], P=0.61; I²=98%; 图 2-14)。我们对于更接近零效应的结果赋予更多的权重 (RR=0.64, 95%CI [0.11, 3.68], P=0.61)。本例中随机效应模型结果为考虑了研究间异质性而分析得出的结果, 因此我们不确定针灸是否对 HBeAg 阳性检测率改变有效应。

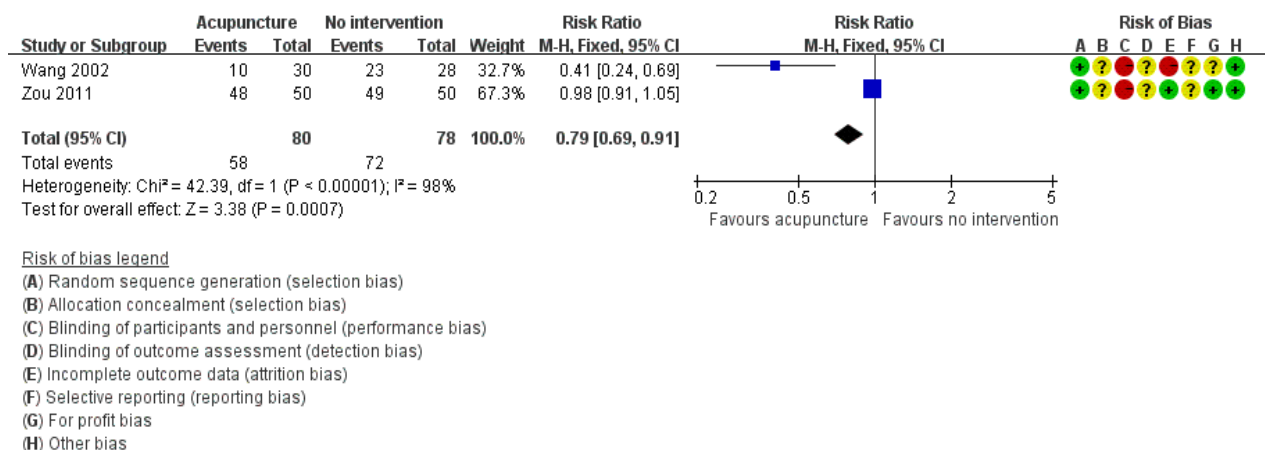


图 2-13 针灸治疗慢性乙型肝炎 HBeAg 阳性参与者固定效应模型森林图

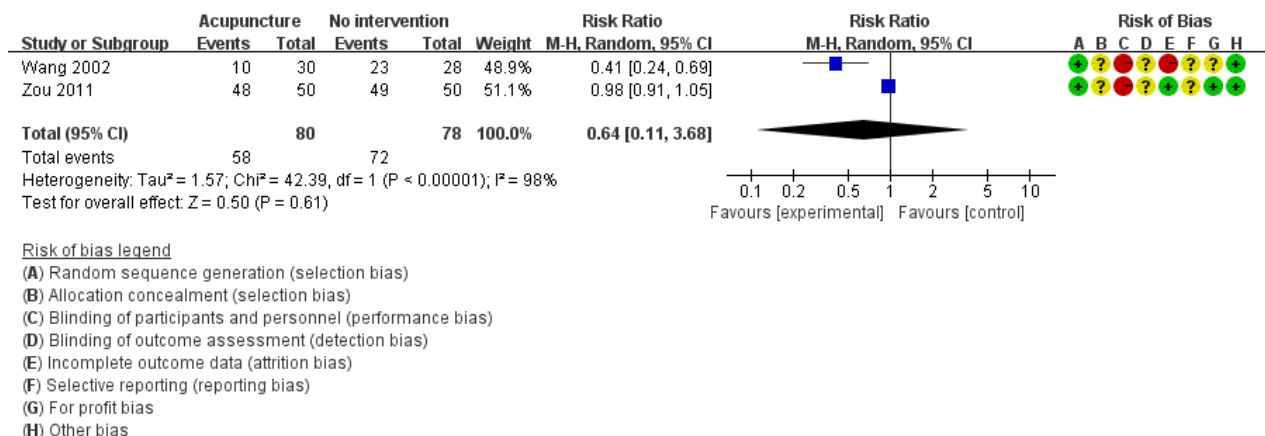


图 2-14 针灸治疗慢性乙型肝炎 HBeAg 阳性参与者随机效应模型森林图

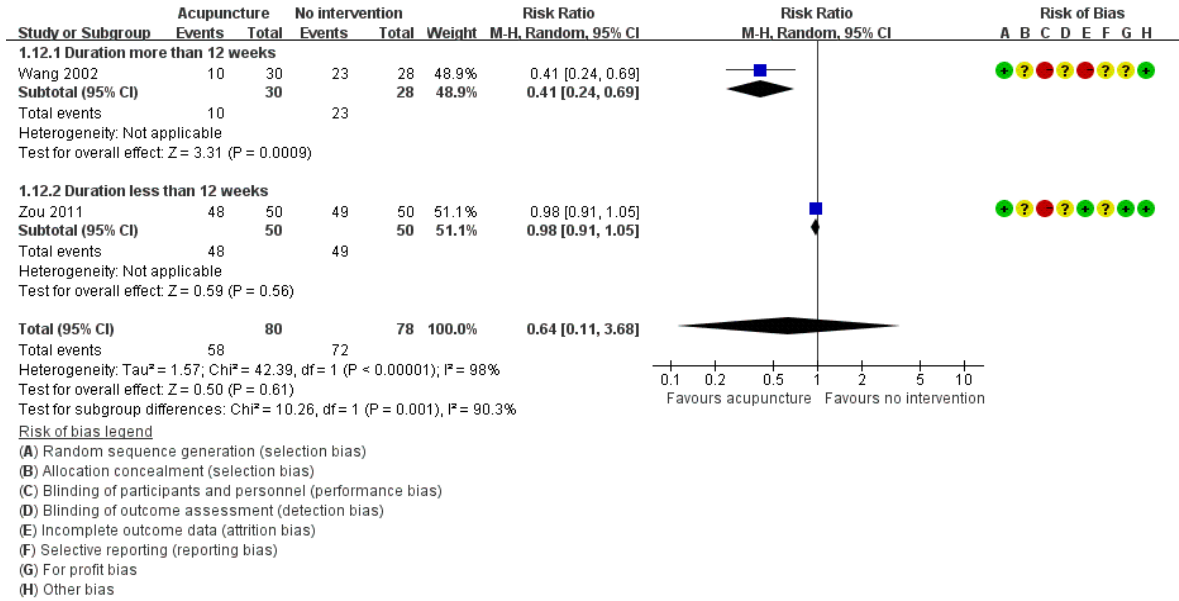


图 2-16 针灸治疗慢性乙型肝炎 HBeAg 阳性参与者-
不同疗程亚组分析森林图

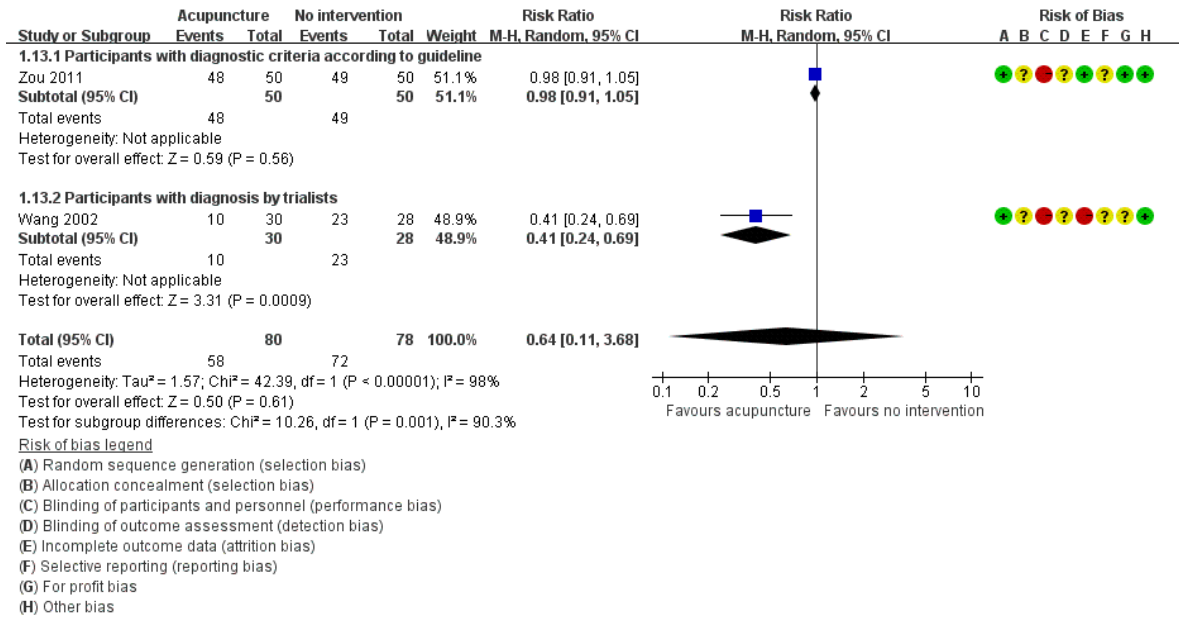


图 2-17 针灸治疗慢性乙型肝炎 HBeAg 阳性参与者-
不同诊断标准亚组分析森林图

敏感性分析

在两项研究报告了所有参与者关于 HBeAg 阳性比例的研究中(Wang 2002; Zou 2011),

一项研究排除了对照组中两名由于出现严重不良事件而退出试验的参与者数据 (Wang 2002)。敏感性分析显示, Wang 2002 研究中排除的 2 例参与者数据并未影响我们的 meta-分析结果 (‘最好-最坏’情况: RR=0.63, 95%CI [0.10, 3.90], P=0.62, 160 名参与者, 2 项研究, 图 2-18; ‘最坏-最好’情况: RR=0.66, 95%CI [0.13, 3.39], P=0.62, 160 名参与者, 图 2-19)。

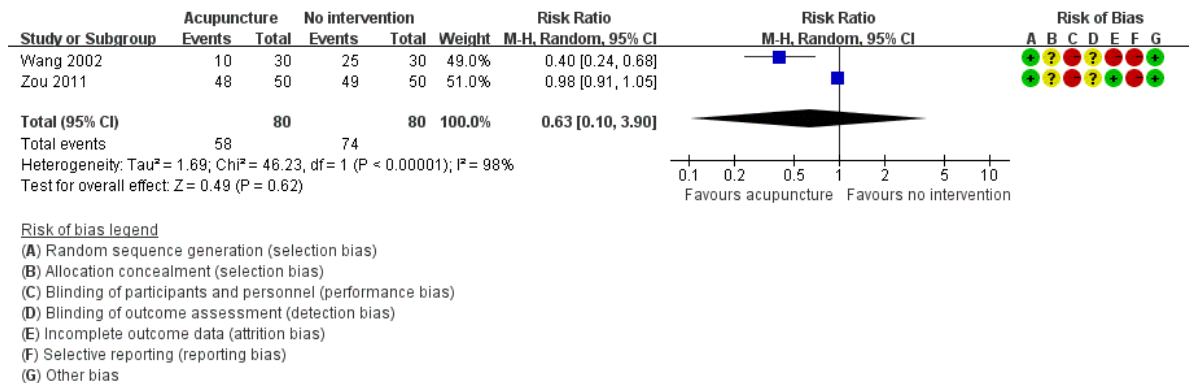


图 2-18 针灸治疗慢性乙型肝炎 HBeAg 阳性参与者最好-最坏情况森林图

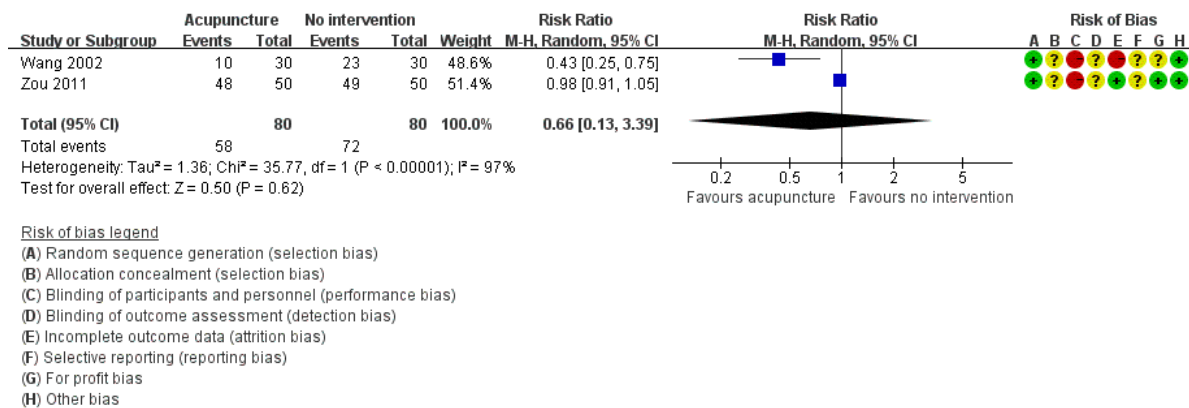


图 2-19 针灸治疗慢性乙型肝炎 HBeAg 阳性参与者最坏-最好情况森林图

运用 GRADE 分析和 TSA 分析评估关于该指标证据的不精确性, 得到一致结果: 运用 GRADE 分析对证据质量不精确性评估时, 由于事件发生数小于 300, 并且可信区间 CI 包含了无效线以及显著疗效 (RR < 0.75) 或显著伤害 (RR > 1.25), 对证据质量降两级^[131, 132]; 运用 TSA 分析对证据质量不精确性评估时, 由于统计差异边界 (有益、有害及无效) 均未被穿过^[98], 对证据质量降两级。

3.3.3 分别报道的非严重不良事件发生率

根据研究方案，对于只纳入一项研究并且数据为二分类变量的分析，我们运用 Fisher 精确检验。由于 Fisher 精确检验与 RevMan 检验结果一直，我们运用 RevMan 检验结果展示相关结局指标的结果。

不良事件乏力的发生率：只有一项研究包含 80 例参与者报告了不良事件乏力的发生率（Wang 2013）。我们不确定针灸是否对不良事件乏力的发生率有效应（RR=0.67, 95%CI[0.12, 3.78]；图 2-20）。

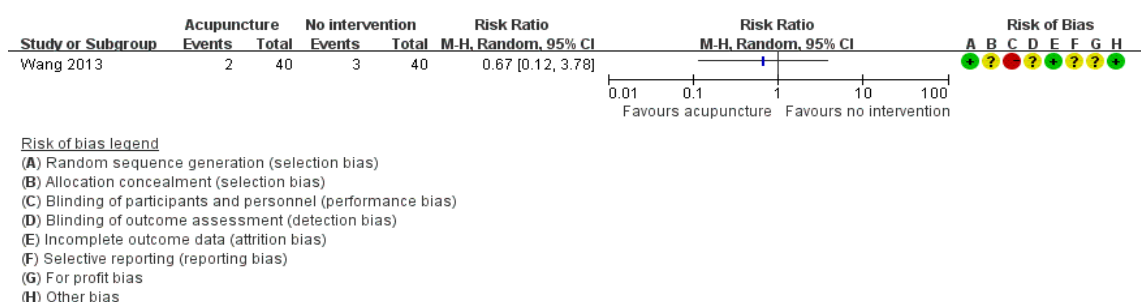


图 2-20 针灸治疗慢性乙型肝炎不良事件乏力的发生率森林图

不良事件腹胀的发生率：只有一项研究包含 80 例参与者报告了不良事件腹胀的发生率（Wang 2013）。Meta 分析结果显示，针灸可升高不良事件腹胀的发生率（RR=4.50, 95%CI[1.04, 19.54]；图 2-21）。

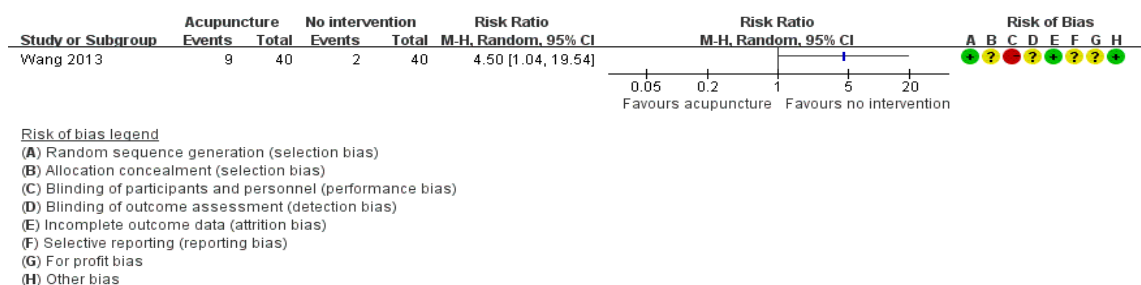
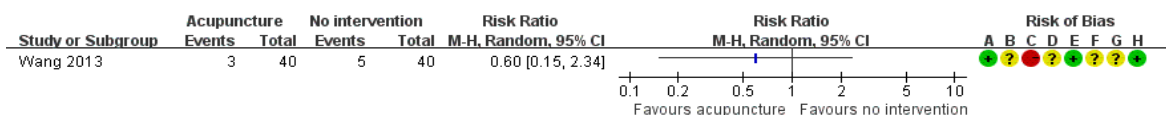


图 2-21 针灸治疗慢性乙型肝炎不良事件腹胀的发生率森林图

不良事件食欲不振的发生率：只有一项研究包含 80 例参与者报告了存在食欲不振的参与者比例（Wang 2013）。我们不确定针灸是否对不良事件食欲不振的发生率有效应（RR=0.60, 95%CI[0.15, 2.34]；图 2-22）。

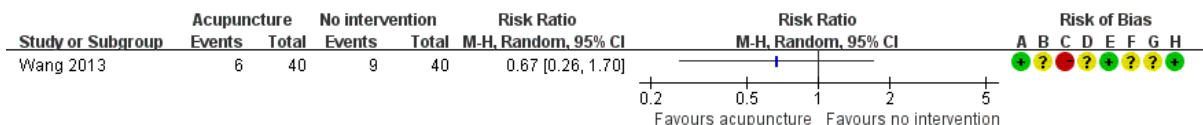


Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) For profit bias
- (H) Other bias

图 2-22 针灸治疗慢性乙型肝炎不良事件食欲不振的发生率森林图

不良事件肋肋疼痛的发生率：只有一项研究包含 80 例参与者报告了不良事件肋肋疼痛的发生率（Wang 2013）。我们不确定针灸是否对不良事件肋肋疼痛的发生率有效应（RR 0.67，95%CI[0.26, 1.70]；图 2-23）。

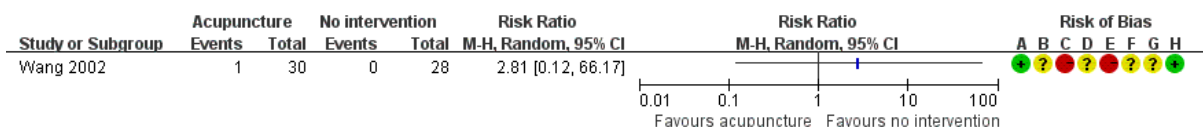


Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) For profit bias
- (H) Other bias

图 2-23 针灸治疗慢性乙型肝炎不良事件肋肋疼痛的发生率森林图

不良事件心动过缓的发生率：只有一项研究包含 58 例参与者报告了不良事件心动过缓的发生率（Wang 2002）。我们不确定针灸是否对不良事件心动过缓的发生率有效应（RR=2.81，95%CI[0.12, 66.77]；图 2-24）。

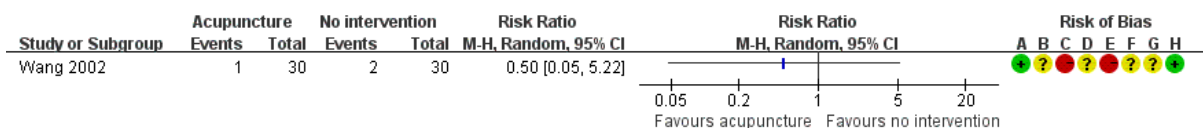


Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) For profit bias
- (H) Other bias

图 2-24 针灸治疗慢性乙型肝炎不良事件心动过缓的发生率森林图

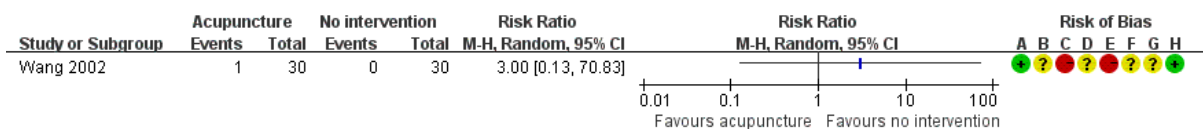
敏感性分析显示，Wang 2002 研究中排除的 2 例参与者数据并未影响该不良事件心动过缓的发生率的 meta 分析结果（‘最好-最坏’情况：RR=0.50，95%CI[0.05, 5.22]，P=0.56，60 名参与者，1 项研究，图 2-25；‘最坏-最好’情况：RR=3.00，95%CI[0.13 to 70.83]，P=0.50，60 名参与者，图 2-26）。



Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) For profit bias
- (H) Other bias

图 2-25 针灸治疗慢性乙型肝炎不良事件心动过缓的发生率最好-最坏情况森林图



Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) For profit bias
- (H) Other bias

图 2-26 针灸治疗慢性乙型肝炎不良事件心动过缓的发生率最坏-最好情况森林图

不良事件血常规显著改变的发生率：只有一项研究包含 58 例参与者报告了不良事件血常规显著改变的发生率（Wang 2002）。我们不确定针灸是否对不良事件血常规显著改变的发生率有效应（RR=0.62，95%CI[0.30, 1.29]；图 2-27）。

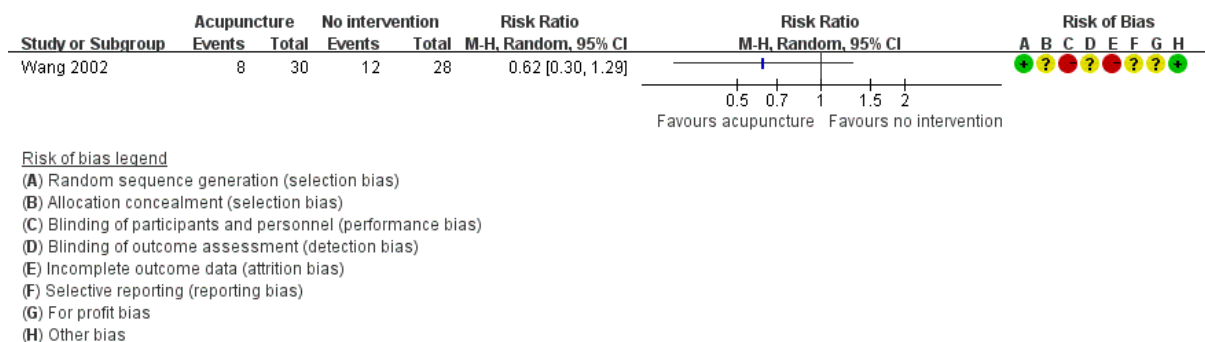


图 2-27 针灸治疗慢性乙型肝炎不良事件血常规显著改变的发生率森林图

敏感性分析显示，Wang 2002 研究中排除的 2 例参与者数据并未影响不良事件血常规显著改变的发生率的 meta 分析结果（‘最好-最坏’情况：RR=0.57，95%CI[0.28, 1.16]，P=0.12，60 名参与者，1 项研究，图 2-28；‘最坏-最好’情况：RR=0.67，95%CI[0.32, 1.39]，P=0.28，60 名参与者，1 项研究，图 2-29）。

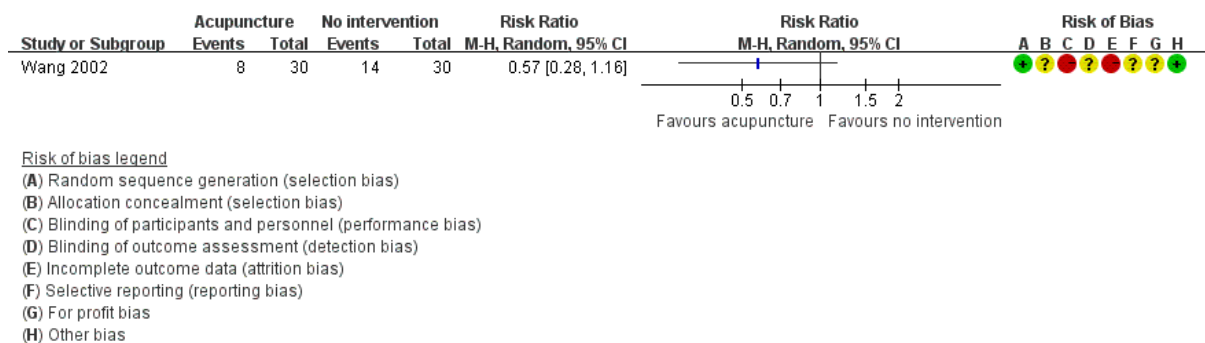
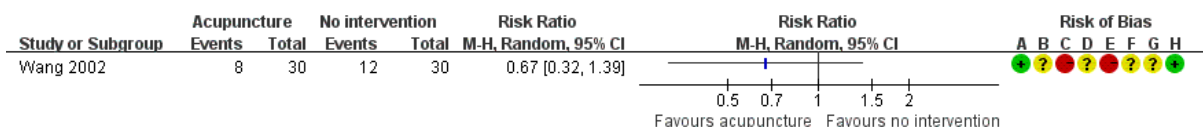


图 2-28 针灸治疗慢性乙型肝炎不良事件血常规显著改变的发生率-最好-最坏情况森林图



Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) For profit bias
- (H) Other bias

图 2-29 针灸治疗慢性乙型肝炎不良事件血常规显著改变的发生率-最坏-最好情况森林图

4.GRADE 分级

我们构建了结果总结表（SOF 表，Summary of finding table）以展示针灸疗法联合常规疗法对比单用常规疗法干预慢性乙型肝炎的结果。然而，由于缺少试验数据，我们只能展示一项或多项非严重不良事件发生率，阳性 HBV-DNA 参与者比例以及阳性 HBeAg 参与者比例相关证据的 GRADE 评估结果。我们针对偏倚风险、无法解释的异质性或者结果的不一致性、不精确性以及发表偏倚三个因素对证据等级进行降级，最终认为所获得的证据确定程度极低（表 2-2）。

表 2-2 针灸治疗慢性乙型肝炎结果总结表(SOF 表)

结果总结:

针灸疗法联合常规疗法对比单用常规疗法干预慢性乙型肝炎人群

患者或人群: 慢性乙型肝炎
 环境: 门诊或住院患者
 干预: 针灸疗法联合常规治疗
 对照: 常规治疗

结局	预期绝对效应量* (95% CI)		相对效果(95% CI)	参与者数量 (研究数量)	证据的确定性 (GRADE)	评论
	单用常规治疗 的风险	常规治疗联合 针灸疗法的 风险				
全因死亡率 - 未报告	-	-	-	-	-	
发生一项或多项严重不良事件参与者比例 - 未报告	-	-	-	-	-	
健康相关生存质量 - 未报告	-	-	-	-	-	
乙肝相关死亡率 - 未报告	-	-	-	-	-	
乙肝相关并发症 - 未报告	-	-	-	-	-	
发生一项或多项非严重不良事件患者比例 随访:范围 2 到 6 个月	330 每 1,000	221 每 1,000 (142 到 350)	相对危险度 0.67 (0.43 到 1.06)	203 (3 随机对照试验 (RCT))	⊕ VERY LOW ^{a,b,c}	b.最优样本量未达到, 样本量不够大 (小于 4000 参与者)
HBV-DNA 阳性检测率 随访:6 个月	821 每 1,000	370 每 1,000 (222 到 608)	相对危险度 0.45 (0.27 到 0.74)	58 (1 随机对照试验 (RCT))	⊕ VERY LOW ^{a,b,c}	b.最优样本量未达到, 样本量不够大 (小于 4000 参与者)
HBeAg 阳性检测率 随访:范围 3 到 6 个月	923 每 1,000	591 每 1,000 (102 到 1000)	相对危险度 0.64 (0.11 到 3.68)	158 (2 随机对照试验 (RCT))	⊕ VERY LOW ^{a,b,c}	b. 事件发生数 < 300, 置信区间包含 重大获益值 (RR < 0.75) 及重大风险 值 (RR > 1.25)

*干预组风险 (及其 95%可信区间) 以对照组假设风险以及干预的相对效果 (及其 95%可信区间) 为基础。
 CI: Confidence interval; RR: Risk ratio

GRADE 工作组证据等级标准:

- 高确定性: 我们非常确定真实效应与估计效应非常接近
- 中等确定性: 我们对估计效应中等程度确定, 即真实效应很可能与估计效应非常接近, 但是可能存在大量差异
- 低确定性: 我们对于估计效应的确定程度有限: 真实效应可能与估计效应存在较大差异
- 非常低确定性: 我们对于估计效应非常不确定: 真实效应非常可能与估计效应存在较大差异

表注:

- a. 由于存在分配隐藏、施盲、选择结局指标报告偏倚而降一级
- b. 由于结果的不精确性, 对于可检测到的 HBeAg 阳性检测率这一结局指标降两级, 对于‘一项或多项非严重不良事件发生率’和‘HBV-DNA 阳性检测率’结局指标降一级
- c. 由于发表偏倚降一级: 所有纳入的研究都是小样本量研究, 强烈怀疑存在发表偏倚

5. 后验性指标 meta 分析

临床上常常通过血清 ALT 及 AST 水平的变化来判断患者肝功能情况, 以进一步判断慢性乙型肝炎的轻重程度及患者的预后。参考临床医生的意见, 我们后验性的对于血清 ALT 及 AST 水平这两个指标进行后验性的 meta 分析, 具体如下:

参与者血清 ALT 水平: 四项研究包含 275 例参与者报告了参与者血清 ALT 水平(Liao 2011, Wang 2013, Zhang 2014, Zhu 2016)。针灸(手针、艾灸、穴位贴敷)可降低参与者血清 ALT 水平 (MD=-21.35, 95%CI[-33.60, -9.09]; 图 2-30)。纳入研究的异质性较大 ($I^2=81%$)，通过分析研究的临床特点发现，四项研究采用的针灸治疗方式不同，Zhu 2016 纳入的患者均合并腹水，而其他三项研究对此未做明确限制，Zhang 2014 疗程 12 个月，而其他三项研究疗程均未超过 6 个月，这可能是导致异质性的原因。

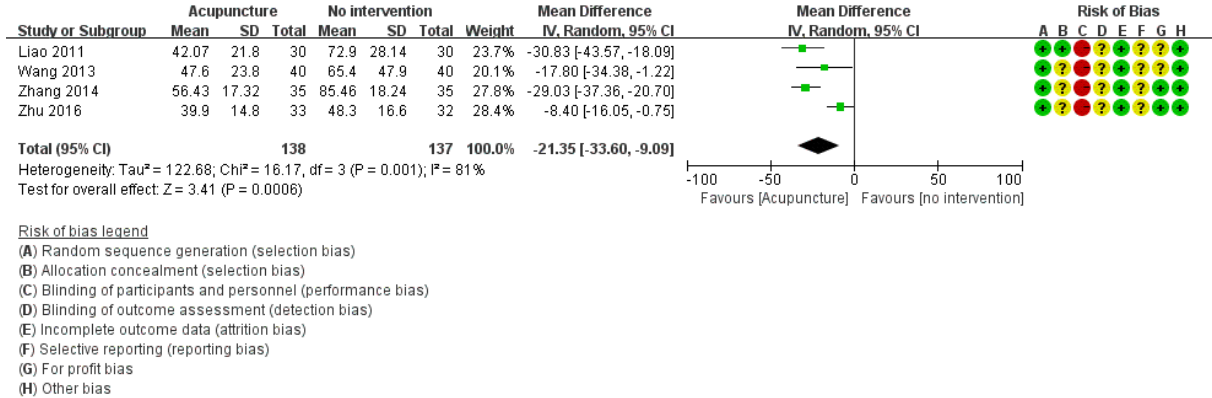


图 2-30 针灸治疗慢性乙型肝炎参与者血清 ALT 水平森林图

试验序贯分析

对于该结局指标，通过试验序贯分析获得的 DARIS 为 1048 名参与者。我们运用如下参数进行分析：对照组 SD 为 34.18，最低标准差差异为 (MIREDF) -17.09，对照组方差为 1168.3， α 值为 5%，把握度为 90%，多样性为 83.88%，结果显示 Z 曲线未到达 DARIS (1048 名参与者)，虽然穿过统计差异边界 (有益)，但随后 Z 曲线又回归到统计差异边界以下，试验序贯分析显示无充分证据支持或者拒绝针灸联合常规疗法对比单用常规疗法可降低参与者血清 ALT 水平 (TSA 校正的 MD=-21.35; 95% CI [-47.77, 5.08]; 图 2-31)。

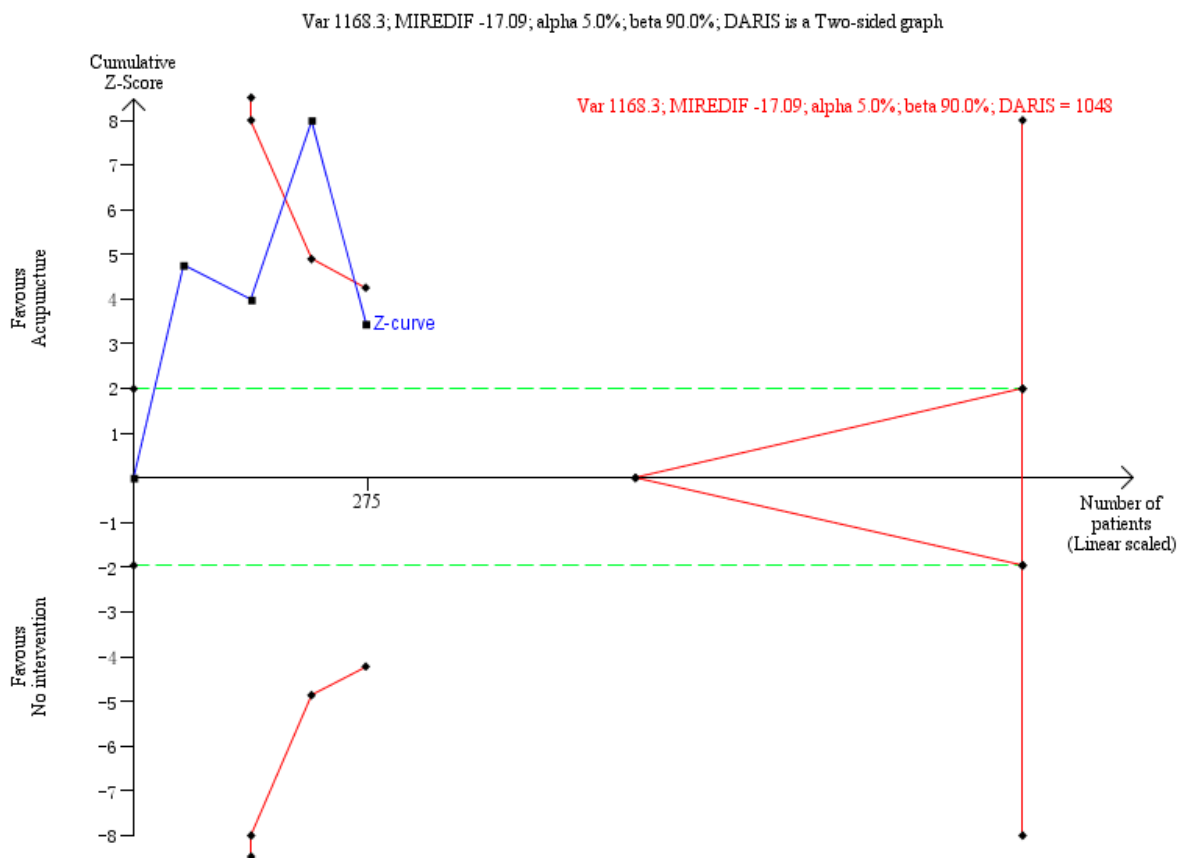
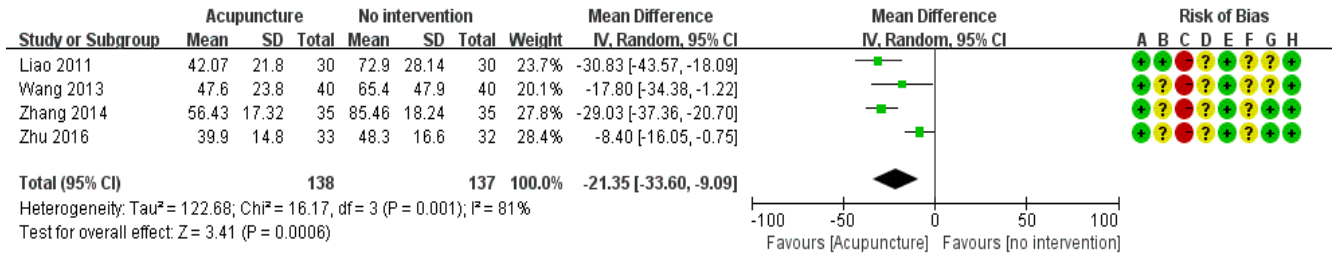


图 2-31 针灸治疗慢性乙型肝炎参与者血清 ALT 水平试验序贯分析图

参与者血清 AST 水平: 四项研究包含 275 例参与者报告了参与者血清 AST 水平(Liao 2011, Wang 2013, Zhang 2014, Zhu 2016)。针灸(手针、艾灸、穴位贴敷)可降低参与者血清 AST 水平(MD=-18.31, 95%CI[-35.44, -1.18]; 图 2-32)。纳入研究的异质性较大($I^2=94%$)，通过分析研究的临床特点发现，四项研究采用的针灸治疗方式不同，Zhu 2016 纳入的患者均合并腹水，而其他三项研究对此未做明确限制，Zhang 2014 疗程 12 个月，而其他三项研究疗程均未超过 6 个月，这可能是导致异质性的原因。



Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) For profit bias
- (H) Other bias

图 2-32 针灸治疗慢性乙型肝炎参与者血清 AST 水平森林图

对于该结局指标，通过试验序贯分析获得的多样性校正后的所需样本量（DARIS）为 2466 名参与者。我们运用如下参数进行分析：对照组 SD 为 34.81，最低标准差差异为(MIREDIF)-17.41, 对照组方差为 1121.7, α 值为 5%, 把握度为 90%, 多样性为 93.68%。结果显示 Z 曲线未到达 DARIS（2466 名参与者），也未穿过统计差异边界（有益、有害及无效）。试验序贯分析显示无充分证据支持或者拒绝针灸联合常规疗法对比单用常规疗法可降低参与者血清 AST 水平(TSA 校正的 MD=-18.31; 95% CI [-88.22, 51.61])。

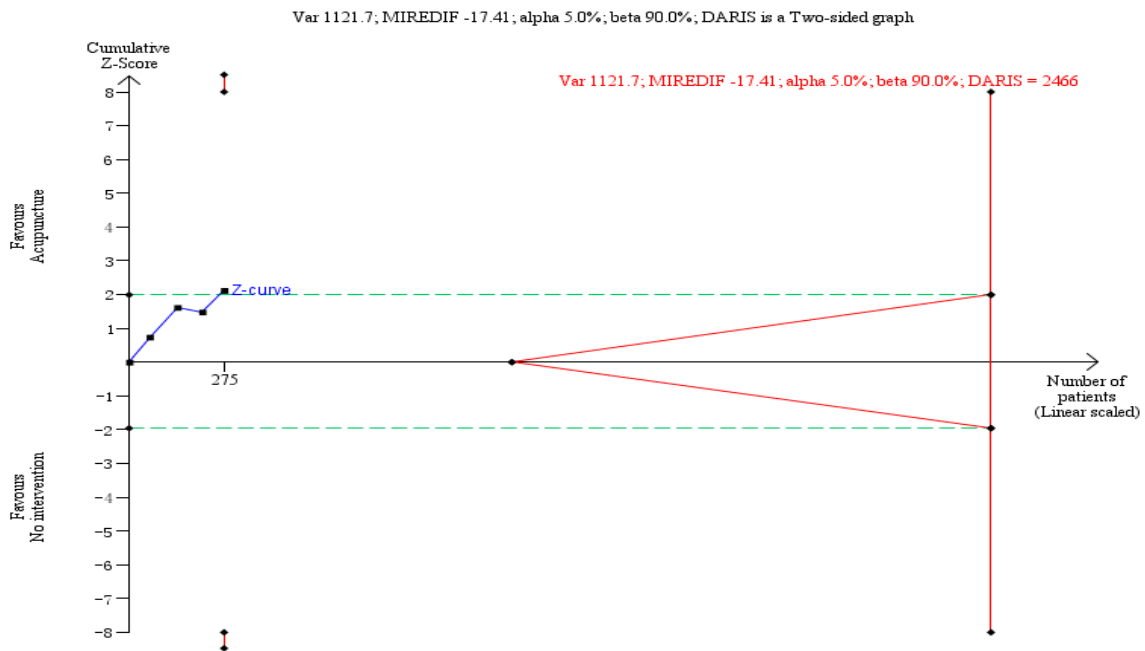


图 2-333 针灸治疗慢性乙型肝炎参与者血清 AST 水平试验序贯分析图

讨 论

1. 效应分析

本研究纳入了 8 项随机对照试验共包含 552 例参与者。全部 8 项研究被评定为高偏倚风险研究。最终只能从 4 项共包含 305 例参与者的研究中提取定量数据。纳入研究均未评估本研究的主要结局指标及两项次要结局指标（乙肝相关并发症及乙肝相关死亡率）。所有纳入研究均比较针灸疗法联合常规疗法对经常规治法。

对于发生一项或多项不良事件参与者比例这一结局指标，基于三项低质量随机对照试验的证据显示针灸疗法联合常规疗法对比单用常规疗法无显著差异性效应。亚组分析显示一致结果。试验序贯分析表明需要更多信息以得到进一步结论。敏感性分析表明，运用 GRADE 分级对该证据的不精确性降两级，而运用试验序贯分析对该证据的不精确性降一级。

对于 HBV-DNA 阳性检测率这一替代结局指标，只基于一项低质量随机对照试验的证据显示针灸疗法联合常规疗法对比单用常规疗法可降低 HBV-DNA 阳性检测率。试验序贯分析表明无充分证据支持或者拒绝针灸疗法对于这一结局指标有效应。敏感性分析显示，不完整结局指标偏倚对该指标结局结果无影响。

对于 HBeAg 阳性检测率这一替代结局指标，我们不确定针灸疗法联合常规疗法对比单用常规疗法是否具有效应。试验序贯分析表明无充分证据支持或者拒绝针灸疗法对于这一结局指标有效应。亚组分析显示一致结果。敏感性分析显示，不完整结局指标偏倚对该指标结局结果无影响。

对于分开报告的不良事件结局，即发生乏力、腹胀、食欲不振、胁肋疼痛、心动过缓、血常规显著异常等事件，只有一项纳入 80 例参与者的研究表明针灸可能导致发生腹胀比例的参与者升高（RR=4.50，95%CI[1.04, 19.54]）。对于其他事件，我们不确定针灸疗法联合常规疗法对比单用常规疗法是否具有效应。敏感性分析显示，不完整结局数据偏倚对该指标结局结果无影响。

对于参与者血清 ALT、AST 水平，我们不确定小柴胡汤联合常规疗法对比单用常规疗法是否有效应。

2. 证据的整体完整性和适用性

我们联系了 87 项研究的作者以获得有关试验设计和方法学的更多信息。但是，我

们没能与任何研究者取得联系，也未能获得任何有用的相关信息。我们发现了两项正在进行的试验，比较针灸疗法联合共同干预措施对照同等实施的共同干预措施。我们没有发现任何未发表的试验。73项未能描述随机序列生成方法，6项研究包含的信息不足以判断它们与本系统综述的相关性，这79项研究缺乏以患者为中心的结果的信息，因此未被纳入我们的系统综述中。

8项试验中包括的参与者均可反映一般慢性乙型肝炎人群的特征。所有试验均未包含儿童。有2项试验纳入了合并肝硬化的参与者（Wang 2013；Zhu 2016），其余6项试验排除了肝硬化、其他肝炎病毒感染和其他严重并发症（如肝癌）的参与者（Wang 2002；Huang 2011；Liao 2011；Zou 2011；Zhang 2014；Jing 2016）。纳入的试验包括手针、针灸注射、草药穴位贴敷和艾灸，试验还包括常用的针灸疗法保留时间、常用的治疗时间和随访时间。我们没有找到关于以患者为中心的结局指标如全因死亡率、健康相关的生活质量、乙型肝炎相关死亡率和乙型肝炎相关的发病率的报告和评估。相反，我们发现了替代结果的数据，例如HBV-DNA阳性检测率、HBeAg阳性检测率、参与者ALT、AST水平。一项试验指出了治疗期间发生了严重不良事件。

3. 证据质量

3.1 研究间偏倚风险

纳入的8项研究中，8项（100%）报告了充分的随机序列生成方法，只有1项（0%）报告了随机分配的隐藏，8项（100%）未予参与者施盲，8项（100%）未报告结局指标评估者施盲，因此被评价为结局指标高偏倚风险，6项（75%）表明未受未完成结局数据影响，10项（100%）均具有选择性报告偏倚偏倚风险，3项（37.5%）表明未受利益偏倚影响，10项（100%）表明无其他偏倚风险。因此，所有研究均被评价为高偏倚风险。

3.2 证据的间接性

本研究中所有的纳入研究均评价针灸疗法干预慢性乙型肝炎患者的疗效。我们发现研究参与者与临床人群十分接近。因此，我们未发现证据间接性方面存在严重问题。

3.3 结果的异质性和不一致性

对于发生一项或多项非严重不良事件结局，无显著研究间异质性（ $I^2=0\%$ ）。

对于阳性HBeAg参与者比例这一结局，具有大量的研究间异质性（ $I^2=98\%$ ）。只有两项研究报告这一结局（Wang 2002；Zou 2011）。通过分析研究特点寻找异质性原因发现，Zou 2011以指南为诊断依据纳入参与者，并且疗程为12周，而Wang 2002并

未报告研究纳入的参与者的诊断标准为何，并且疗程为 24 周，这可能是导致异质性的原因。

对于阳性 HBV-DNA 参与者比例这一结局，只有一项研究进行了报告，所以不涉及研究间异质性问题。

对于参与者血清 ALT、AST 水平结局，纳入研究的异质性较大 ($I^2 > 80\%$)，通过分析研究的临床特点发现，通过分析研究的临床特点发现，Zhu 2016 纳入的患者均合并腹水，而其他三项研究对此未做明确限制，Zhang 2014 疗程 12 个月，而其他三项研究疗程均未超过 6 个月，这可能是导致异质性的原因。

3.4 研究结果的不精确性

根据 GRADE 标准，对于发生一项或多项不良事件患者比例这一结局，由于事件发生数小于 300，并且可信区间 CI 包含了无效线以及显著疗效 ($RR < 0.75$) 或显著伤害 ($RR > 1.25$)，对证据质量降两级^[131, 132]；对于阳性 HBV-DNA 参与者比例结局及阳性 HBeAg 参与者比例结局，由于最优样本量未达到，样本量小于 4000，我们将证据质量均降一级^[131, 132]。

运用 TSA 分析对证据质量不精确性评估时，由于统计差异边界(有益、有害及无效)均未被穿过，对发生一项或多项不良事件患者比例、阳性 HBV-DNA 参与者比例以及阳性 HBeAg 参与者比例结局，我们将证据质量降两级^[98]。

3.5 发表偏倚风险

对于本研究结局指标，均包含小于 3 项研究，因此我们无法构建倒漏斗图。由于纳入研究均为小样本研究，并且有大多数研究表明特异的阳性效应，我们怀疑纳入研究具有发表偏倚^[113, 114]。

4. 本研究中可能存在的偏倚

未发表的具有阴性结果的文章可能是潜在的偏倚来源。我们试图联系纳入本研究的 8 项研究及潜在可能为随机对照试验的 79 项研究的作者以获得方法学或者结局指标相关缺失数据，但未与任何一项研究者成功取得联系。

严重不良事件报告的缺失可能会导致对以患者为中心的结局指标结论的缺失。

对于非严重不良事件，我们选择试验组或者对照组出现的最高事件数量以计算发生一个或多个非严重不良事件参与者比例。这一解决措施可能由于低估了发生非严重不良事件的参与者比例而存在问题。

我们进行了 7 项亚组分析并纳入了大量的替代结局指标。因此每个结局指标的 7 项亚组分析可能存在重复分析问题^[133]。

由于本研究进行的序贯分析中，所有分析的所需样本量以及统计差异边界均未被穿过，因此不能排除本研究涉及结局指标结果的随机错误风险。

小 结

1. 以患者为中心的结局指标如全因死亡率，严重不良事件，健康相关生存质量，乙肝相关死亡率以及乙肝相关并发症的相关数据并未被报告。通过评价低质量的证据，我们发现针灸疗法联合常规疗法对比单用常规疗法可降低 HBV-DNA 阳性检测率、降低参与者血清 ALT、AST 水平，但对于非严重不良事件和 HBeAg 阳性检测率这两个结局指标，目前尚不能表明针灸疗法（手针、艾灸、穴位注射、中药穴位贴敷）联合常规治疗对比常规治疗具有显著有效性。由于检索到的研究具有高风险偏倚，HBV-DNA 阳性检测率、参与者血清 ALT、AST 水平是未经证实的替代结局指标，且只有一项研究能够提供数据，未达到所需样本量，纳入的研究其确定性较低，因此关于针灸疗法对慢性乙型肝炎的有益性结论尚不能确证。而对于针灸疗法增加慢性乙型肝炎患者腹胀事件的风险，同样由于只有一项研究能够提供数据，且纳入的研究其确定性较低，其结论同样不能确证。

2. 纳入的针灸疗法治疗慢性乙型肝炎的研究均具有高风险偏倚，研究质量均很低。在今后的临床随机对照试验中，对于随机序列生成方法、随机分配的隐藏以及盲法应进行严格设计、实施和详尽报告，应尽量避免结局指标报告不完整或者存在缺失数据，未来研究报告的透明度和报告水平亟需提高。对于针灸疗法干预慢性乙型肝炎的临床随机对照试验：应设计如激光针刺等更多疗法，或者多种疗法联合应用于干预慢性乙型肝炎的研究；更多设计针刺疗法对比假针刺疗法的临床试验；应该对于研究人群的诊断标准以及合并疾病进行更严格清晰的限定，并且应针对不同的合并疾病进行分层随机；应设计更长随访期和更大样本量的临床随机对照试验；应设计更多针对以患者为中心的结局指标（如死亡率、严重不良事件、健康相关生存质量等）的临床试验，并应设计临床随机对照试验以验证替代结局指标与以患者为中心的结局指标的关系。通过提高相关中医临床研究质量，以进一步获得关于针灸疗法干预慢性乙型肝炎的高质量临床证据，从而指导针灸疗法的临床应用实施。

论文三 三维误差证据矩阵在中医临床证据评价中应用的探索分析

最可靠的证据应该具有最高的方法学质量^[203]及影响真实性的最低风险,应当最接近真实情况,具有指导临床实践的最强影响力量^[204]。基于高质量的随机对照试验进行 meta 分析的系统综述,并通过 GRADE 分级得出的证据,被公认为是最高等级证据^[204],还有一些观点认为,基于网络 meta 分析的系统综述得出的证据为最高等级证据。最高等级证据应被应用于临床指南的形成^[205-207]。自 2000 年起,有学者大量地将系统综述和 GRADE 分级引入中医药的有效性和安全性研究中^[208,209]。临床决策制定者应当获得最可靠证据以指导临床实践^[5,8],包括中医药领域。心律失常抑制研究(CAST)是一项随机临床对照试验,该试验以死亡率为结局指标,得出了震惊医疗界及学术界的结果,即抗心律失常干预对比安慰剂治疗增加患者 3 倍死亡率^[210,211]。CAST 研究结论,最终使得心梗后预防猝死而应用抗心律失常药物这一干预措施的停用。尽管在此前的一项基于替代指标的随机临床对照试验得出抗心律失常干预措施即利多卡因可以抑制初发室颤,并因此得以在临床实施数十年。高质量的临床随机对照试验通常以患者为中心的结局为结局指标^[203,212],更接近真实情况,是基于有力证据而得出正确和安全的临床决策的根本^[85,203]。

可靠证据应具有内部真实性和外部真实性。内部真实性是指一项研究能够真正回答一个临床问题的能力^[213]。外部真实性是研究结果可以向外推广的程度^[24]。没有内部真实性,研究就不会具有外部真实性。内部真实性及外部真实性都可以被系统误差(即偏倚)、随机误差(即机遇)以及设计误差(即运用错误的设计去回答提出的问题,或者错误的情景)影响^[5,8,22,20]。Cochrane 的偏倚风险工具‘Risk of bias tool’是一个被广泛认可的评估系统误差的评估工具^[205,214]标准误(SE)则被应用于评估随机误差^[22]。对于设计误差评估的方法,主要围绕设计要点(结局指标、参与者选择、干预措施方案等)进行,常用的评估方法如下:INCLIN 的公平滤镜^[215],PROGRESS 评估方法^[216],以及 Bornhöft 检查清单^[217]。Keus 及同事提出了证据矩阵(即误差矩阵,一个可视化评估工具)以评估临床证据的三个维度误差(即系统误差、随机误差以及设计误差)^[22]。

本研究中,我们以两种干预措施为例,即小柴胡汤与针灸疗法(手针、艾灸、穴位注射、穴位贴敷疗法),运用误差矩阵(证据矩阵)方法^[22]以概观中医药干预慢性乙型

肝炎可获得证据的内部真实性及外部真实性。我们通过结局指标重要程度对临床证据的设计误差风险进行排序。此外，通过展示设计误差的各个影响要素的详细情况而进一步评估证据体的设计误差，以进一步改进误差矩阵评估方法。通过系统误差、随机误差、设计误差三个维度全面评估中医临床证据体的质量，补充常用的证据质量评估方法对随机误差和设计误差重视不足的缺陷，从系统误差、随机误差、设计误差三个维度提出中医临床研究设计改进方法，提高中医临床研究质量，促进中医疗法的临床实施。

资料与方法

以小柴胡汤及针灸疗法分别干预慢性乙型肝炎为例，我们构建可视化三维误差证据矩阵以评估中医药干预慢性乙型肝炎临床证据的可靠性。我们通过以下三个维度误差构建三维误差证据矩阵。

1. 系统误差

系统误差（即偏倚）是指结果或推断距离真实情况的偏离。系统误差可导致过高或者过低的估计真实干预措施效应^[214]。系统误差的风险和证据的强度取决于临床试验设计类型（系统综述，随机对照试验，观察性研究等）^[218, 219]。目前学术界广泛认可干预措施效应的证据等级划分（见表 3-1）^[22, 207, 214]。临床研究的方法学质量越好，研究类型所在等级就越高，其越可能得到接近真实的观察值，则存在系统误差的风险越低，证据强度越大^[218, 219]。本研究中，我们运用证据等级高低以评估证据的系统误差。此外，我们还运用 Cochrane 的 ‘risk of bias tool’ 以详细展示纳入的随机对照试验的系统误差。

表 3-1 根据系统误差（偏倚）对临床研究的证据等级划分^[22]

证据等级	研究类型
Level 1a	基于低偏倚风险的随机对照试验形成的 Meta 分析
Level 1b	低偏倚风险的随机对照试验
Level 1c	基于全部随机对照试验的 Meta 分析
Level 1d	高偏倚风险的随机对照试验
Level 2a	队列研究的 Meta 分析
Level 2b	队列研究
Level 3a	病例对照研究的 Meta 分析
Level 3b	病例对照研究
Level 4	病例系列研究

2. 随机误差

随机误差（即不精确性）是指多次重复相同的研究会导致不同的评估效应^[214]。临床研究中，随机误差通常由零散数据或者重复检验引起^[133]。小样本研究结果可常常被抽样错误影响，因此结果通常不够精确^[22, 214, 220]。我们运用标准误（SE）作为不确定性的测量指标以评估随机误差风险。标准误常用来表明样本均值与真实人群均值的接近程度。我们根据 Cochrane 协作组的统计分析方法进行评估^[221]，具体如下：

2.1 对于二分类变量

每项研究的风险比为：

$$RR_i = \frac{a_i / n_{1i}}{c_i / n_{2i}}$$

（其中 a 和 c 是发生事件的患者数目， b 和 d 是未发生事件的患者数目， n_{1i} 和 n_{2i} 分别是第 i 项研究干预组和对照组总患者数目）。

风险比 \log 值的标准误为：

$$SE\{\ln(RR_i)\} = \sqrt{\frac{1}{a_i} + \frac{1}{c_i} - \frac{1}{n_{1i}} - \frac{1}{n_{2i}}}$$

2.2 对于连续变量

均差为：

$$MD_i = m_{1i} - m_{2i}$$

（ m_{1i} 和 m_{2i} 为第 i 项研究干预组和对照组的均值）。

均差的标准误为：

$$SE\{MD_i\} = \sqrt{\frac{sd_{1i}^2}{n_{1i}} + \frac{sd_{2i}^2}{n_{2i}}}$$

（ sd_{1i} 和 sd_{2i} 为第 i 项研究干预组合对照组的标准偏差）。

在队列研究以及 meta 分析中，标准误为：

$$SE\{\hat{\theta}\} = \frac{\ln(UCL) - \ln(LCL)}{2 \times \phi(1 - \alpha/2)}$$

(其中 UCL 为干预组上限, LCL 为干预组下限, $\hat{\theta}$ 为 $\ln(RR_{MH})$ 或 Hazard Ratio, ϕ 为标准正态偏差)。

2.3 零数据处理

当零数据导致标准误差计算有问题时, 我们将该数据标记为 Z 而并不填补零数据。

2.4 随机误差评估

SE 为 0.0-0.1 时, 为可忽略的随机误差; SE 为 0.1-0.2 时, 为随机误差低风险; SE 为 0.2-0.3 时, 为中等风险的随机误差; SE 为 0.3-0.5 时, 为大量风险的随机误差; SE \geq 0.5 时为高风险的随机误差^[22]。

3. 设计误差

设计误差指设计要素的错误选择导致临床研究的外部真实性出现问题^[214]。设计误差可由不同的设计要素如 PICOTA (参与者、干预措施、对照、结局指标、时间、分析) (详见表 3-3) 引起^[22, 214, 222]。在众多需要考虑的设计要素中, 结局指标可作为临床设计的核心要素^[22, 223]。根据 GRADE 分级, 我们运用结局指标的优先级别 (即危急、重要、不重要) 来评估证据的设计误差 (表 6)。此外, 在改进后的误差矩阵中, 我们还详细展示可导致设计误差的其他设计要素的具体情况 (表 3-3)。

表 3-2 根据对乙肝患者重要程度对结局指标进行分级

重要程度	结局指标分类	结局指标
1	主要结局指标 (对于临床决策很重要)	全因死亡率 严重不良事件 健康相关生存质量
2	次要结局指标 (重要但是对于临床决策不重要)	乙肝相关死亡率 乙肝相关并发症 非严重不良事件
3	探索性结局指标 (对临床决策不重要对患者也不重要)	可检测到的 HBV-DNA 可检测到的 HBeAg

表 3-3 用于评估设计误差的变量

项目	内容	会导致对干预措施效应过高估计或估计不足的原因
1	结局指标测量	滥用替代结局指标 ^[205, 206] 和复合指标 ^[252, 253] , 或者未清楚

		定义主要结局指标和次要结局指标
2	参与者	诊断标准、纳入标准以及排除标准定义不清晰，或者参与者患有极重或极轻的疾病 ^[207]
3	试验组干预措施	干预措施的实施不清楚或者与临床实践不相符，或者未清楚定义的根据患者个体的化裁过程 ^[254]
4	对照组干预措施	干预措施的实施不清楚或者与临床实践不相符
5	试验中心或研究实施地点	单临床试验中心或者试验中心患者患有过重或者过轻的疾病 ^[207, 255]
6	研究目标-探索性试验或实践性试验	探索性试验结果很少应用于临床实践中 ^[229, 255]
7	试验结构-平行对照，交叉对照	不稳定的病情或者不可逆的干预措施，或者不统一的药理或者心理方面的结转效应，或者在交叉对照试验中干预反馈相互关联 ^[256, 257]
8	研究目的-优效性，等效性或劣效性	等效性研究或者非优效性研究的滥用 ^[258]
9	分析单元	数据分析单元与随机分配的不一致 ^[85]

4. 证据矩阵

我们通过三个误差维度构建可视化证据矩阵。首先，我们根据系统误差维度和随机误差维度对检索到的文献进行排序。其次，我们根据表 3-2 中的结局指标等级对研究进一步排序。我们用一个摩天大楼代表一项研究，摩天大楼的高度代表着研究结局指标的等级。为了进一步展示设计误差的其他构成要素，我们将摩天大楼平均分为 8 层，每一层代表一个设计误差要素。如果该要素导致设计误差风险较低，则该层将被点亮，否则，该层是黑暗的。根据表 3-3 列出的原因，如果某一设计要素会导致过高或过低的估计干

预效应，则该要素导致设计误差的风险较高。最后，我们通过构建以摩天大楼为基本元素的可视化三维误差证据矩阵以评估证据体。由于摩天大楼的高度代表着结局指标的优先级，坐落于误差矩阵西北角的最高的摩天大楼代表最佳证据。

5. 纳入研究标准

我们通过两个例子探索中医药干预措施证据体质量：即小柴胡汤干预慢性乙型肝炎及针灸疗法干预慢性乙型肝炎。通过 PICO 结构问题来阐释这两个例子及纳入标准：P（Participants，参与者）：依据研究者诊断为慢性乙型肝炎的参与者，无论性别和年龄；I（Intervention，干预措施）：对于小柴胡汤一例，小柴胡汤应至少包含四味主药即柴胡、半夏、人参、黄芩，对于剂量、急性和实施方案无限制；对于针灸一例，通过手针、穴位按压、穴位注射、电刺激、激光刺激、热灸或药物贴敷等方式进行的针灸干预，对于实施方式和疗程无限制；C（Control，对照）：对于小柴胡汤一例，对照组干预措施为安慰剂对照或无干预措施；对于针灸一例，对照组干预措施为假针灸或无干预措施；O（Outcome，结局指标）：如表 6 所示。

6. 纳入研究检索方式

通过建立检索式（附表 1、附表 2）进行如下数据库检索：英文数据库：Cochrane 肝胆病组对照试验注册库（The Cochrane Hepato-Biliary Group Controlled Trials Register (Gluud 2018)）、Cochrane 图书馆中的 Cochrane 对照试验注册中心（Cochrane Central Register of Controlled Trials (CENTRAL) in The Cochrane Library）、MEDLINE Ovid、Embase Ovid、LILACS Bireme (Latin American and Caribbean Health Science Information database)、Science Citation Index Expanded (Web of Science)、Conference Proceedings Citation Index-Science (Web of Science)；中文数据库：中国知网数据库（China National Knowledge Infrastructure Database, CNKI）、万方数据库、重庆维普中文科技期刊数据库（Chongqing VIP Chinese Science and Technology Periodical Database, VIP）、中国生物医学文献数据库（Chinese Biomedical Literature Database, CBM）。我们检索了相关系统综述、meta 分析以及符合纳入标准的文章的参考文献列表。我们还检索了世界卫生组织国际临床试验注册平台 World Health Organization International Clinical Trials Registry Platform (www.who.int/ictrp)、ClinicalTrial.gov (www.clinicaltrials.gov/) 以及中国临床试验注册中心（ChiCTR）以检索正在进行中的研究和未发表研究。

7. 数据收集

作者两人一组根据纳入研究标准独立进行文章筛选，当有意见分歧时，进行第三发咨询。在与研究作者联系已获得缺失或者不清晰的数据后，我们根据研究者报告提取最大随访期限的数据并以此进行数据分析。

研究结果

1. 小柴胡汤干预慢性乙型肝炎

根据表 3-1，由于纳入的两项 meta 分析研究无研究方案，并且研究中所纳入的随机对照试验存在高偏倚风险，两项 meta 分析研究的证据级别为 1c。Qin 2010^[224]具有可被忽略的随机误差风险（即 SE=0.0-0.1）。Hu 2011^[225]具有小风险的随机误差（即 SE=0.1-0.2）。两项 meta 分析研究显示在参与者^[224]、试验组干预措施^[224]、结局指标^[224, 225]、临床试验中心^[224, 225]、试验目标^[224, 225]以及试验目的^[224, 225]这些设计误差方面具有高风险。

52 项临床试验探索了小柴胡汤干预慢性乙型肝炎的获益和风险。我们评估了每项临床试验的系统误差（图 3-1、图 3-2）。图 3-1 展示了纳入的临床试验的系统误差。只有十项研究报告了运用电脑或者数字表产生随机序列。没有一项研究报告如何进行随机隐藏。所有纳入的研究除联合干预措施外，对照组均未施加任何干预措施，而小柴胡汤组则实施小柴胡汤干预，极易导致揭盲，因此在参与者施盲方面具有高偏倚风险。纳入研究由于缺少相关信息的报告，在结局指标评估施盲方面具有高偏倚风险。只有 Jia 1990^[226]及 Zhao 2014^[227]由于缺少部分参与者数据，因而在不完整结局数据方面表现高风险。所有研究由于缺少提前发表的研究方案以及未报告以患者为中心的结局指标数据如死亡率、研究不良事件、健康相关生活质量，因此存在选择性报告高风险^[214]。

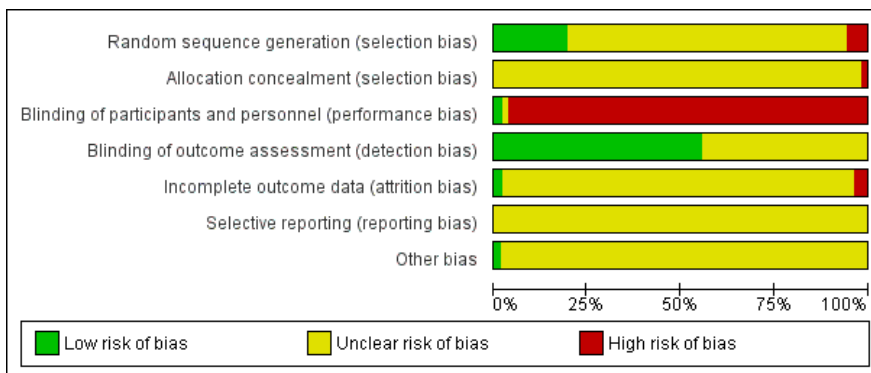


图 3-1 构建证据矩阵纳入研究的风险偏倚评估（百分比） - 以小柴胡汤干预慢性乙型肝炎为例

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Bo 2006	?	?	?	?	?	?	?
Chen 2008	?	?	?	?	?	?	?
Chen 2011	?	?	?	?	?	?	?
Chen 2013	?	?	?	?	?	?	?
Chen 2014	?	?	?	?	?	?	?
Chen 2017	?	?	?	?	?	?	?
Cheng 2015	?	?	?	?	?	?	?
He 2008	?	?	?	?	?	?	?
Hirayama 1989	?	?	?	?	?	?	?
Hong 2005	?	?	?	?	?	?	?
Hu 2017	?	?	?	?	?	?	?
Huang 2001	?	?	?	?	?	?	?
Jia 1990	?	?	?	?	?	?	?
Jiang 2015	?	?	?	?	?	?	?
Kang 2016	?	?	?	?	?	?	?
Li 2001a	?	?	?	?	?	?	?
Li 2001b	?	?	?	?	?	?	?
Li 2009	?	?	?	?	?	?	?
Li 2015	?	?	?	?	?	?	?
Lin 2017	?	?	?	?	?	?	?
Liu 2011	?	?	?	?	?	?	?
Liu 2017	?	?	?	?	?	?	?
Mao 2014	?	?	?	?	?	?	?
Qiu 2010	?	?	?	?	?	?	?
Sata 1994	?	?	?	?	?	?	?
Shan 2006	?	?	?	?	?	?	?
Shi 2006	?	?	?	?	?	?	?
Sun 2004	?	?	?	?	?	?	?
Sun 2005	?	?	?	?	?	?	?
Tajiri 1991	?	?	?	?	?	?	?
Tian 2010	?	?	?	?	?	?	?
Wang 1992	?	?	?	?	?	?	?
Wang 2009	?	?	?	?	?	?	?
Wang 2012	?	?	?	?	?	?	?
Wang 2013a	?	?	?	?	?	?	?
Wang 2013b	?	?	?	?	?	?	?
Wang 2014a	?	?	?	?	?	?	?
Wang 2014b	?	?	?	?	?	?	?
Wang 2016	?	?	?	?	?	?	?
Wei 2007	?	?	?	?	?	?	?
Wu 1994	?	?	?	?	?	?	?
Wu 2015	?	?	?	?	?	?	?
Xiong 2003	?	?	?	?	?	?	?
Yang 2009	?	?	?	?	?	?	?
Yu 2000	?	?	?	?	?	?	?
Zeng 2015	?	?	?	?	?	?	?
Zhang 2001	?	?	?	?	?	?	?
Zhang 2004	?	?	?	?	?	?	?
Zhao 2014	?	?	?	?	?	?	?
Zhou 2015a	?	?	?	?	?	?	?
Zhou 2015b	?	?	?	?	?	?	?
Zou 2016	?	?	?	?	?	?	?

图 3-2 构建证据矩阵纳入研究的风险偏倚评估（单个研究详情）
-以小柴胡汤干预慢性乙型肝炎为例

表 3-4 展示了纳入的临床证据的随机误差。以 HBeAg 阳性参与者结局指标为例，如图 3-3 所示，9 项随机对照试验表明高随机误差风险；7 项研究表明大量的随机误差风险；12 项研究表明中等程度的随机误差风险；12 项研究表明可忽略或较小的随机误差风险^[22]。

表 3-4 根据证据等级（系统误差）、标准误（随机误差）、结局指标（设计误差）排列纳入证据-以小柴胡汤干预慢性乙型肝炎为例

序号	研究 ID	证据等级	标准误（随机误差）							
			全因死亡率	严重不良事件	健康相关生存质量	乙肝相关死亡率	乙肝相关并发症	非严重不良事件	HBV-DNA 阳性检测率	HBeAg 阳性检测率
1	Hirayama 1989 ^[259]	1d	NA	NA	NA	NA	NA	1.22	NA	NA
2	Jia 1990 ^[226]	1d	NA	NA	NA	NA	NA	NA	1.11	0.54
3	Tajiri 1991 ^[260]	1d	NA	NA	NA	NA	NA	NA	NA	0.22
4	Wang 1992 ^[261]	1d	NA	NA	NA	NA	NA	NA	0.30	0.30
5	Sata 1994 ^[262]	1d	NA	NA	NA	NA	NA	0.06	0.31	0.21
6	Wu 1994 ^[263]	1d	NA	NA	NA	NA	NA	NA	0.39	0.40
7	Yu 2000 ^[264]	1d	NA	NA	NA	NA	NA	NA	0.20	0.14
8	Huang 2001 ^[265]	1d	NA	NA	NA	NA	NA	0.20	0.22	0.21
9	Li 2001a ^[266]	1d	NA	NA	NA	NA	NA	NA	NA	NA
10	Li 2001b ^[267]	1d	NA	NA	NA	NA	NA	NA	0.14	0.15
11	Zhang 2001 ^[268]	1d	NA	0.61	NA	NA	NA	0.45	NA	NA
12	Xiong 2000 ^[269]	1d	NA	NA	NA	NA	NA	NA	NA	NA
13	Sun 2004 ^[270]	1d	NA	NA	NA	NA	NA	NA	NA	NA
14	Zhang 2004 ^[271]	1d	NA	NA	NA	NA	NA	NA	0.13	0.18
15	Hong 2005 ^[272]	1d	NA	NA	NA	NA	NA	NA	0.09	0.08

16	Sun 2005 ^[273]	1d	NA	NA	NA	NA	NA	NA	NA	NA
17	Bo 2006 ^[274]	1d	NA	NA	NA	NA	NA	NA	0.21	0.20
18	Shan 2006 ^[275]	1d	NA	NA	NA	NA	NA	NA	0.21	NA
19	Shi 2006 ^[276]	1d	NA	NA	NA	NA	NA	NA	0.15	0.17
20	Wei 2007 ^[277]	1d	NA	NA	NA	NA	NA	NA	NA	NA
21	Chen 2008 ^[278]	1d	NA	NA	NA	NA	NA	Z	0.13	0.11
22	He 2008 ^[279]	1d	NA	NA	NA	NA	NA	NA	NA	NA
23	Li 2009 ^[280]	1d	NA	NA	NA	NA	NA	NA	0.52	0.22
24	Wang 2009 ^[281]	1d	NA	NA	NA	NA	NA	NA	NA	NA
25	Yang 2009 ^[282]	1d	NA	NA	NA	NA	NA	NA	NA	NA
26	Qin 2010 ^[224]	1c	NA	NA	NA	NA	NA	NA	0.07	0.07
27	Qiu 2010 ^[283]	1d	NA	NA	NA	NA	NA	NA	0.10	0.11
28	Tian 2010 ^[284]	1d	NA	NA	NA	NA	NA	NA	0.29	0.26
29	Chen 2011 ^[285]	1d	NA	NA	NA	NA	NA	Z	0.12	0.21
30	Hu 2011 ^[225]	1c	NA	NA	NA	NA	NA	0.12	0.10	0.12
31	Liu 2011 ^[286]	1d	NA	NA	NA	NA	NA	NA	0.23	0.27
32	Wang 2012 ^[287]	1d	NA	NA	NA	NA	NA	NA	NA	NA
33	Chen 2013 ^[288]	1d	NA	NA	NA	NA	NA	NA	0.51	0.18
34	Wang 2013a ^[289]	1d	NA	NA	NA	NA	NA	NA	0.37	0.41
35	Wang 2013b ^[290]	1d	NA	NA	NA	NA	NA	Z	NA	NA
36	Chen 2014 ^[291]	1d	NA	NA	NA	NA	NA	NA	NA	NA
37	Mao 2014 ^[292]	1d	NA	NA	NA	NA	NA	NA	NA	NA
38	Wang 2014a ^[293]	1d	NA	NA	NA	NA	NA	NA	NA	NA

39	Wang 2014b ^[294]	1d	NA	NA	NA	NA	NA	NA	0.28	0.35
40	Zhao 2014 ^[227]	1d	NA	NA	NA	NA	NA	NA	0.20	0.13
41	Cheng 2015 ^[295]	1d	NA	NA	NA	NA	NA	NA	0.41	0.52
42	Jiang 2015 ^[296]	1d	NA	NA	NA	NA	NA	Z	NA	NA
43	Li 2015 ^[297]	1d	NA	NA	NA	NA	NA	NA	0.72	NA
44	Wu 2015 ^[298]	1d	NA	NA	NA	NA	NA	NA	NA	NA
45	Zeng 2015 ^[299]	1d	NA	NA	NA	NA	NA	NA	NA	NA
46	Zhou 2015a ^[300]	1d	NA	NA	NA	NA	NA	0.59	0.30	0.25
47	Zhou 2015b ^[301]	1d	NA	NA	NA	NA	NA	NA	NA	NA
48	Kang 2016 ^[302]	1d	NA	NA	NA	NA	NA	NA	0.32	0.28
49	Wang 2016 ^[303]	1d	NA	NA	NA	NA	NA	NA	0.74	NA
50	Zou 2016 ^[304]	1d	NA	NA	NA	NA	NA	NA	0.30	0.30
51	Chen 2017 ^[305]	1d	NA	NA	NA	NA	NA	NA	NA	NA
52	Hu 2017 ^[306]	1d	NA	NA	NA	NA	NA	NA	0.23	NA
53	Lin 2017 ^[307]	1d	NA	NA	NA	NA	NA	NA	0.29	0.26
54	Liu 2017 ^[308]	1d	NA	NA	NA	NA	NA	1.20	NA	NA

表注: Z:结局指标结局为零事件, 无法计算 SE; NA: 无法获取数据;
 某些结局指标可能存在相关关系 (例.全因死亡率包含乙肝相关死亡率);
 本例中运用 $\ln RR_i$ 的标准误计算公式计算单个研究的标准误, 运用 $\ln RR_{MH}$ 的标准误计算公式计算 meta 分析的标准误。

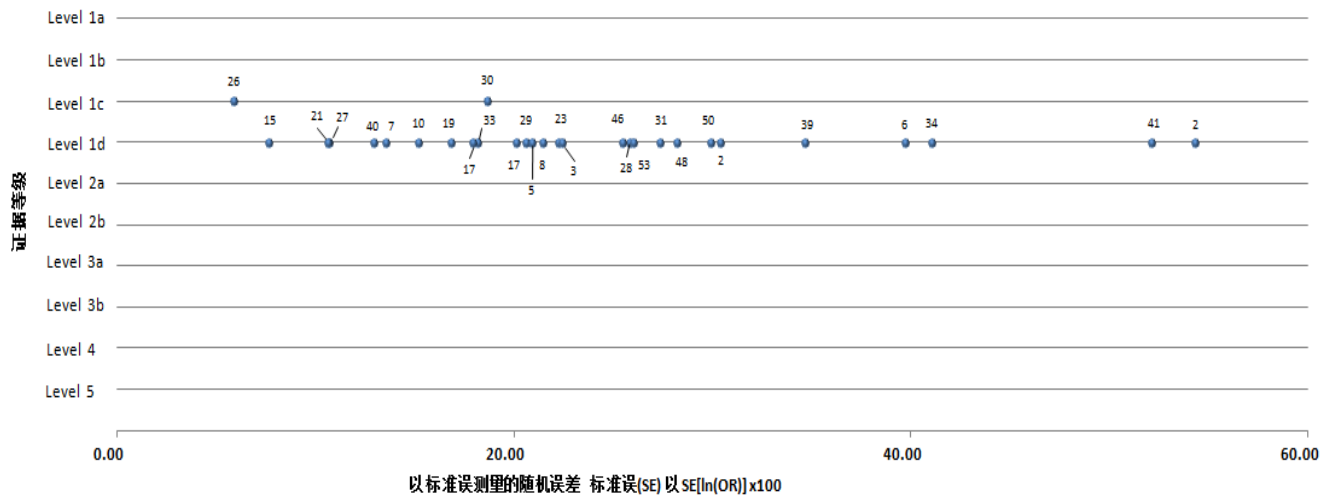


图 3-3 根据证据的系统误差和随机误差对证据进行排列-以小柴胡汤干预慢性乙型肝炎为例

如图 3-4 所示，所有研究均未报告关于死亡率、乙肝相关发病率或健康相关生活质量。只有 1 项研究^[228]报告了严重不良事件，只有 11 项研究报告了非严重不良事件。39 项研究只报告了替代结局指标。

证据等级	主要结局指标			次要结局指标			探索性结局指标	
	全因死亡率	严重不良事件	健康相关生存质量	乙肝相关死亡率	乙肝相关并发症	非严重不良事件	可检测的 HBV-DNA	可检测的 HBeAg
Level 1a								
Level 1b								
Level 1c							Qin 2010; Hu 2011	Qin 2010; Hu 2011
Level 1d		Zhang 2001				Hirayama 1989; Sata 1994; Huang 2001; Zhang 2001; Chen 2008; Chen 2011; Hu 2011; Wang 2013b; Jiang 2015; Zhou 2015a; Liu 2017	Jia 1990; Wang 1992; Sata 1994; Wu 1994; Yu 2000; Huang 2001; Li 2001b; Zhang 2004; Hong 2005; Bo 2006; Shan 2006; Shi 2006; Chen 2008; Li 2009; Qiu 2010; Tian 2010; Chen 2011; Liu 2011; Chen 2013; Wang 2013a; Wang 2014b; Cheng 2015; Li 2015; Zhou 2015a; Kang 2016; Wang 2016; Zou 2016; Hu 2017; Lin 2017	Jia 1990; Tajiri 1991; Wang 1992; Sata 1994; Wu 1994; Yu 2000; Huang 2001; Li 2001b; Zhang 2004; Hong 2005; Bo 2006; Shi 2006; Chen 2008; Li 2009; Qiu 2010; Tian 2010; Chen 2011; Liu 2011; Chen 2013; Wang 2013a; Wang 2014b; Zhao 2014; Cheng 2015; Zhou 2015a; Kang 2016; Zou 2016; Lin 2017
Level 2a								
Level 2b								
Level 3a								
Level 3b								
Level 4								
Level 5								

图 3-4 根据证据的系统误差和设计误差对证据进行排列-以小柴胡汤干预慢性乙型肝炎为例

表 3-5 展示了每项随机对照试验的设计误差详情。23 项随机对照试验设定了诊断标

准、纳入标准和排除标准。31 项随机对照试验（59.6%）清楚地报告了联合干预措施的详细实施方案。10 项随机对照试验（19%）为清楚地报告如何对小柴胡汤根据个体患者进行化裁。由于缺少研究方案，我们无法评估纳入的随机对照试验是优效性研究还是实用性研究^[229]。没有一项研究在 WHO 或中国临床注册中心（ChiCTR）进行过注册。所有纳入的研究均为平行对照设计，并以个体参与者为数据分析单位。只有 Tajiri1991^[230]和 Sata 1994^[231]报告他们的试验为多中心试验。

表 3-5 每项研究的设计误差详情-以小柴胡汤干预慢性乙型肝炎为例

序号	研究 ID	结局指标	参与者	试验组干预	对照组干预	临床中心	目标-探索性或实用性	研究结构	目标-优效性、等效性、非劣性	分析单元
1	Hirayama 1989	替代结局指标, 不良事件	诊断标准、纳入标准、排除标准清晰	干预措施清晰, 个体化裁清晰	定义清晰	多中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
2	Jia 1990	替代结局指标, 复合结局指标	无纳入标准、排除标准	联合干预措施不清楚	不清楚的联合干预措施	试验中心不清楚	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
3	Tajiri 1991	替代结局指标, 不良事件	无纳入标准、排除标准	干预措施清晰	定义清晰	多中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
4	Wang 1992	替代结局指标	无纳入标准、排除标准	小柴胡汤组成不清楚, 个体化裁不清楚	定义清晰	试验中心不清楚	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
5	Sata 1994	替代结局指标	诊断标准、纳入标准、排除标准清晰	干预措施清晰	定义清晰	多中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
6	Wu 1994	替代结局指标	无纳入标准、排除标准	小柴胡汤组成不清楚, 个体化裁不清楚	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
7	Yu 2000	替代结局指标	诊断标准、纳入标准	联合干预措施不清楚	不清楚的联合干预措施	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体

			标准、排除标准 清晰		合干预措施					
8	Huang 2001	不良事件, 替代结局指标	无纳入标准、排除标准	干预措施清晰	定义清晰	试验中心不清楚	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
9	Li 2001a	替代结局指标	无纳入标准、排除标准	小柴胡汤组成不清楚, 个体化裁不清楚	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
10	Li 2001b	替代结局指标	无排除标准	小柴胡汤组成不清楚, 个体化裁不清楚	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
11	Zhang 2001	不良事件, 替代结局指标	无纳入标准、排除标准	联合干预措施不清楚	不清楚的联合干预措施	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
12	Xiong 2003	不良事件, 替代结局指标	无纳入标准、排除标准	联合干预措施不清楚	不清楚的联合干预措施	试验中心不清楚	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
13	Sun 2004	替代结局指标	无纳入标准、排除标准	联合干预措施不清楚	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
14	Zhang 2004	替代结局指标	诊断标准、纳入标准、排除标准 清晰	不清楚的小柴胡汤组、个体化裁、联合干预措施	不清楚的联合干预措施	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
15	Hong 2005	不良事件, 替代结局指标	诊断标准、纳入标准、排除标准 清晰	联合干预措施不清楚	不清楚的联合干预措施	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
16	Sun 2005	复合结局指标	无诊断标准、纳入标准、排除标准	个体化裁清晰	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
17	Bo 2006	不良事件, 替代结局指标	无纳入标准、排除标准	小柴胡汤组成不清楚, 个体化裁不清楚	定义清晰	试验中心不清楚	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
18	Shan 2006	不良事件, 替代结局指标	无排除标准	干预措施清晰	定义清晰	试验中心不清楚	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体

		代结局指标				清楚					
19	Shi 2006	替代结局指标	无纳入标准、排除标准	干预措施清晰	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体	
20	Wei 2007	替代结局指标	诊断标准、纳入标准、排除标准清晰	联合干预措施不清楚	不清楚的联合干预措施	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体	
21	Chen 2008	替代结局指标	诊断标准、纳入标准、排除标准清晰	联合干预措施不清楚	不清楚的联合干预措施	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体	
22	He 2008	替代结局指标	无纳入标准、排除标准	联合干预措施不清楚	不清楚的联合干预措施	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体	
23	Li 2009	替代结局指标	诊断标准、纳入标准、排除标准清晰	干预措施清晰	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体	
24	Wang 2009	替代结局指标	无排除标准	小柴胡汤组成不清楚, 个体化裁不清楚	定义清晰	试验中心不清楚	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体	
25	Yang 2009	不良事件, 替代结局指标	无排除标准	联合干预措施不清楚	不清楚的联合干预措施	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体	
26	Qin 2010	不良事件, 替代结局指标	诊断标准、纳入标准、排除标准清晰	小柴胡汤组成不清楚	定义清晰	纳入的研究均未报告为多中心	纳入的研究均无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体	
27	Qiu 2010	替代结局指标	诊断标准、纳入标准、排除标准清晰	联合干预措施不清楚	不清楚的联合干预措施	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体	
28	Tian 2010	替代结局指标	无排除标准	干预措施清晰	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体	
29	Chen 2011	不良事件, 替	诊断标准、纳入	联合干预措施不清楚	不清楚的联	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体	

		代结局指标	标准、排除标准 清晰		合干预措施					
30	Hu 2011	不良事件, 替代结局指标	诊断标准、纳入标准、排除标准 清晰	干预措施清晰	定义清晰	纳入的研究均未报告为多中心	纳入的研究均无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
31	Liu 2011	替代结局指标	诊断标准、纳入标准、排除标准 清晰	对照组干预措施方案不清楚	对照组干预措施方案不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
32	Wang 2012	替代结局指标	诊断标准、纳入标准、排除标准 清晰	个体化裁清晰	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
33	Chen 2013	替代结局指标	无排除标准	小柴胡汤组成不清楚, 个体化裁不清楚	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
34	Wang 2013a	替代结局指标	诊断标准、纳入标准、排除标准 清晰	联合干预措施不清楚	不清楚的联合干预措施	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
35	Wang 2013b	不良事件, 替代结局指标	无诊断标准、纳入标准、排除标准	联合干预措施不清楚	不清楚的联合干预措施	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
36	Chen 2014	替代结局指标	诊断标准、纳入标准、排除标准 清晰	联合干预措施不清楚	不清楚的联合干预措施	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
37	Mao 2014	不良事件, 复合结局指标	诊断标准、纳入标准、排除标准 清晰	不清楚的小柴胡汤组、个体化裁、联合干预措施	不清楚的联合干预措施	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
38	Wang 2014a	替代结局指标,	无诊断标准、纳	干预措施清晰	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体

		复合结局指标	入标准、排除标准								
39	Wang 2014b	替代结局指标	无纳入标准、排除标准	干预措施清晰	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体	
40	Zhao 2014	不良事件, 替代结局指标	诊断标准、纳入标准、排除标准清晰	干预措施清晰	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体	
41	Cheng 2015	替代结局指标	诊断标准、纳入标准、排除标准清晰	干预措施清晰	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体	
42	Jiang 2015	不良事件, 替代结局指标, 复合结局指标	无诊断标准、纳入标准、排除标准	干预措施清晰	定义清晰	试验中心不清楚	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体	
43	Li 2015	替代结局指标	诊断标准、纳入标准、排除标准清晰	干预措施清晰	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体	
44	Wu 2015	替代结局指标	无诊断标准、纳入标准、排除标准	小柴胡汤组成不清楚, 个体化裁不清楚	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体	
45	Zeng 2015	替代结局指标	诊断标准、纳入标准、排除标准清晰	干预措施清晰	不清楚的联合干预措施	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体	
46	Zhou 2015a	不良事件, 替代结局指标	诊断标准、纳入标准、排除标准清晰	联合干预措施不清楚	不清楚的联合干预措施	试验中心不清楚	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体	
47	Zhou 2015b	复合结局指标	无诊断标准、纳入标准、排除标准	干预措施清晰	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体	

			入标准、排除标准							
48	Kang 2016	替代结局指标, 复合结局指标	无诊断标准、纳入标准	干预措施清晰	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
49	Wang 2016	复合结局指标	诊断标准、纳入标准、排除标准清晰	干预措施清晰	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
50	Zou 2016	替代结局指标, 复合结局指标	诊断标准、纳入标准、排除标准清晰	小柴胡汤组成不清楚	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
51	Chen 2017	替代结局指标	诊断标准、纳入标准、排除标准清晰	联合干预措施不清楚	不清楚的联合干预措施	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
52	Hu 2017	替代结局指标	无诊断标准	联合干预措施不清楚	不清楚的联合干预措施	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
53	Lin 2017	替代结局指标, 复合结局指标	无纳入标准、排除标准	干预措施清晰	定义清晰	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
54	Liu 2017	不良事件, 替代结局指标, 复合结局指标	诊断标准、纳入标准、排除标准清晰	联合干预措施不清楚	不清楚的联合干预措施	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体

我们为证据体构建证据矩阵（图 3-5,图 3-6）。所有纳入的证据被评估为内部真实性和外部真实性存在高偏倚风险。位于 1c 级别的最佳等级证据（Qin 2010）及（Hu 2011）分别具有可忽略的随机误差风险和小随机误差风险，但是具有高系统误差和设计误差风险。

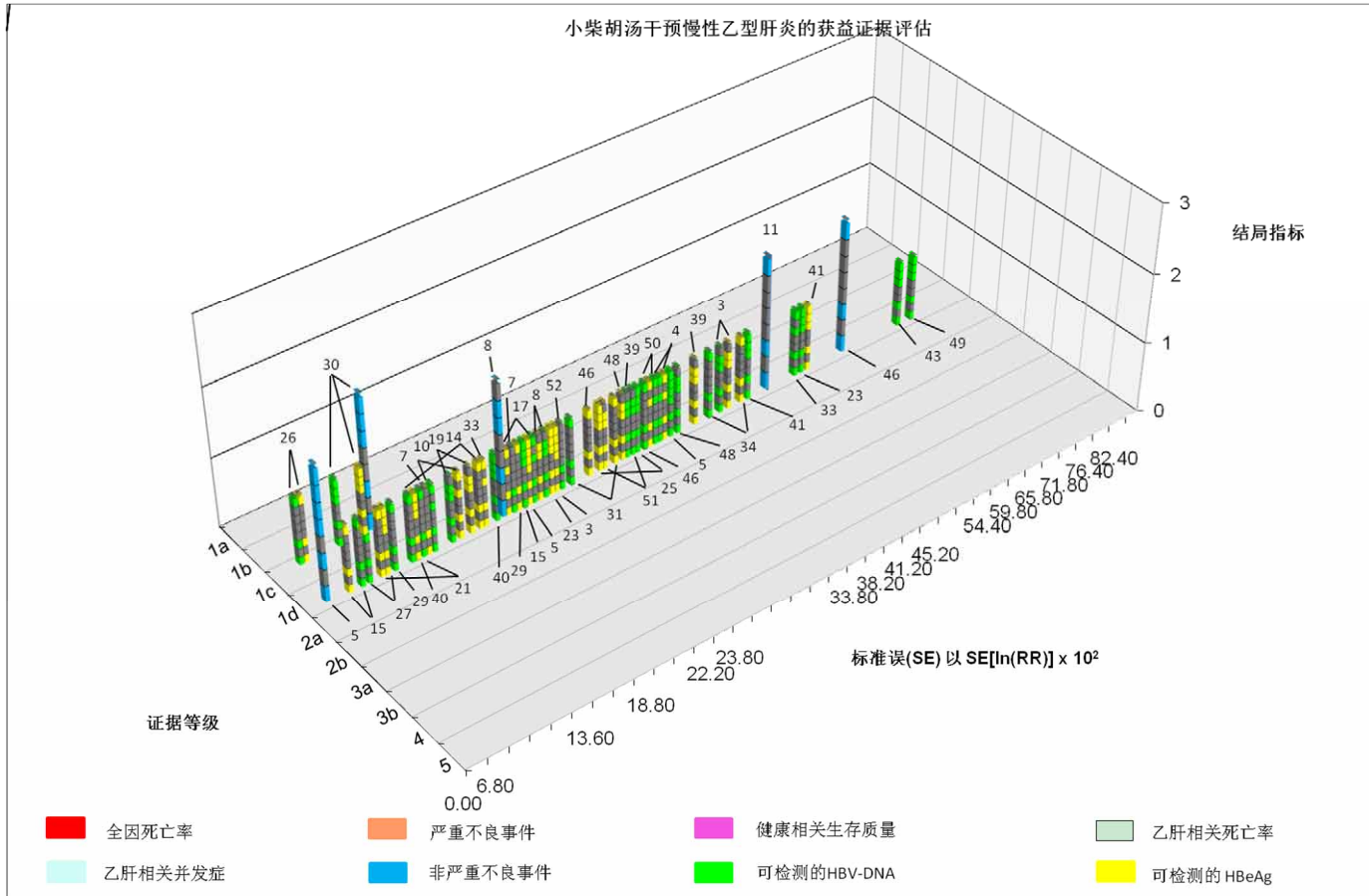


图 3-5 根据系统误差，随机误差和设计误差构建的三维误差证据矩阵-
小柴胡汤干预慢性乙型肝炎的获益

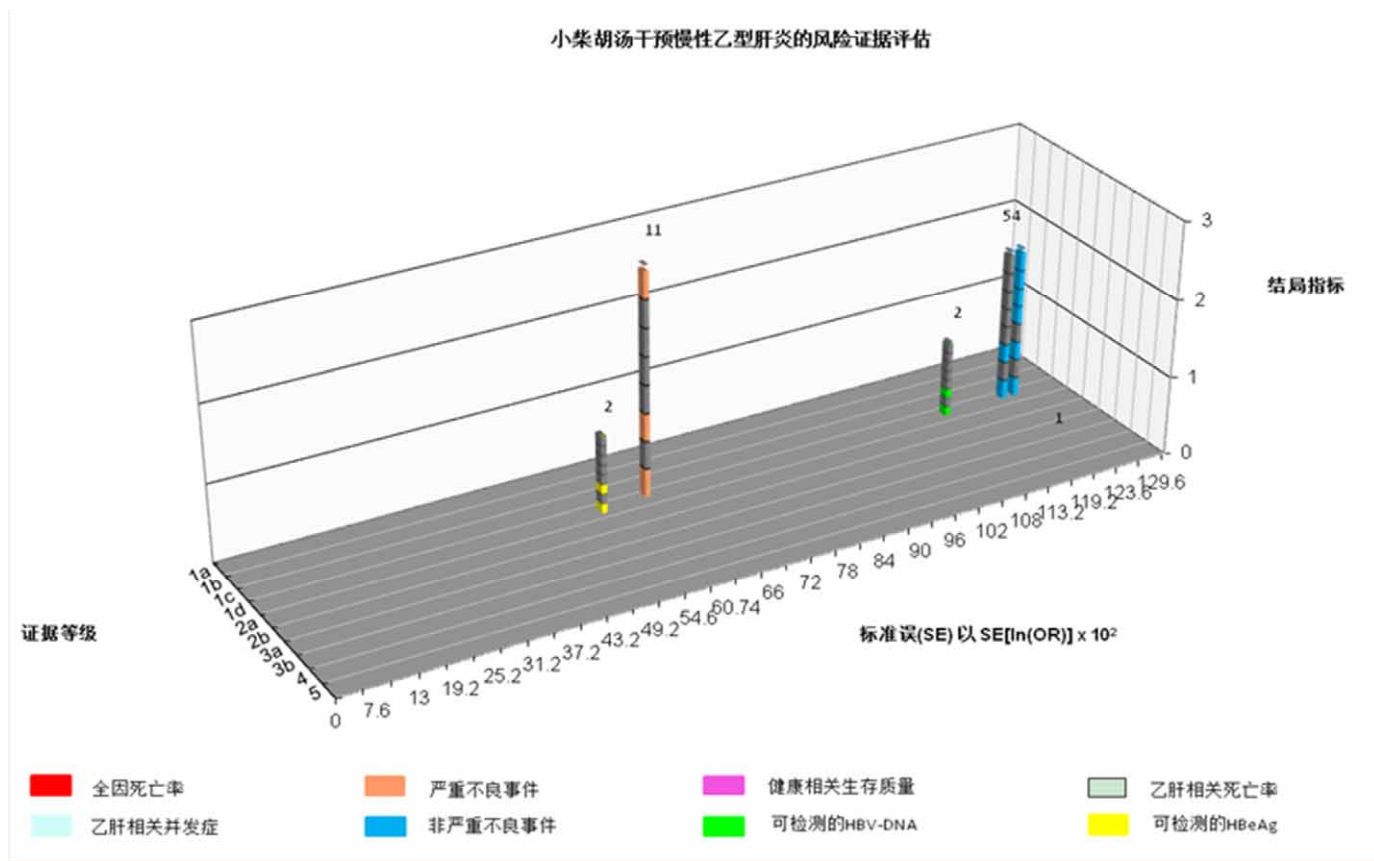


图 3-6 根据系统误差，随机误差和设计误差构建的三维误差证据矩阵-
小柴胡汤干预慢性乙型肝炎的风险

2. 针灸干预慢性乙型肝炎

基于低质量的临床随机对照试验的两项 meta 分析、77 项临床随机对照试验以及两项病例系列报告研究探索了针灸干预慢性乙型肝炎的有效性和安全性。图 3-7、图 3-8 展示了纳入的随机对照试验的系统误差详细情况。纳入的随机对照试验中，大约 13% (10/77) 通过电脑或者随机数字表生成随机序列。只有 Liao 2011^[232] 报告运用不透明的信封进行随机分配的隐藏，只有 Chen 2015a^[233] 报告对参与者进行施盲。纳入的随机对照试验中，大约 34% (26/77) 表现为结局指标评估的低偏倚风险。只有 Zhang 2016b^[234] 报告在分析中纳入全部的参与者数据。所有随机对照研究由于缺少提前发表的研究方案以及未报告以患者为中心的结局指标数据如死亡率、研究不良事件、健康相关生活质量，因此存在选择性报告高风险。

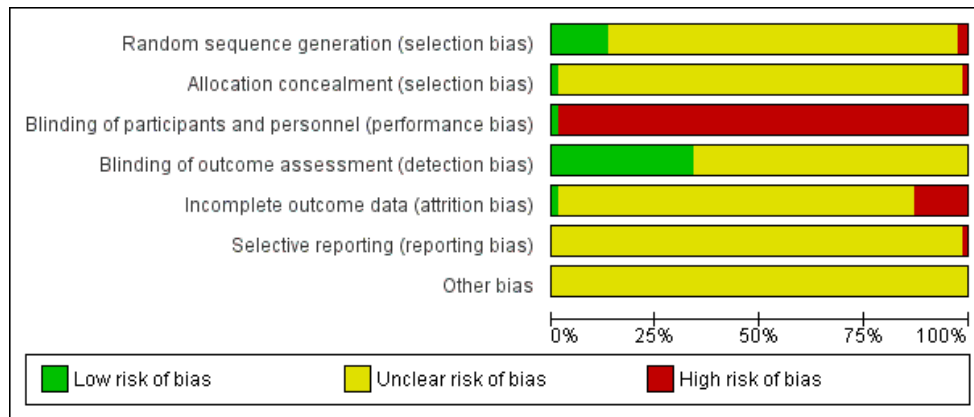


图 3-7 纳入研究的风险偏倚评估（百分比）-以针灸干预慢性乙型肝炎为例

Author	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Cai 2016	?	?	?	?	?	?	?
Chen 1998	?	?	?	?	?	?	?
Chen 2005	?	?	?	?	?	?	?
Chen 2014	?	?	?	?	?	?	?
Chen 2015a	?	?	?	?	?	?	?
Chen 2015b	?	?	?	?	?	?	?
Cheng 2000	?	?	?	?	?	?	?
Cheng 2004	?	?	?	?	?	?	?
Cheng 2016	?	?	?	?	?	?	?
Cheng 2009	?	?	?	?	?	?	?
Du 2004	?	?	?	?	?	?	?
Du 2005	?	?	?	?	?	?	?
Fang 2012	?	?	?	?	?	?	?
Han 2011	?	?	?	?	?	?	?
He 2013	?	?	?	?	?	?	?
Hu 2014	?	?	?	?	?	?	?
Hu 2015	?	?	?	?	?	?	?
Huang 2005	?	?	?	?	?	?	?
Huang 2011	?	?	?	?	?	?	?
Ji 1989	?	?	?	?	?	?	?
Jing 2016	?	?	?	?	?	?	?
Lei 2014	?	?	?	?	?	?	?
Li 2000	?	?	?	?	?	?	?
Li 2003	?	?	?	?	?	?	?
Li 2007	?	?	?	?	?	?	?
Li 2011a	?	?	?	?	?	?	?
Li 2011b	?	?	?	?	?	?	?
Li 2014	?	?	?	?	?	?	?
Li 2015a	?	?	?	?	?	?	?
Li 2015b	?	?	?	?	?	?	?
Li 1989	?	?	?	?	?	?	?
Li 2003	?	?	?	?	?	?	?
Li 2006	?	?	?	?	?	?	?
Li 2011	?	?	?	?	?	?	?
Li 2015	?	?	?	?	?	?	?
Li 2011	?	?	?	?	?	?	?
Li 2016a	?	?	?	?	?	?	?
Li 2016b	?	?	?	?	?	?	?
Ma 2001	?	?	?	?	?	?	?
Miao 2007	?	?	?	?	?	?	?
Peng 2002	?	?	?	?	?	?	?
Peng 2003	?	?	?	?	?	?	?
Ruan 2009	?	?	?	?	?	?	?
Shao 2009	?	?	?	?	?	?	?
Shen 2004	?	?	?	?	?	?	?
Sun 2000	?	?	?	?	?	?	?
Sun 2008	?	?	?	?	?	?	?
Tan 2009	?	?	?	?	?	?	?
Tang 2001	?	?	?	?	?	?	?
Tang 2014	?	?	?	?	?	?	?
Wang 2002	?	?	?	?	?	?	?
Wang 2003	?	?	?	?	?	?	?
Wang 2013	?	?	?	?	?	?	?
Wang 2014	?	?	?	?	?	?	?
Wang 2015b	?	?	?	?	?	?	?
Wang 2017	?	?	?	?	?	?	?
Wei 2014	?	?	?	?	?	?	?
Wu 2012	?	?	?	?	?	?	?
Xue 2007	?	?	?	?	?	?	?
Xue 2008	?	?	?	?	?	?	?
Yang 2002	?	?	?	?	?	?	?
Yang 2011	?	?	?	?	?	?	?
Zhang 2005	?	?	?	?	?	?	?
Zhang 2006	?	?	?	?	?	?	?
Zhang 2007	?	?	?	?	?	?	?
Zhang 2010	?	?	?	?	?	?	?
Zhang 2012	?	?	?	?	?	?	?
Zhang 2014a	?	?	?	?	?	?	?
Zhang 2014b	?	?	?	?	?	?	?
Zhang 2015	?	?	?	?	?	?	?
Zhang 2016a	?	?	?	?	?	?	?
Zhang 2016b	?	?	?	?	?	?	?
Zhou 2008	?	?	?	?	?	?	?
Zhu 2011	?	?	?	?	?	?	?
Zhu 2016	?	?	?	?	?	?	?
Zhuo 2012	?	?	?	?	?	?	?
Zou 2011	?	?	?	?	?	?	?

图 3-8 纳入研究的风险偏倚评估（单个研究详情）-以针灸干预慢性乙型肝炎为例

如表 3-6 所示，纳入证据的 22%（18/81）表明存在随机误差的高风险（即证据的结局指标 SE 值大于等于 0.5）。图 3-9 以可检测的 HBeAg 阳性参与者数目为例，根据证据的系统误差和随机误差对证据进行排列，以进一步构建三维误差矩阵。

表 3-6 根据证据等级（系统误差）、标准误（随机误差）、结局指标（设计误差）排列纳入证据-以针灸干预慢性乙型肝炎为例

序号	研究 ID	证据等级	标准误（随机误差）							
			全因死亡率	严重不良事件	健康相关生存质量	乙肝相关死亡率	乙肝相关并发症	非严重不良事件	HBV-DNA 阳性检测率	HBeAg 阳性检测率
1	Xu 1996 ^[309]	4	NA	NA	NA	NA	NA	NA	NA	NA
2	Chen 1998 ^[310]	1d	NA	NA	NA	NA	NA	NA	NA	NA
3	Ji 1999 ^[311]	1d	NA	NA	NA	NA	NA	0.58	NA	NA
4	Liu 1999 ^[236]	1d	NA	NA	NA	NA	NA	NA	0.14	0.15
5	Cheng 2000 ^[312]	1d	NA	NA	NA	NA	NA	NA	NA	NA
6	Li 2000 ^[313]	1d	NA	NA	NA	NA	NA	NA	NA	NA
7	Sun 2000 ^[314]	1d	NA	NA	NA	NA	NA	NA	NA	0.14
8	Ma 2001 ^[315]	1d	ND	ND	ND	ND	ND	ND	ND	ND
9	Tang 2001 ^[316]	1d	NA	NA	NA	NA	NA	NA	NA	NA
10	Peng 2002 ^[317]	1d	NA	NA	NA	NA	NA	NA	0.17	0.16
11	Wang 2002 ^[318]	1d	NA	NA	NA	NA	NA	NA	0.26	0.28
12	Yang 2002 ^[319]	1d	NA	NA	NA	NA	NA	NA	0.1	0.08
13	Li 2003 ^[320]	1d	NA	1.1	NA	NA	NA	0.85	0.15	0.16
14	Liu 2003 ^[321]	1d	NA	NA	NA	NA	NA	NA	0.15	NA
15	Peng 2003 ^[237]	1d	NA	NA	NA	NA	NA	0.5	0.36	0.48

16	Wang 2003 ^[322]	1d	NA	NA	NA	NA	NA	NA	NA	NA
17	Cheng 2004 ^[323]	1d	NA	NA	NA	NA	NA	NA	NA	NA
18	Du 2004 ^[324]	1d	0.17	NA	NA	NA	NA	NA	0.4	0.38
19	Shen 2004 ^[325]	1d	NA	NA	NA	NA	NA	NA	0.21	0.25
20	Chen 2005 ^[326]	1d	NA	NA	NA	NA	NA	NA	NA	NA
21	Du 2005 ^[327]	1d	NA	NA	NA	NA	NA	NA	NA	NA
22	Huang 2005 ^[328]	1d	NA	NA	NA	NA	NA	NA	NA	0.06
23	Zhang 2005 ^[329]	1d	NA	0.36	NA	NA	NA	0.37	NA	NA
24	Liu 2006 ^[330]	1d	NA	NA	NA	NA	NA	NA	0.05	0.05
25	Zhang 2006 ^[235]	1d	0.89	NA	NA	0.89	NA	NA	0.12	0.13
26	Li 2007 ^[331]	1d	NA	NA	NA	NA	NA	NA	NA	NA
27	Miao 2007 ^[332]	1d	NA	NA	NA	NA	NA	NA	0.12	NA
28	Xue 2007 ^[333]	1d	NA	NA	NA	NA	NA	NA	NA	NA
29	Zhang 2007 ^[334]	1d	NA	NA	NA	NA	NA	NA	NA	NA
30	Sun 2008 ^[335]	1d	NA	Z	NA	NA	NA	NA	NA	NA
31	Xue 2008 ^[336]	1d	NA	NA	NA	NA	NA	NA	0.2	0.18
32	Zhou 2008 ^[337]	1d	NA	NA	NA	NA	NA	0.6	NA	NA
33	Deng 2009 ^[338]	1d	NA	NA	NA	NA	NA	NA	NA	NA
34	Ruan 2009 ^[339]	1d	NA	0.6	NA	NA	NA	0.89	0.07	0.06
35	Shao 2009 ^[340]	1d	NA	NA	NA	NA	NA	NA	0.14	0.14
36	Tan 2009 ^[341]	1d	NA	NA	NA	NA	NA	0.5	NA	NA
37	Zhang 2010 ^[342]	1d	NA	NA	NA	NA	NA	NA	NT	NT
38	Han 2011 ^[343]	1d	NA	0.5	NA	NA	NA	0.54	0.08	0.1
39	Huang 2011 ^[344]	1d	NA	NA	NA	NA	NA	NA	NA	NA

40	Li 2011a ^[238]	1d	NA	0.4	NA	NA	NA	0.37	NA	NA
41	Li 2011b ^[345]	1d	NA	NA	NA	NA	NA	NA	NA	NA
42	Liao 2011 ^[232]	1d	NA	Z	0.19	NA	NA	Z	NA	NA
43	Liu 2011 ^[346]	1d	NA	NA	0.07	NA	NA	NA	NA	NA
44	Lu 2011 ^[347]	1d	NA	Z	0.02	NA	NA	Z	NA	NA
45	Yang 2011 ^[348]	1d	NA	NA	NA	NA	NA	0.65	0.20	0.05
46	Zhu 2011 ^[349]	1d	NA	Z	NA	NA	Z	Z	NA	0.18
47	Zou 2011 ^[350]	1d	NA	NA	NA	NA	Z	Z	NA	0.04
48	Fang 2012 ^[351]	1d	NA	NA	NA	NA	NA	NA	NA	NA
49	Han 2012 ^[352]	4	NA	NA	NA	NA	NA	NA	NA	NA
50	Wu 2012 ^[353]	1d	NA	NA	NA	NA	NA	NA	NT	NT
51	Zhang 2012 ^[354]	1d	NA	NA	NA	NA	NA	NA	NA	NA
52	Zhuo 2012 ^[355]	1d	NA	NA	NA	NA	NA	NA	0.4	0.08
53	He 2013 ^[356]	1d	NA	NA	NA	NA	NA	NA	0.22	0.21
54	Wang 2013 ^[357]	1d	NA	NA	NA	NA	NA	0.42	NA	NA
55	Chen 2014 ^[358]	1d	NA	Z	0.33	NA	NA	0.7	0.37	NA
56	Hu 2014 ^[359]	1d	NA	NA	NA	NA	NA	NA	0.5	0.13
57	Lei 2014 ^[360]	1d	NA	Z	NA	NA	NA	0.71	NA	NA
58	Li 2014 ^[361]	1d	NA	Z	NA	NA	NA	0.57	NA	NA
59	Tang 2014 ^[362]	1d	NA	Z	NA	NA	Z	0.71	NA	NA
60	Wang 2014 ^[363]	1d	NA	0.56	NA	NA	NA	0.62	0.55	0.11
61	Wei 2014 ^[364]	1d	NA	NA	0.16	NA	NA	0.57	0.23	0.08
62	Zhang 2014a ^[365]	1d	NA	NA	0.08	NA	NA	NA	NA	NA

63	Zhang 2014b ^[366]	1d	NA	NA	NA	NA	NA	NA	NA	NA
64	Chen 2015a ^[233]	1d	NA	Z	0.07	NA	NA	Z	NA	NA
65	Chen 2015b ^[367]	1d	NA	NA	NA	NA	NA	NA	Z	NA
66	Hu 2015 ^[368]	1d	NA	NA	NA	NA	NA	0.75	NA	NA
67	Li 2015a ^[369]	1d	NA	NA	NA	NA	NA	NA	NA	NA
68	Liu 2015 ^[370]	1d	NA	NA	NA	NA	NA	0.37	0.25	NA
69	Wang 2015a ^[371]	1c	NA	NA	NA	NA	NA	NA	NA	NA
70	Wang 2015b ^[372]	1d	NA	NA	0.03	NA	NA	NA	NA	NA
71	Zhang 2015 ^[373]	1d	NA	NA	NA	NA	NA	NA	NA	NA
72	Cai 2016 ^[374]	1d	NA	NA	NA	NA	0.5	0.5	NA	NA
73	Cheng 2016 ^[375]	1d	NA	NA	NA	NA	NA	NA	NA	NA
74	Jing 2016 ^[376]	1d	NA	NA	NA	NA	NA	NA	NA	NA
75	Lu 2016a ^[377]	1d	NA	NA	NA	NA	1.06	1.06	NA	NA
76	Lu 2016b ^[378]	1d	NA	NA	NA	NA	NA	Z	NA	NA
77	Zhang 2016a ^[379]	1d	NA	z	0.15	NA	NA	z	NA	NA
78	Zhang 2016b ^[234]	1d	NA	NA	NA	NA	NA	NA	NA	NA
79	Zhu 2016 ^[380]	1d	NA	NA	NA	NA	0.43	0.43	NA	NA
80	Wang 2017 ^[381]	1d	NA	Z	NA	NA	NA	0.27	NA	NA
81	Wu 2017 ^[382]	1c	NA	NA	NA	NA	NA	NA	NA	NA

表 Z:结局指标结局为零事件，无法计算 SE； NA: 无法获取数据；

注： 某些结局指标可能存在相关关系 (例.全因死亡率包含乙肝相关死亡率)；

本例中运用 $\ln RR_i$ 的标准误计算公式计算单个研究的标准误，运用 $\ln RR_{MH}$ 的标准误计算公式计算 meta 分析的标准误。

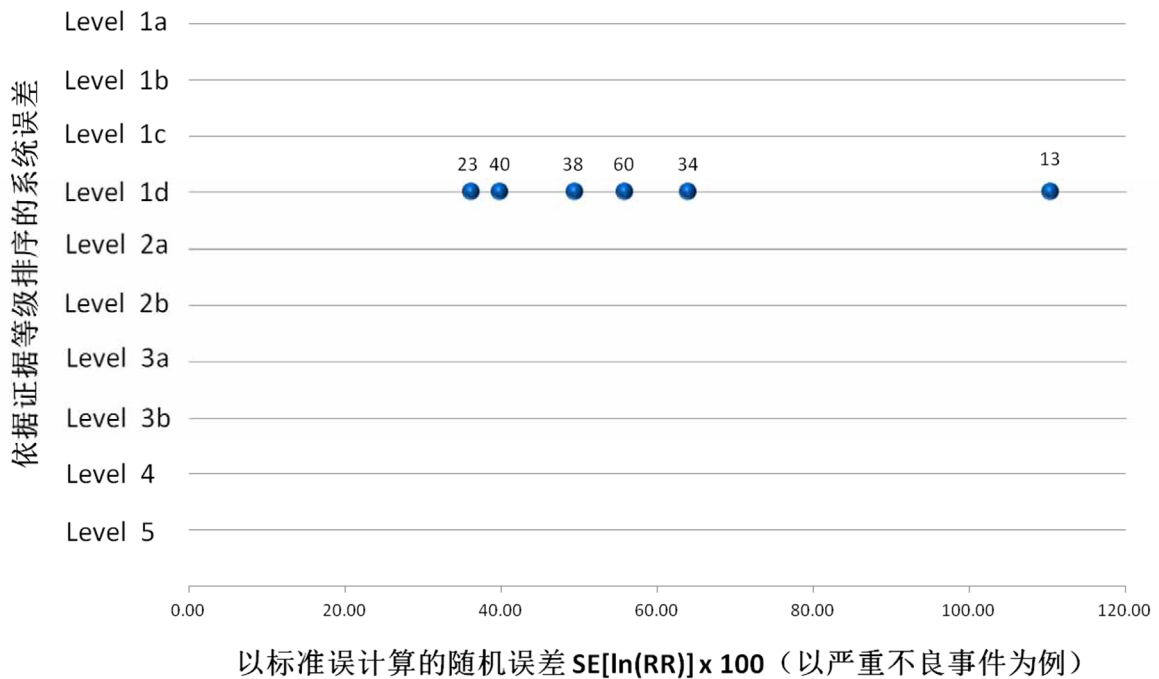


图 3-9 根据证据的系统误差和随机误差对证据进行排列-
以针灸干预慢性乙型肝炎为例

如图 3-10 所示，只有 Zhang 2006^[235] 报告了全因死亡率以及乙肝相关并发症。大约 19% (15/81) 的研究报告了严重不良事件，36% (29/81) 的研究报告了非严重不良事件，12% (10/81) 的研究报告了健康相关生活质量，7% (6/81) 的研究报告了乙肝相关并发症。

证据等级	主要结局指标			次要结局指标			探索性结局指标	
	全因死亡率	严重不良事件	健康相关结局指标	乙肝相关死亡率	乙肝相关并发症	非严重不良事件	可检测的HBV-DNA	可检测的HBeAg
Level 1a								
Level 1b								
Level 1c								
Level 1d	Zhang 2006	Li 2003; Ruan 2009; Li 2011a; Liao 2011; Lu 2011; Zhu 2011; Chen 2014; Lei 2014; Li 2014; Tang 2014; Wang 2014; Chen 2015a; Zhang 2016a; Wang 2017	Han 2011; Liao 2011; Liu 2011; Lu 2011; Chen 2014; Wei 2014; Zhang 2014a; Chen 2015a; Wang 2015b; Zhang 2016a	Zhang 2006	Zhu 2011; Zou 2011; Tang 2014; Cai 2016; Lu 2016a; Zhu 2016	Ji 1999; Li 2003; Peng 2003; Zhang 2005; Ruan 2009; Han 2011; Li 2011a; Liao 2011; Lu 2011; Yang 2011; Zhu 2011; Zou 2011; Wang 2013; Chen 2014; Lei 2014; Li 2014; Tang 2014; Wang 2014; Wei 2014; Chen 2015a; Hu 2015; Liu 2015; Cai 2016; Lu 2016a; Lu 2016b; Zhu 2016; Zhang 2016a; Wang 2017	Liu 1999; Lan 2001; Peng 2002; Wang 2002; Yang 2002; Li 2003; Liu 2003; Peng 2003; Du 2004; Shen 2004; Liu 2006; Zhang 2006; Xue 2008; Zhou 2008; Ruan 2009; Shao 2009; Han 2011; Yang 2011; Zhuo 2012; He 2013; Chen 2014; Hu 2014; Wang 2014; Wei 2014; Chen 2015b; Liu 2015	Liu 1999; Sun 2000; Lan 2001; Peng 2002; Wang 2002; Yang 2002; Li 2003; Peng 2003; Du 2004; Shen 2004; Huang 2005; Liu 2006; Zhang 2006; Xue 2008; Ruan 2009; Shao 2009; Han 2011; Yang 2011; Zou 2011; Zhu 2011; Liu 2011; Zhuo 2012; He 2013; Hu 2014; Wei 2014; Miao 2007
Level 2a								
Level 2b								
Level 3a								
Level 3b								
Level 4								
Level 5								

图 3-10 根据证据的系统误差和设计误差对证据进行排列-以针灸干预慢性乙型肝炎为例

表 3-7 展示了每项研究的设计误差详情。58% (47/81) 的随机对照试验设定了诊断标准、纳入标准和排除标准。44% (36/81) 的随机对照试验清楚地报告了联合干预措施的详细实施方案。对于非药物干预措施的研究, 3 项研究 (3/33) 清楚地报告如何根据个体患者进行化裁治疗方案。所有研究均未报告提供医疗措施人员的明确资格标准。由于缺少研究方案, 我们无法评估纳入的随机对照试验是优效性研究还是实用性研究^[229]。没有研究曾在 WHO 或中国临床注册中心 (ChiCTR) 进行过注册。所有纳入的研究均为平行对照设计, 并以个体参与者为数据分析单位。只有 Liu 1999^[236], Peng 2003^[237], 和 Li 2011a^[238] 报告他们的试验包含两个试验中心。

表 3-7 每项研究的设计误差详情-以针灸干预慢性乙型肝炎为例

序号	研究 ID	干预类别	结局指标	参与者	试验组干预	对照组干预	临床中心	目标-探索性或实用性	研究结构	目标-优效性、等效性、非劣性	分析单元
1	Xu 1996	穴位按压	替代指标, 复合指标	无纳入标准、排除标准, 干预措施实施者资格不清楚	不清楚的个体化裁	无对照组	单中心	无法满足实用性试验标准	无对照组	无试验方案, 难以判断	个体
2	Chen 1998	艾灸	替代指标	干预措施实施者资格不清楚	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
3	Ji 1999	针刺	不良事件, 替代指标	无纳入标准、排除标准, 干预措施实施者资格不清楚	不清楚的个体化裁	定义清楚	单中心	无法满足实用性试验标准	无对照组	无试验方案, 难以判断	个体
4	Liu 1999	穴位注射	不良事件, 替代指标, 复合指标	无排除标准	清楚的干预措施	定义清楚	双中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
5	Cheng 2000	针刺	不良事件, 替代指标	无诊断标准、纳入标准、排除标准, 干预措施实施者资格不清楚	联合干预措施不清楚, 不清楚的个体化裁	联合干预措施不清楚	单中心	无法满足实用性试验标准	无对照组	无试验方案, 难以判断	个体
6	Li 2000	穴位注射	替代指标	无诊断标准、纳入标准、排除标准	穴位注射方案不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体

				除标准		施不清楚		性试验标准		难以判断	
7	Sun 2000	穴位注射	替代指标	无排除标准	清楚的干预措施	定义清楚	单中心	无法满足实用性试验标准	无对照组	无试验方案, 难以判断	个体
8	Ma 2001	电针, 穴位按压	复合指标	无纳入标准、排除标准	清楚的个体化裁	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
9	Tang 2001	电针	不良事件, 替代指标	无纳入标准、排除标准, 干预措施实施者资格不清楚	联合干预措施不清楚, 不清楚的个体化裁	联合干预措施不清楚	单中心	无法满足实用性试验标准	无对照组	无试验方案, 难以判断	个体
10	Peng 2002	艾灸	替代指标	无纳入标准、排除标准, 干预措施实施者资格不清楚	联合干预措施不清楚, 不清楚的个体化裁	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
11	Wang 2002	针刺	不良事件, 替代指标	干预措施实施者资格不清楚	联合干预措施不清楚, 不清楚的个体化裁	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
12	Yang 2002	穴位注射	替代指标	无纳入标准、排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	无对照组	无试验方案, 难以判断	个体
13	Li 2003	穴位注射	不良事件, 替代指标	无纳入标准、排除标准	清楚的干预措施	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体 12
14	Liu 2003	艾灸	替代指标	干预措施实施者资格不清楚	联合干预措施不清楚, 不清楚的个体化裁	联合干预措施不清楚	临床中心 不清楚	无法满足实用性试验标准	无对照组	无试验方案, 难以判断	个体
15	Peng 2003	针刺	不良事件, 替代指标	干预措施实施者资格不清楚	不清楚的个体化裁	定义清楚	双中心	无法满足实用性试验标准	无对照组	无试验方案, 难以判断	个体
16	Wang 2003	穴位注射	替代指标	清楚的诊断标准、纳入标准、排除标准	清楚的干预措施	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
17	Cheng 2004	艾灸	不良事件, 替代指标, 复合指标	无排除标准, 干预措施实施者资格不清楚	联合干预措施不清楚, 不清楚的个体化裁	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体

18	Du 2004	穴位注射	全因死亡率, 替代指标	无纳入标准、排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
19	Shen 2004	穴位注射	替代指标	清楚的诊断标准、纳入标准、排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
20	Chen 2005	针刺, 艾灸	替代指标	干预措施实施者资格不清楚	清楚的个体化裁	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
21	Du 2005	穴位注射	替代指标	清楚的诊断标准、纳入标准、排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
22	Huang 2005	穴位贴敷	替代指标	无纳入标准、排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
23	Zhang 2005	穴位注射	不良事件, 替代指标	清楚的诊断标准、纳入标准、排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
24	Liu 2006	针刺	替代指标	无诊断标准、纳入标准、排除标准, 干预措施实施者资格不清楚	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
25	Zhang 2006	穴位注射	替代指标	清楚的诊断标准、纳入标准、排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
26	Li 2007	穴位注射	不良事件, 替代指标	无排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
27	Miao 2007	穴位注射	替代指标	清楚的诊断标准、纳入标准、排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	无对照组	无试验方案, 难以判断	个体
28	Xue 2007	穴位注射	替代指标	无纳入标准、排除标准	联合干预措施不清楚	联合干预措施不清楚	不清楚临	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体

29	Zhang 2007	穴位注射	替代指标	清楚的诊断标准、纳入标准、排除标准	联合干预措施不清楚	联合干预措施不清楚	床中心	性试验标准		难以判断	
30	Sun 2008	穴位贴敷	不良事件, 替代指标, 复合指标	清楚的诊断标准、纳入标准、排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
31	Xue 2008	穴位注射	替代指标	无纳入标准、排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	无对照组	无试验方案, 难以判断	个体
32	Zhou 2008	穴位贴敷	不良事件, 替代指标	无纳入标准、排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
33	Deng 2009	穴位注射	不良事件, 替代指标	无排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
34	Ruan 2009	穴位注射	不良事件, 替代指标	清楚的诊断标准、纳入标准、排除标准	清楚的干预措施	定义清楚	不清楚 床中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
35	Shao 2009	针刺	替代指标	干预措施实施者资格不清楚	不清楚的个体化裁	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
36	Tan 2009	穴位注射	不良事件, 替代指标, 复合指标	清楚的诊断标准、纳入标准、排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
37	Zhang 2010	穴位注射	替代指标	无诊断标准、纳入标准、排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
38	Han 2011	艾灸	不良事件, 替代指标	干预措施实施者资格不清楚	联合干预措施不清楚, 不清楚的个体化裁	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
39	Huang 2011	穴位注射	替代指标	清楚的诊断标准、纳入标准、排除标准	清楚的干预措施	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体

			复合指标	排除标准						性试验标准	难以判断
40	Li 2011a	艾灸	不良事件, 替代指标	无纳入标准、排除标准, 干预措施实施者资格不清楚	不清楚的个体化裁	定义清楚	双中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
41	Li 2011b	穴位贴敷	不良事件, 替代指标	无排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
42	Liao 2011	穴位贴敷	生存质量, 不良事件, 替代指标	清楚的诊断标准、纳入标准、排除标准	清楚的干预措施	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
43	Liu 2011	穴位贴敷	生存质量, 替代指标	清楚的诊断标准、纳入标准、排除标准	清楚的干预措施	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
44	Lu 2011	穴位贴敷	生存质量, 不良事件, 复合指标	清楚的诊断标准、纳入标准、排除标准	定义清楚	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
45	Yang 2011	穴位注射	不良事件, 替代指标	无排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
46	Zhu 2011	穴位注射	替代指标	清楚的诊断标准、纳入标准、排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
47	Zou 2011	穴位贴敷	不良事件, 替代指标	清楚的诊断标准、纳入标准、排除标准	不清楚的个体化裁	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
48	Fang 2012	艾灸, 穴位贴敷	替代指标	无纳入标准、排除标准, 干预措施实施者资格不清楚	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案, 难以判断	个体
49	Han 2012	艾灸	不良事件	无纳入标准、排除标准, 干预措施实施者资格不清楚	No experimental group	无对照组	单中心	观察性研究	无对照组	无试验方案, 难以判断	个体

50	Wu 2012	穴位注射	替代指标	无纳入标准、排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,难以判断	个体
51	Zhang 2012	穴位注射	替代指标	清楚的诊断标准、纳入标准、排除标准	清楚的干预措施	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,难以判断	个体
52	Zhuo 2012	针刺	替代指标	干预措施实施者资格不清楚	不清楚的个体化裁	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,难以判断	个体
53	He 2013	穴位注射	不良事件,替代指标	无排除标准	清楚的干预措施	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,难以判断	个体
54	Wang 2013	艾灸	不良事件,替代指标	干预措施实施者资格不清楚	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,难以判断	个体
55	Chen 2014	穴位注射	不良事件,生存质量,替代指标,复合指标	清楚的诊断标准、纳入标准、排除标准	清楚的干预措施	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,难以判断	个体
56	Hu 2014	穴位注射	不良事件,替代指标	清楚的诊断标准、纳入标准、排除标准	清楚的干预措施	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,难以判断	个体
57	Lei 2014	艾灸	不良事件,替代指标	干预措施实施者资格不清楚	不清楚的个体化裁	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,难以判断	个体
58	Li 2014	穴位贴敷	不良事件,复合指标	清楚的诊断标准、纳入标准、排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,难以判断	个体
59	Tang 2014	艾灸	不良事件,替代指标	干预措施实施者资格不清楚	不清楚的个体化裁	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,难以判断	个体
60	Wang 2014	艾灸	不良事件,	干预措施实施者资格不清楚	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,难以判断	个体

61	Wei 2014	艾灸	替代指标 生存质量, 不良事件, 替代指标	干预措施实施者资格不清楚	清楚的干预措施	施不清楚 定义清楚	单中心	性试验标准 无法满足实用 性试验标准	平行对照	难以判断 无试验方案, 难以判断	个体
62	Zhang 2014a	针刺	生存质 量, 替代 指标	干预措施实施者资格不清楚	联合干预措施不清楚, 不清楚的个体化裁	联合干预措 施不清楚	单中心	无法满足实用 性试验标准	平行对照	难以判断 无试验方案, 难以判断	个体
63	Zhang 2014b	针刺	替代指标	无诊断标准、纳入标准、排 除标准, 干预措施实施者资 格不清楚	不清楚的个体化裁	定义清楚	单中心	无法满足实用 性试验标准	平行对照	难以判断 无试验方案, 难以判断	个体
64	Chen 2015a	针刺	生存质量, 不良事件, 替代指标, 复合指标	干预措施实施者资格不清楚	不清楚的个体化裁	定义清楚	单中心	无法满足实用 性试验标准	平行对照	难以判断 无试验方案, 难以判断	个体
65	Chen 2015b	穴位注射	替代指标	无纳入标准、排除标准	联合干预措施不清楚	联合干预措 施不清楚	单中心	无法满足实用 性试验标准	平行对照	难以判断 无试验方案, 难以判断	个体
66	Hu 2015	电针, 穴 位贴敷	不良事件, 替代指标	清楚的诊断标准、纳入标准、 排除标准	联合干预措施不清楚	联合干预措 施不清楚	单中心	无法满足实用 性试验标准	平行对照	难以判断 无试验方案, 难以判断	个体
67	Li 2015a	针刺	不良事件, 替代指标	干预措施实施者资格不清楚	清楚的个体化裁	定义清楚	单中心	无法满足实用 性试验标准	平行对照	难以判断 无试验方案, 难以判断	个体
68	Liu 2015	穴位贴敷	不良事件, 替代指标, 复合指标	清楚的诊断标准、纳入标准、 排除标准	清楚的干预措施	定义清楚	单中心	无法满足实用 性试验标准	平行对照	难以判断 无试验方案, 难以判断	个体

69	Wang 2015a	穴位注射	替代指标	无诊断标准、纳入标准、排除标准	联合干预措施不清楚	联合干预措施不清楚	纳入的研究均未报告应用多中心	所有纳入的研究均无法满足实用性试验标准	平行对照	无试验方案,难以判断	个体
70	Wang 2015b	穴位贴敷	生存质量	清楚的诊断标准、纳入标准、排除标准	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,难以判断	个体
71	Zhang 2015	穴位注射	不良事件,替代指标	无纳入标准、排除标准	清楚的干预措施	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,难以判断	个体
72	Cai 2016	穴位贴敷	乙肝相关并发症,不良事件,替代指标,复合指标	无排除标准	清楚的干预措施	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,难以判断	个体
73	Cheng 2016	艾灸	不良事件,替代指标	干预措施实施者资格不清楚	清楚的干预措施	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,难以判断	个体
74	Jing 2016	穴位注射	替代指标	清楚的诊断标准、纳入标准、排除标准	清楚的干预措施	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,难以判断	个体
75	Lu 2016a	针刺, 艾灸	乙肝相关并发症,不良事件,替代指标	干预措施实施者资格不清楚	不清楚的个体化裁	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,难以判断	个体
76	Lu 2016b	穴位贴敷	不良事件,替代指标	清楚的诊断标准、纳入标准、排除标准	清楚的干预措施	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,难以判断	个体
77	Zhang	针刺	生存质量,	干预措施实施者资格不清楚	联合干预措施不清楚,	联合干预措施不清楚	单中心	无法满足实用性	平行对照	无试验方案,	个体

	2016a		替代指标		不清楚的个体化裁	施不清楚		性试验标准		难以判断	
78	Zhang	穴位注射	替代指标	清楚的诊断标准、纳入标准、排除标准	清楚的干预措施	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,	个体
	2016b									难以判断	
79	Zhu 2016	穴位贴敷	不良事件, 替代指标	清楚的诊断标准、纳入标准、排除标准	清楚的干预措施	定义清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,	个体
80	Wang 2017	艾灸	不良事件, 复合指标	无纳入标准、排除标准, 干预措施实施者资格不清楚	联合干预措施不清楚	联合干预措施不清楚	单中心	无法满足实用性试验标准	平行对照	无试验方案,	个体
81	Wu 2017	穴位贴敷	不良事件, 替代指标, 复合指标	清楚的诊断标准、纳入标准、排除标准	联合干预措施均不清楚	联合干预措施不清楚	未报告应用多中心	无法满足实用性试验标准	平行对照	无试验方案,	个体

图 3-11、图 3-12 展示了可视化证据矩阵。处在最高级别的证据（即两项 meta 分析）并未对以患者为中心的结局指标（如全因死亡率、不良事件、并发症）或者替代结局指标（如可检测到的 HBV-DNA 以及 HBeAg）进行任何分析和报告，因此在本证据矩阵中未有所体现。所有检索到的证据均被评估为设计误差高风险。

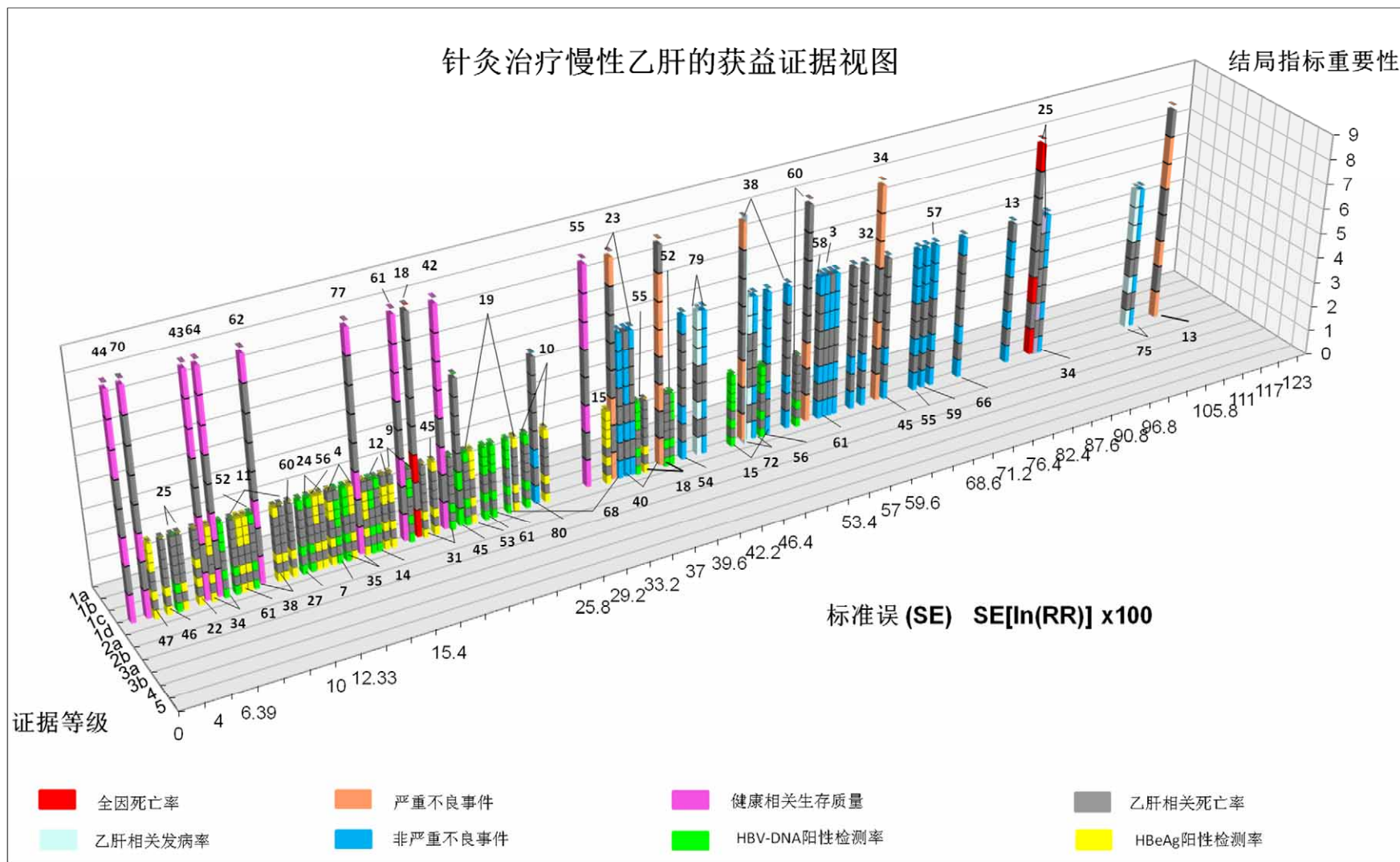


图 3-11 根据系统误差，随机误差和设计误差构建的三维误差证据矩阵-针灸干预慢性乙型肝炎的获益

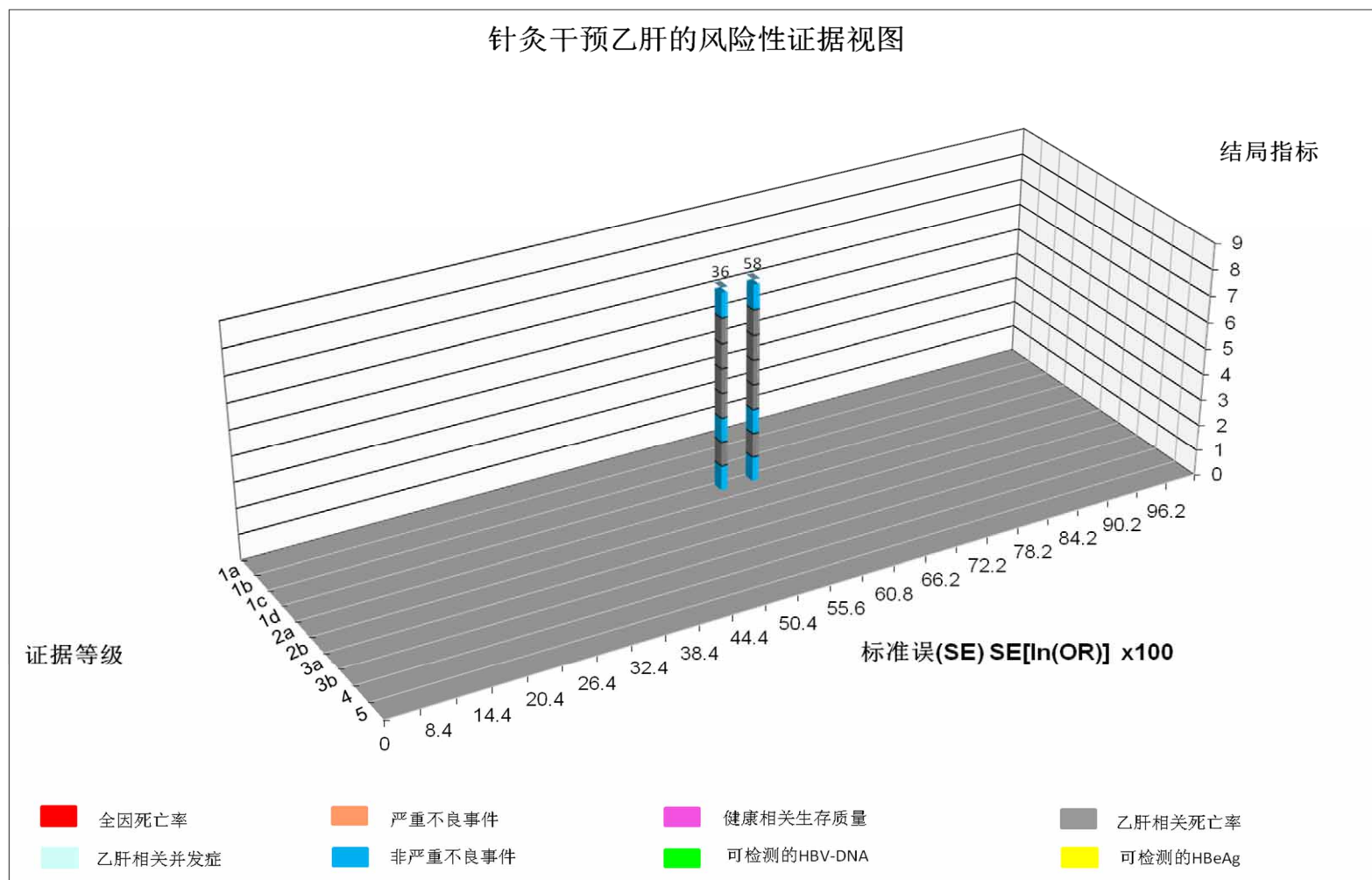


图 3-12 根据系统误差，随机误差和设计误差构建的三维误差矩阵-针灸干预慢性乙型肝炎的风险

讨 论

通过本研究，我们未能找到关于小柴胡汤或者针灸这两种常用的干预措施治疗慢性乙型肝炎的任何高质量系统综述。在缺少高质量的证据前对患者实施干预措施是一种存在于西医界和中医界的普遍现象^[239-241]。因此，我们建议在临床实践中，实施任何干预措施前，先从高质量的系统综述中获得证据，因为高质量的系统综述能够使各种误差最小化。

随机临床对照试验，作为干预措施效应证据的基石^[23, 96, 214]，应当使能够影响证据可靠性的多种误差风险（即系统误差、随机误差、设计误差）最小化。如随机序列的产生、随机分配的隐藏、盲法^[9, 242]、不完整的结局指标^[243, 244]、选择性报告^[245, 246]以及商业涉入这些会导致系统误差的项目的高风险，会使干预措施效应评估产生偏倚。这些偏倚会导致过高评估干预措施的有效性和过低评估干预措施的有害性。图 3-1、图 3-2、图 3-7、图 3-8 表明纳入的随机对照试验在以上各方面均存在大量的偏倚风险。因此，我们建议未来的临床试验应对能够以上导致高风险系统误差的试验要素进行严谨的设计和实施。

对临床随机对照试验证据评估表明对于替代结局指标，证据存在高风险、大量风险或中等程度风险的随机误差。样本量不充足会导致随机误差^[247]。对于二分类变量，随机对照试验中等样本量估计应考虑以下参数：I 类错误(α)和 II 类错误(β)风险的最大值以及对于主要结局指标预期的干预措施效应量(μ)。对于连续变量，主要结局指标的方差同样需要被考虑^[23, 248]。此外，由于在小样本中预后因素分布不均以及累积数据的重复测量^[104, 220, 249]，在达到预期样本量前停止随机对照试验（即序贯分析），还应当应用更严格的调整后的 P 值^[250]。对于随机对照试验的 meta 分析，试验序贯分析进一步发展了序贯分析的方法学，并且评估所需样本量是否足够接受或者拒绝一项干预措施的假设效应^[18, 104, 105]。我们建议研究者设计和实施具有足够样本量的随机对照试验，并且在进行系统综述时运用试验序贯分析。在进行试验序贯分析的同时，还应进行 GRADE 分级以辅助评价证据体的不精确性^[108, 110]。

替代结局指标可能会展现出干预措施的显著有效性，然而以患者为中心的结局指标可能展现出干预措施无效甚至有害^[205]。替代结局指标结果是否真实反映临床有效性的提高是充满疑问，并且有待证实。本研究中被检索到的研究大多关注于未经证实的替代结局指标。关于设计误差的其他要素（具体见表 3-3），检索到的证据同样表明具有大量的偏倚风险，因而可能影响证据的外推性。此外，只有当主要结局指标表明无临床及统

计学显著差异时，次要结局指标结果才可以被用于干预措施的临床实施推荐^[22]。因此，未来的临床试验应当更多的关注以患者为中心的结局指标，更加明确定义及报告的诊断标准、纳入标准及排除标准，以及实施多中心临床试验。在探索干预措施的效应时，应当设计和实施实效性以及有效性研究^[205, 207]。此外，对于通常由不同水平的医疗实施者根据患者个体进行量体化裁中医药干预措施，也应当在试验设计和实施阶段考虑个体化裁的细节以及医疗实施者的资格标准。

证据矩阵从系统误差、随机误差以及设计误差三个维度为检索到的证据提供了一个综合的、可视化的质量评估工具。该证据矩阵还为进一步系统评价证据质量的五个方面（即研究间偏倚、证据的间接性、不可解释的异质性和结果的不一致性、结果的不精确性、发表偏倚）的 GRADE 评估提供详细的各个误差要素的可视化信息。研究员应当尽力“点亮”证据矩阵的所有楼层以获得最可靠的证据。证据矩阵可以帮助我们找到现存的临床证据中的主要问题，为未来的临床研究指明研究方向，以进一步促进临床实践中安全和正确的干预措施的实施。

尽管 PRISMA 声明^[251]和 CONSORT 声明^[247]已经指出如何提高系统综述和随机对照试验的报告质量，本研究中纳入的大部分研究仍然缺失诸如研究方案、随机序列产生、分配隐藏以及缺失数据处理等大量信息。不幸的是，当我们试图联系纳入研究的研究者时，我们仅从一项研究的研究者那里收到了缺失数据的反馈信息。这种现象导致我们很难判断临床研究的各种误差风险究竟是由研究的方法学低质量还是由于研究报告的低质量引起。

由于未考虑多重结局指标检验以及累积数据的重复测量，标准误可能会引起随机误差^[22]。不同的临床问题，其在一个水平等级内的结局指标（及紧急、重要或不重要）可以内部调换。此外，结局指标测量，如全因死亡率以及乙肝相关并发症通常是相互关联的，而这种关联往往容易被研究者忽略。

异质性，包含统计学异质性^[95, 97]以及临床异质性（如各研究间 PICOTA 要素的不一致性）是在评估证据质量时需要考虑的另一个方面。关于异质性的评估，在进行系统综述和对证据进行 GRADE 评价时也应对此进行考虑。在评估发表偏倚风险时，也同样应考虑异质性的评估^[218, 219]。

小 结

1. 共有 2 项 meta 分析, 52 项随机对照试验探讨小柴胡汤治疗慢性乙型肝炎的有效性及安全性。通过本研究的可视化三维误差证据矩阵评估, 由于存在系统误差和设计误差高风险, 和小风险随机误差, 两项 meta 分析被评为低质量研究, 并且两项 meta 分析均不数据系统综述。纳入的 52 项随机对照试验展示系统误差和设计误差的高偏倚风险, 其中 9 项研究显示具有高随机误差风险。
2. 两项 meta 分析、77 项临床随机对照试验以及两项病例系列报告研究探索了针灸干预慢性乙型肝炎的有效性和安全性。通过本研究的可视化误差矩阵评估方法, 由于存在系统误差和设计误差高风险, 和小风险随机误差, 两项 meta 分析被评为低质量研究, 并且两项 meta 分析均不数据系统综述。纳入的 72 项随机对照试验展示系统误差和设计误差的高偏倚风险, 其中 18 项研究显示具有高随机误差风险。
3. 我们建议在临床实践中, 实施任何干预措施前, 先从高质量的系统综述中获得证据以使各种误差最小化, 并且在进行系统综述时运用试验序贯分析以及 GRADE 分级以辅助评价证据体的质量。我们提出提高中医临床证据质量的相关建议: 在未来的临床随机对照试验设计中应对导致高风险系统误差的试验要素(如随机化、盲法、随机分配隐藏、不完整的结局指标、选择性报告)进行严谨的设计、实施和报告; 应设计足够样本量以及多样性参与者的随机对照试验, 以减少随机误差; 应更多设计以患者为中心的结局指标(如死亡率、严重不良事件、健康相关生存质量等)的临床试验; 应在研究设计阶段即明确疾病的诊断标准、参与者纳入标准及排除标准, 并详细报告; 应设计实施多中心临床试验, 尽可能的进行实效性及优效性研究; 针对中医辨证论治的特色, 对于中医药疗法的临床研究, 应当在试验设计和实施阶段清楚定义个体化裁的细节, 并制定医疗干预措施实施者的资格标准。
4. 本研究创立的“三维误差证据矩阵”与目前常用的偏倚评估工具“Risk of bias tool”以及 Grade 分级标准既相互支撑, 又可弥补后两者对于设计误差具体影响因素评估的不足。

结 论

1. 中药复方小柴胡汤干预慢性乙型肝炎的循证评价

获得的小柴胡汤干预慢性乙型肝炎的证据缺少对于全因死亡率、健康相关生存质量、严重不良事件、乙肝相关死亡率、乙肝相关发病率这些结局指标的评估和报告。通过评价低质量的证据表明，我们发现小柴胡汤联合常规疗法对比单用常规疗法可降低 HBV-DNA 阳性检测率，但对于非严重不良事件和 HBeAg 阳性检测率改变这两个结局指标，目前尚不能表明小柴胡汤联合常规疗法对比单用常规疗法具有显著有效性。由于检索到的研究具有高风险偏倚，HBV-DNA 阳性检测率改变是未经证实的替代结局指标，并且纳入的研究其确定性较低，因此结论尚不能确证。

2. 针灸疗法干预慢性乙型肝炎的循证评价

获得的针灸疗法（穴位注射、中药贴敷、手针、艾灸疗法）干预慢性乙型肝炎的证据缺少对于全因死亡率、健康相关生存质量、严重不良事件、乙肝相关死亡率、乙肝相关发病率这些结局指标的评估和报告。通过评价低质量的证据表明，我们不同发现针灸疗法联合常规疗法对比单用常规疗法可降低 HBV-DNA 阳性检测率，但对于非严重不良事件和 HBeAg 阳性检测率这两个结局指标，目前尚不能表明针灸疗法联合常规疗法对比单用常规疗法具有显著有效性。针灸疗法联合常规治疗与单用常规治疗相比较，可能增加不良事件腹胀的发生率。但由于检索到的研究具有高风险偏倚，HBV-DNA 阳性检测率是未经证实的替代结局指标，且只有一项研究能够提供数据，纳入的研究其确定性较低，因此关于针灸疗法对慢性乙型肝炎的有益性结论尚不能确证。而对于针灸疗法增加慢性乙型肝炎患者腹胀事件的风险，同样由于只有一项研究能够提供数据，且纳入的研究其确定性较低，其结论同样不能确证。

3. 中医临床证据体质量评估方法和临床研究设计改进方法研究

我们建议在临床实践中，实施任何干预措施前，先从高质量的系统综述中获得各种误差最小化，质量最高的证据，并且在进行系统综述时运用试验序贯分析以及 GRADE 分级以辅助评价证据体的质量。

运用本研究的可视化“三维误差证据矩阵”发现目前小柴胡汤和针灸疗法干预慢性乙型肝炎的临床证据在系统误差、随机误差、设计误差三个方面存在较严重问题。我们提出提高中医临床证据质量的相关建议：在未来的临床随机对照试验设计中应对导致高

风险系统误差的试验要素（如随机化、盲法、随机分配隐藏、不完整的结局指标、选择性报告）进行严谨的设计、实施和报告；应设计足够样本量以及多样性参与者的随机对照试验，以减少随机误差；应更多设计以患者为中心的结局指标（如死亡率、严重不良事件、健康相关生存质量等）的临床试验；应在研究设计阶段即明确疾病的诊断标准、参与者纳入标准及排除标准，并详细报告；应设计实施多中心临床试验，尽可能的进行实效性及优效性研究；针对中医辨证论治的特色，对于中医药疗法的临床研究，应当在试验设计和实施阶段清楚定义个体化裁的细节，并制定医疗干预措施实施者的资格标准。

本研究创立的“三维误差证据矩阵”与目前常用的偏倚评估工具“Risk of bias tool”以及 Grade 分级标准既相互支撑，又可弥补后两者对于设计误差具体影响因素评估的不足。

本研究创新性的自我评价

1. 目前只有三篇未经严格方法学设计的 meta 分析评价小柴胡汤干预慢性乙型肝炎的疗效：三项研究的研究人群患病均较单一；均未试图对以患者为中心的结局指标（如死亡率、不良事件、生存质量等）进行评估；运用的风险偏倚评估工具不够全面和严格；没有运用 TSA 分析来控制 meta 分析的随机误差；亦没有运用 GRADE 分级来定义证据的确定性。本研究首次通过严格方法学设计，构建 Cochrane 系统综述对小柴胡汤干预慢性乙型肝炎的获益和风险进行循证评价；首次关注患者为中心的结局指标（全因死亡率、健康相关生存质量、严重不良事件、乙肝相关死亡率、乙肝相关发病率）的效应评估；首次运用 TSA 分析来探索目前是否已满足获得“充分证据”证明小柴胡汤干预慢性乙型肝炎的获益风险所需要的所需样本量大小，并对获得的证据进行统计显著性检验和无效性检验，以判断是否需要进行更多的临床试验以进一步提供证据；首次运用 GRADE 分级对获得的证据质量确定性进行评估，为临床实践提供推荐意见；首次提出对于小柴胡汤干预慢性乙型肝炎的临床随机对照试验：应针对不同的小柴胡汤方剂组成、不同的剂量以及不同的疗程进行试验设计；更多设计对比小柴胡汤与安慰剂干预慢性乙型肝炎的临床试验。

2. 目前只有两篇未经严格方法学设计的 meta 分析评价穴位注射疗法和艾灸疗法分别干预慢性乙型肝炎的疗效：两项研究的研究人群患病均较单一；均未试图对以患者为中心的结局指标（如死亡率、不良事件、生存质量等）进行评估；运用的风险偏倚评估工具不够全面和严格；没有运用 TSA 分析来控制 meta 分析的随机误差；亦没有运用 GRADE 分级对证据进行评估。本研究首次通过严格方法学设计，构建 Cochrane 系统综述试图对针灸疗法（手针、艾灸、电针、激光针刺、穴位按压、穴位注射、穴位贴敷等）干预慢性乙型肝炎的获益和风险进行循证评价；首次关注患者为中心的结局指标（全因死亡率、健康相关生存质量、严重不良事件、乙肝相关死亡率、乙肝相关发病率）的效应评估；首次运用 TSA 分析来探索目前是否已满足获得“充分证据”证明不同针灸疗法干预慢性乙型肝炎的获益风险所需要的所需样本量大小，并对获得的证据进行统计显著性检验和无效性检验，以判断是否需要进行更多的临床试验以进一步提供证据；首次运用 GRADE 分级对获得的证据质量确定性进行评估，为临床实践提供推荐意见；首次提出对于针灸疗法干预慢性乙型肝炎的临床随机对照试验：应设计如激光针刺等更多疗法，或

者多种疗法联合应用干预慢性乙型肝炎的研究；更多设计针刺疗法对比假针刺疗法的临床试验。

3. 通过本研究发现，在临床实践过程中，往往缺乏最高质量证据支持。我们建议在临床实践中，实施任何干预措施前，先从高质量的系统综述中获得证据以使各种误差最小化，并且在进行系统综述时运用序贯分析以及 GRADE 分级以辅助评价证据体的质量。

4. 本研究建立的可视化“三维误差证据矩阵”可以从系统误差、随机误差、设计误差三个方面全面发现目前临床证据存在的问题。可通过系统误差、随机误差、设计误差三个维度提出中医临床研究设计改进方法，提高中医临床证据质量，促进中医疗法的临床实施，为未来中医药临床研究指明研究方向。

5. 我们首次从系统误差、随机误差、设计误差三个维度提出提高中医临床证据质量的相关建议：在未来的临床随机对照试验设计中应对随机化（应针对不同的合并疾病进行分层随机）、盲法、随机分配隐藏、不完整的结局指标数据、结局指标选择性报告等方面进行严谨的设计、实施和报告；应设计足够样本量以及多样性参与者的随机对照试验，以减少随机误差，并对于研究人群的诊断标准以及合并疾病进行更严格清晰的限定，设立更长随访期；为了减少设计误差风险：应更多设计以患者为中心的结局指标（如死亡率、严重不良事件、健康相关生存质量等）的临床试验；应在研究设计阶段即明确疾病的诊断标准、参与者纳入标准及排除标准，并详细报告；应设计实施多中心临床试验，尽可能的进行实效性及优效性研究；针对中医辨证论治的特色，对于中医药疗法的临床研究，应当在试验设计和实施阶段清楚定义个体化裁的细节，并制定医疗干预措施实施者的资格标准。

参考文献

- [1] Geneva: World Health Organization. Global Hepatitis Report,2017. Licence: CC BY-NC-SA 3.0 IGO.
- [2] Lozano R, Naghavi M, Foreman K, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012, 380:2095-2128.
- [3] Liang X, Bi S, Yang W, et al. Epidemiological serosurvey of hepatitis B in China-declining HBV prevalence due to hepatitis B vaccination. *Vaccine*. 2009, 27: 6550-6557.
- [4] Sackett DL, Rosenberg WM, Gray JA, et al. Evidence based medicine: what it is and what it isn' t. *BMJ*. 1996, 312(7023):71 - 2.
- [5] Silvio Garattini, Janus C. Jakobsen, Jørn Wetterslev, et al. Evidence-based clinical practice: overview of threats to the validity of evidence and how to minimise them. *European Journal of Internal Medicine*. 2016, Vol. 32, 13-21.
- [6] Evidence-Based Medicine Working Group: Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA*. 1992, 268:2420-5.
- [7] Sutton AJ, Higgins JPT. Recent developments in meta-analysis. *Stat Med*. 2008, 27:625-50.
- [8] Straus SE, Richardson WS, Glasziou P, et al. Evidence-based medicine. How to practice and teach EBM. Edinburgh, UK: Churchill Livingstone, 2005.
- [9] Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. The Cochrane Collaboration. 2008.
- [10]刘为民,刘保延,谢雁鸣.循证医学在中医药国际化中的应用研究.世界科学技术,2006,(03):42-46.
- [11]刘建平,夏芸.中文期刊发表的中医药系统综述或 Meta-分析文章的质量评价.中国中西医结合杂志,2007,(04):306-311.
- [12] Puhan MA, Schunemann HJ, Murad MH, et al. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. *BMJ*. 2014,349: g5630.
- [13] Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines 3: rating the quality of evidence - introduction. *Journal of Clinical Epidemiology*. 2011,64:401-6.
- [14] Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction - GRADE evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology*. 2011, 64(4):383-94.

- [15]宋海阔,张玄,郭义,等.制定《中医治未病实践指南》的方法学探讨. 河南中医. 2017,(12):2066-2067.
- [16] Borm GF, Donders AR. Updating meta-analyses leads to larger type I errors than publication bias. *J Clin Epidemiol.* 2009, 62:825-830.
- [17] Lan KK, DeMets D. Discrete sequential monitoring boundaries for clinical trials. *Biometrika.* 1983:659-663.
- [18] Jørn Wetterslev, Kristian Thorlund, Jesper Brok, et al. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. *Journal of Clinical Epidemiology.* 2008, 61:64-75.
- [19] Jørn Wetterslev, Kristian Thorlund, Jesper Brok, et al. Estimating required information size by quantifying diversity in random-effects model meta-analyses. *BMC Medical Research Methodology.*2009, 9:86.
- [20] David MK, Georgios K. Against pragmatism: on efficacy, effectiveness and the real world. *Trials.* 2009, 10: 48.
- [21] Glasgow RE, Green LW, Klesges LM, et al. External validity: we need to do more. *Ann Behav Med.* 2006, 31: 105-108.
- [22] Frederik Keus, Jørn Wetterslev, Christian Gluud, et al. Evidence at a glance: error matrix approach for overviewing available evidence. *BMC Medical Research Methodology.* 2010, 10:90.
- [23] Jakobsen JC, Gluud C. The necessity of randomized clinical trials. *Br J Med Res.* 2013, 3:1453-68.
- [24] Rothwell PM. External validity of randomized controlled trials: ‘to whom do the results of this trial apply?’. *Lancet.* 2005, 365: 82-93.
- [25] 吴宇侠,康德英,洪旗,等.临床随机对照试验的外部真实性及其评价的探讨. 中华流行病学杂志. 2011,32:514-518.
- [26] Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med.* 2013, 158(3):200-7.
- [27] SPIRIT 中文工作组. SPIRIT2013 声明: 定义临床研究方案的标准条目. 中国中西医结合杂志. 2014, 34: 115-122.
- [28] World Health Organization. Global hepatitis report, 2017. www.who.int/hepatitis/publications/global-hepatitis-report2017/en/ 2017 (accessed 20 July 2018).
- [29] World Health Organization. Guidelines for the prevention, care and treatment of persons with chronic hepatitis B infection. www.who.int/hiv/pub/hepatitis/hepatitis-b-guidelines/en/ 2015 (accessed 20 July 2018).
- [30] Terrault NA, Bzowej NH, Chang KM, et al. AASLD guidelines for treatment of chronic

- hepatitis B. *Hepatology* (Baltimore, Md.). 2016,63(1):261-83.
- [31] European Association for the Study of the Liver. EASL 2017 clinical practice guidelines on the management of hepatitis B virus infection. *Journal of Hepatology*. 2017, 67(2):370-98.
- [32] Alizadeh AHM, Ranjbar M, Yadollahzadeh M. Patient concerns regarding chronic hepatitis B and C infection. *La Revue de Santé de la Méditerranée Orientale*. 2008,14(5):1142-7.
- [33] Lu JJ, Xu AQ, Wang J, et al. Direct economic burden of hepatitis B virus related diseases: evidence from Shandong, China. *BMC Health Services Research*. 2013, 31(13):37.
- [34] Keshavarz K, Kebriaeezadeh A, Alavian SM, et al. Economic burden of hepatitis B virus-related diseases: evidence from Iran. *Hepatitis Monthly*. 2015,15(4):e25854.
- [35] Ezbarami ZT, Hassani P, Tafreshi MZ, et al. A qualitative study on individual experiences of chronic hepatitis B patients. *Nursing Open*. 2017, 4(4):310-8.
- [36] Peng WW, editor(s). *Infectious Diseases – Hepatitis B*. 5 edition. Beijing: People's Health Publishing House, 2000.
- [37] Nassal M. Hepatitis B viruses: reverse transcription a different way. *Virus Research*. 2008;134:235-49.
- [38] Moraleda G, Saputelli J, Aldrich CE, et al. Lack of effect of antiviral therapy in nondividing hepatocyte cultures on the closed circular DNA of woodchuck hepatitis virus. *Journal of Virology*. 1997, 71: 9392-9.
- [39] Delmas J, Schorr O, Jamard C, et al. Inhibitory effect of adefovir on viral DNA synthesis and covalently closed circular DNA formation in duck hepatitis B virus-infected hepatocytes in vivo and in vitro. *Antimicrobial Agents and Chemotherapy*. 2002, 46(2): 425-33.
- [40] Gripon P, Rumin S, Urban S, et al. Infection of a human hepatoma cell line by hepatitis B virus. *Proceedings of the National Academy of Sciences of the United States of America*. 2002, 99(24):15655-60.
- [41] Zoulim F. New insight on hepatitis B virus persistence from the study of intrahepatic viral cccDNA. *Journal of Hepatology*. 2005,42(3):302-8.
- [42] Peters van Ton AM, Gevers TJ, Drenth JP. Antiviral therapy in chronic hepatitis E: a systematic review. *Journal of Viral Hepatitis*. 2015,22(12):965-73.
- [43] Ubiña-Aznar E, Fernández-Moreno N, Rivera-Irigoín R, et al. Pulmonary sarcoidosis associated with pegylated interferon in the treatment of chronic hepatitis C. *Gastroenterologia Hepatologia*. 2005;28(8):450-2.
- [44] Nadeem A, Hussain MM, Aslam M, et al. Interferon-alpha induced and ribavirin induced thyroid dysfunction in patients with chronic hepatitis C. *Hepatitis Monthly*

2010,10(2):132-40.

[45] Orságová I, Rožnovský L, Petroušová L, et al. Investigation of autoimmunity markers during interferon alpha therapy of chronic hepatitis B and C - twenty years of experience. *Klinicka Mikrobiologie a Infekcni Lekarstvi*. 2016,22(2):61-7.

[46] Cattie JE, Letendre SL, Woods SP, et al. Persistent neurocognitive decline in a clinic sample of hepatitis C virus-infected persons receiving interferon and ribavirin treatment. *Journal of Neurovirology*. 2014,20(6):561-70.

[47] Yuen MF, Sablon E, Hui CK, et al. Factors associated with hepatitis B virus DNA breakthrough in patients receiving prolonged lamivudine therapy. *Hepatology (Baltimore, Md.)*. 2001,34:785-91.

[48] Papatheodoridis GV, Dimou E, Laras A, et al. Course of virologic breakthroughs under long-term lamivudine in HBeAg-negative precore mutant HBV liver disease. *Hepatology (Baltimore, Md.)*. 2002,36(1):219-26.

[49] Liaw YF, Sung JJ, Chow WC, et al. Lamivudine for patients with chronic hepatitis B and advanced liver disease. *New England Journal of Medicine*. 2004,351(15):1521-31.

[50] Hongthanakorn C, Chotiyaputta W, Oberhelman K, et al. Virological breakthrough and resistance in patients with chronic hepatitis B receiving nucleos(t)ide analogues in clinical practice. *Hepatology (Baltimore, Md.)*. 2011,53(6):1854-63.

[51] Miyauchi T, Kanda T, Shinozaki M, et al. Efficacy of lamivudine or entecavir against virological rebound after achieving HBV DNA negativity in chronic hepatitis B patients. *International Journal of Medical Sciences*. 2013,10(6):647-52.

[52] Zhang HY, Liu LG, Ye CY, et al. Evolution of drug-resistant mutations in HBV genomes in patients with treatment failure during the past seven years (2010-2016). *Virus Genes*. 2017,11:1-7.

[53] World Health Organization. Hepatitis B. www.who.int/news-room/fact-sheets/detail/hepatitis-b-2000 (accessed 20 July 2018).

[54] Zheng X, Wang J, Yang D. Antiviral therapy for chronic hepatitis B in China. *Medical Microbiology and Immunology*. 2015,204:115-20.

[55] 张学海,郭秀珠.小柴胡汤对慢乙肝 HBVM 及 ALT 的影响.中国中医药信息杂志,1998,(08):36.

[56] 余安清,王克奇,张文鼎,等.加味小柴胡汤与病毒唑联用治疗慢性乙型肝炎.浙江中西医结合杂志,2000,(03):10-12.

[57] 熊凡,孙静,熊伟.干扰素联合小柴胡汤治疗肝纤维化的临床观察.湖北中医杂志,2003,(10):10-11.

[58] Tajiri H, Kozaiwa K, Ozaki Y, et al. Effect of sho-saiko-to (xiao-chai-hu-tang) on

- HBeAg clearance in children with chronic hepatitis B virus infection and with sustained liver disease. *American Journal of Chinese Medicine*. 1991,19(2):121-9.
- [59] Yamashiki M, Kosaka Y, Nishimura A, et al. Efficacy of a herbal medicine 'sho-saiko-to' on the improvement of impaired cytokine production of peripheral blood mononuclear cells in patients with chronic viral hepatitis. *Journal of Clinical & Laboratory Immunology*. 1992,37(3):111-21.
- [60]张仲景.伤寒论.北京:人民卫生出版社,2005.
- [61] The Ministry of Health, Labour and Welfare. The Japanese Pharmacopoeia, seventeenth edition. www.mhlw.go.jp/file/06-Seisakujouhou-11120000-Iyakushokuhinkyoku/JP17_REV_1.pdf 2016 (accessed 20 July 2018).
- [62]温志坚,曹毓,彭龙玲.小柴胡汤提取物体外抗 HBV 活性研究.中西医结合肝病杂志,2000,(04):29-30.
- [63]盖丽丽.小柴胡片对慢性乙肝患者外周血树突状细胞抗原递呈的影响.广州.广州中医药大学,2007.
- [64] Liu ZJ, Wang ZL, Sun XH, et al. Effect of modified Xiaochaihu decoction on expression of IL-12 , IL-4, and IFN- γ mRNA in HBV transgenic mice. *Chinese Journal of Integrated Traditional and Western Medicine on Liver Diseases*. 2010,20:289-90.
- [65] Bachem MG, Meyer D, Schäfer W, et al. The response of rat liver perisinusoidal lipocytes to polypeptide growth regulator changes with their transdifferentiation into myofibroblast-like cells in culture. *Journal of Hepatology*. 1993,18:40-52.
- [66] Ma YR, Shimizu I, Ito S, et al. Effect of Xiaochaihu decoction on experimental hepatic fibrosis in rats. *Journal of Luzhou Medical College*. 1997,20:81-6.
- [67] Ono M, Miyamura M, Kyotani S, et al. Effect of Sho-saiko-to extract on HGF and TGF-beta levels of intra organs in liver-injured rats after partial hepatectomy. *Journal of Pharmacy and Pharmacology*. 2000,52:111-8.
- [68]张学敏,谢斌.加味小柴胡汤对 D-氨基半乳糖所致小鼠肝损伤的保护作用.中国热带医学,2006,(07):1134-1135.
- [69] Yano H, Mizoguchi A, Fukuda K, et al. The herbal medicine sho-saiko-to inhibits proliferation of cancer cell lines by inducing apoptosis and arrest at the G0/G1 phase. *Cancer Research*. 1994,54:448-54.
- [70] Cao HY. The new mechanism of Xiaochaihu decoction in inhibiting cancer. *Foreign Medical Sciences* 2003;24:232-3.
- [71] Wang SJ, Ai Q. Effect of Xiaochaihu decoction on hepatic fibrosis and liver cancer. *Progress in Japanese Medicine*. 2004,25:42-4.

- [72] Mao M, Fu H, Huang X-S, et al. Small Radix Bupleuri decoction inducing bearing cancer mice S180 cell apoptosis and influence of it on cell cycle. *Modern Journal of Integrated Traditional Chinese and Western Medicine*. 2005,14:2646-8.
- [73] 吴俏青.小柴胡汤化裁治疗乙型肝炎肝硬化 69 例临床观察. *浙江中医药大学学报*,2009,33(04):517-518.
- [74] Takada N, Arai S, Kusuhara N, et al. A case of sho-saiko-to-induced pneumonitis, diagnosed by lymphocyte stimulation test using bronchoalveolar lavage fluid. *Nihon Kyobu Shikkan Gakkai Zasshi*. 1993,31(9):1163-9.
- [75] Hatakeyama S, Tachibana A, Morita M, et al. Five cases of pneumonitis induced by Sho-Saiko-To. *Nihon Kyobu Shikkan Gakkai Zasshi*. 1997,35(5):505-10.
- [76] Sato A, Toyoshima M, Kondo A, et al. Pneumonitis induced by the herbal medicine Sho-saiko-to in Japan. *Nihon Kyobu Shikkan Gakkai Zasshi*. 1997,35(4):391-5.
- [77] Tsumura & Co. TSUMURA Shosaikoto Extract Granules for Ethical Use. www.tsumura.co.jp/english/products/pi/JPR_T009.pdf 2014 (accessed 20 July 2018).
- [78] Itoh S, Marutani K, Nishijima T, et al. Liver injuries induced by herbal medicine, syo-saiko-to (xiao-chai-hu-tang). *Digestive Diseases and Sciences*.1995,40(8):1845-8.
- [79] Stickel F, Egerer G, Seitz HK. Hepatotoxicity of botanicals. *Public Health Nutrition*. 2000,3(2):111.
- [80] Hsu L-M, Huang Y-S, Tsay S-H, et al. Acute hepatitis induced by Chinese hepatoprotective herb, xiao-chai-hu-tang. *Journal of the Chinese Medical Association*. 2006,69(2):86-8.
- [81] Kiguchi T, Kimura F, Niiya K, et al. Acute thrombocytopenic purpura after ingestion of Sho-saiko-to for hepatitis. *Liver*. 2000,20(6):491.
- [82] Sakamoto O, Ichikado K, Kohroggi H, et al. Clinical and CT characteristics of Chinese medicine-induced acute respiratory distress syndrome. *Respirology*. 2003,8(3):344-50.
- [83] International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH harmonised tripartite guideline. Clinical safety data management: definitions and standards for expedited reporting E2A. In: www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E2A/Step4/E2A_Guideline.pdf. (accessed 2 November 2018).
- [84] International Council on Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. ICH harmonised guideline. Integrated Addendum to ICH E6(R1): Guideline for good clinical practice E6(R2).. In: www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R2_Step_4_2016_1109.pdf. (accessed 02 October 2018).

- [85] Higgins JP, Green S, editor(s). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.
- [86] <https://hbg.cochrane.org/>.
- [87] Review Manager 5 (RevMan 5) [Computer program]. Version 5.3. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2014.
- [88] Thorlund K, Engström J, Wetterslev J, et al. User manual for Trial Sequential Analysis (TSA). ctu.dk/tsa/files/tsa_manual.pdf 2011 (accessed 16 January 2017).
- [89] TSA - Trial Sequential Analysis [Computer program]. Version 0.9.5.10 Beta. Copenhagen: Copenhagen Trial Unit, 2011. <https://www.ctu.dk/tsa/downloads.aspx>.
- [90] Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ (Clinical Research Ed.)*. 2009,339:b2700.
- [91] Schulz KF, Chalmers I, Hayes RJ, et al. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA*. 1995,273(5):408-12.
- [92] Moher D, Pham B, Jones A, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet*. 1998,352(9128):609-13.
- [93] Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. *Annals of Internal Medicine*. 2001,135(11):982-9.
- [94] Wood L, Egger M, Gluud LL, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ (Clinical Research Ed.)*. 2008,336(7644):601-5.
- [95] Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*. 2002,21(11):1539-58.
- [96] Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. *BMJ (Clinical Research Ed.)*. 2003,327(7414):557-60.
- [97] Wetterslev J, Thorlund K, Brok J, et al. Estimating required information size by quantifying diversity in a random-effects meta-analysis. *BMC Medical Research Methodology*. 2009,9:86.
- [98] Jakobsen J, Wetterslev J, Winkel P, et al. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. *BMC Medical Research Methodology*. 2014,14:120.

- [99] Fisher RA. On the interpretation of X^2 from contingency tables, and the calculation of P. *Journal of the Royal Statistical Society*. 1922,85(1):87-94.
- [100] Student. The probable error of a mean. *Biometrika*. 1908,6(1):1-25.
- [101] Wetterslev J, Thorlund K, Brok J, et al. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. *Journal of Clinical Epidemiology*. 2008,61:64-75.
- [102] Thorlund K, Imberger G, Walsh M, et al. The number of patients and events required to limit the risk of overestimation of intervention effects in meta-analysis – a simulation study. *PLoS One*. 2011,6(10):e25491.
- [103] Wetterslev J, Thorlund K, Brok J, et al. Estimating required information size by quantifying diversity in a random-effects meta-analysis. *BMC Medical Research Methodology*. 2009,9:86.
- [104] Wetterslev J, Jakobsen JC, Gluud C. Trial sequential analysis in systematic reviews with meta-analysis. *BMC Medical Research Methodology*. 2017,17(1):39.
- [105] Brok J, Thorlund K, Gluud C, et al. Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses. *Journal of Clinical Epidemiology*. 2008,61:763-9.
- [106] Brok J, Thorlund K, Wetterslev J, et al. Apparently conclusive meta-analyses may be inconclusive - trial sequential analysis adjustment of random error risk due to repetitive testing of accumulating data in apparently conclusive neonatal meta-analyses. *International Journal of Epidemiology*. 2009,38(1):287-98.
- [107] Thorlund K, Anema A, Mills E. Interpreting meta-analysis according to the adequacy of sample size. An example using isoniazid chemoprophylaxis for tuberculosis in purified protein derivative negative HIV-infected individuals. *Clinical Epidemiology*. 2010,2:57-66.
- [108] Castellini G, Nielsen EE, Gluud C. Comment on: 'Cell therapy for heart disease: trial sequential analyses of two Cochrane reviews'. *Clinical Pharmacology and Therapeutics*. 2017,102:21-4.
- [109] Thorlund K, Engstrøm J, Wetterslev J, et al. User manual for Trial Sequential Analysis (TSA). ctu.dk/tsa/files/tsa_manual.pdf 2011 (accessed 16 January 2017).
- [110] Castellini G, Bruschetti M, Gianola S, et al. Assessing imprecision in Cochrane systematic reviews: a comparison of GRADE and Trial Sequential Analysis. *BMC* 2018;7:110. [Other: <https://doi.org/10.1186/s13643-018-0770-1>]
- [111] Gartlehner G, Nussbaumer-Streit B, Wagner G, et al. Increased risks for random errors are common in outcomes graded as high certainty of evidence. *Journal of Clinical Epidemiology* 2018 Oct 18 [Epub ahead of print]. [DOI: 10.1016/j.jclinepi. 2018.10.009].
- [112] Bachem MG, Meyer D, Schäfer W, et al. The response of rat liver perisinusoidal lipocytes to polypeptide growth regulator changes with their transdifferentiation into

myofibroblast-like cells in culture. Journal of Hepatology. 1993,18:40-52.

[113] GRADEpro GDT [Computer program]. Hamilton (ON): McMaster University (developed by Evidence Prime), 2015. Available at grade.org.

[114] Guyatt GH, Ebrahim S, Alonso-Coello P, et al. GRADE guidelines 17: assessing the risk of bias associated with missing participant outcome data in a body of evidence. Journal of Clinical Epidemiology 2017;87:14-22. [PubMed: 28529188]

[115]孙晓娜,孙俊波.小柴胡汤联合甘利欣注射液治疗肝炎肝纤维化 60 例.中国中医药信息杂志,2004,(06):534-535.

[116]中华医学会肝病学会,中华医学会感染病学会.慢性乙型肝炎防治指南.世界临床药物,2006,27(2):36-41.

[117]中国中西医结合学会肝病专业委员会.肝纤维化中西医结合诊疗指南.药品评价,2007,14(4):1052-1056.

[118]郑筱萸.中药新药临床研究指导原则.北京:中国医药科技出版社,2002:143-150.

[119]中华医学会肝病学会,中华医学会感染病学会.慢性乙型肝炎防治指南(2010年版).中华肝脏病杂志.2011,19(1):13-24.

[120]中华医学会传染寄生虫病学会.病毒性肝炎诊断标准.中华传染病杂志.1995,13(4):242.

[121]中华医学会肝病学会,中华医学会感染病学会.慢性乙型肝炎防治指南.中国新药与临床杂志.2006,25(2):147-160.

[122]陈艳,周学东.小柴胡汤联合拉米夫定治疗慢性乙型肝炎合并肺结核 60 例.河南中医.2014,34(07):1291-1292.

[123]陈小中.小柴胡汤加味治疗慢性乙型肝炎肝纤维化 32 例.湖南中医杂志.2017,33(05):57-58.

[124]康昭恒.小柴胡汤联合恩替卡韦治疗慢性乙型肝炎的随机对照研究.实用中西医结合临床.2016,16(02):71-72.

[125]刘乐.小柴胡汤治疗慢性乙肝临床治疗的效果探讨.医药前沿.2017,7(23):351-352.

[126]毛果,蔡光先,陈斌.基于中医证候疗效评价的慢性乙型肝炎肝郁脾虚证最佳方药的探究.湖南中医杂志.2014,30(08):12-15.

[127]王成东.小柴胡汤联合抗病毒药物在乙肝治疗中的应用探讨.实用中西医结合临床.2013,13(07):11-13.

[128]王莹.小柴胡汤联合西药对慢性乙型肝炎患者肝功能及 HBV-DNA 拷贝数的影响.新中医.2016,48(07):74-76.

[129]吴昊星.小柴胡汤联合拉米夫定治疗乙型肝炎肝硬化疗效观察.世界临床医学.2015,9(6):114.

[130]赵春宏.加味小柴胡汤联合阿德福韦酯治疗 HBeAg 阳性慢性乙型肝炎 40 例.河南中医.2014,34(10):1924-1925.

[131] The GRADE Working Group 2013. GRADE handbook for grading quality of evidence

- and strength of recommendations. //gdt.gradepr.org/app/handbook/handbook.html.
- [132] Schünemann HJ. Interpreting GRADE's levels of certainty or quality of the evidence: GRADE for statisticians, considering review information size or less emphasis on imprecision? *Journal of Clinical Epidemiology*. 2016,75:6-15.
- [133] Imberger G, Vejlbj AD, Hansen SB, et al. Statistical multiplicity in systematic reviews of anaesthesia interventions: a quantification and comparison between Cochrane and non-Cochrane reviews. *Plos One*. 2011,6(12):e28422.
- [134] Robinson N, Lorenc A, Ding W, et al. Exploring practice characteristics and research priorities of practitioners of traditional acupuncture in china and the EU-A survey. *Journal of Ethnopharmacology*. 2012,140:458-68.
- [135] Woo JM, Park EJ, Lee M, et al. Changes in attitudes toward and patterns in traditional Korean medicine among the general population in SouthKorea: a comparison between 2008 and 2011. *BMC Complementary and Alternative Medicine*. 2014,14:436.
- [136] Xia Y, Cao XD, Wu GC, et al. *Acupuncture Therapy for Neurological Diseases: A Neurobiological View*. Beijing: TsingHua University Press, 2010.
- [137] Hopton AK, Curnoe S, Kanaan M, et al. Acupuncture in practice: Mapping the providers, the patients and the settings in a national cross-sectional survey. *BMJ Open*. 2012,2(1):e000456.
- [138] Baldry P. *Acupuncture, Trigger Points And Musculoskeletal Pain: A Scientific Approach To Acupuncture For Use By Doctors And Physiotherapists In The Diagnosis And Management Of Myofascial Trigger Point Pain*. Churchill Livingstone: Elsevier, 2005.
- [139] Biersack HJ, Freeman LM. *Clinical Nuclear Medicine*. Berlin Heidelberg New York: Springer, 2007.
- [140] Longhurst JC. Defining meridians: a modern basis of understanding. *Journal of Acupuncture and Meridian Studies*. 2010,3(2):67-74.
- [141] Lin YC, Hsu SE. *Acupuncture for Pain Management*. Heidelberg, Dordrecht, London, New York: Springer, 2014.
- [142] Yang ES, Li PW, Nilius B, et al. Ancient Chinese medicine and mechanistic evidence of acupuncture physiology. *Pflügers Archiv: European Journal of Physiology*. 2011,462(5):645-53.
- [143] Wang H, Du Y. *Acupuncture*. Beijing: Chinese Press of Traditional Chinese Medicine, 2013.
- [144] 孔胜利,张维彬,李成林,等.大剂量乙肝疫苗穴位注射治疗 HBV 携带者的临床研究. *中西医结合肝病杂志*.2005,(05):9-10.
- [145] Han XH, Wang MY, Song SJ, et al. Clinical research of application of moxibustion on Ganshu (BL18) and Qimen (LR14) for chronic hepatitis B. *Zhong Yi Zhong Yao*. 2011,49(19):86-8.

- [146] He YX, Cui N, Li YY. Interferon acupoint manipulation for treatment chronic hepatitis B. *Modern Journal of Integrated Traditional Chinese and Western Medicine*. 2013,22:2358-60.
- [147] Wang YM, Yao LL, Han XH, et al. Clinical study on treatment of chronic hepatitis B with adefovir diptergent combined with moxibustion liver Yuhe Xuemen point. *China Practical Medical*. 2014,9(15):148-9.
- [148]高婷婷,胡东胜. 艾灸治疗乙型肝炎肝硬化患者失眠的疗效观察. *中国肝脏病杂志(电子版)*. 2016,8(02):123-126.
- [149] Xu LF, Yu ZC, Shen MH, et al. Effect of moxibustion on immune function of 60 patients with esophageal carcinoma after radiotherapy. *Journal of Nanjing University of Chinese Medicine*. 2008,01:12-4.
- [150] Zhou F, Huang JY, Xue L. Observation on the efficacy of combined acupuncture and drugs for leukopenia after chemotherapy. *Practical Preventive Medicine*. 2010,17(06):1168-70.
- [151] Zhao CL, Peng LJ, Zhang ZL. Effect of acupuncture and moxibustion on peripheral blood T lymphocyte subsets and NK cell activity in patients with colon cancer liver metastases. *Chinese Acupuncture & Moxibustion*. 2010,30(01):10-2.
- [152] Cao QB, Liu Z, Geng QL, et al. Effect of acupuncture on immune function after laparoscopic surgery. *Modern Journal of Integrated Traditional Chinese Medicine*. 2011,20(1):1-2.
- [153] Zhang SY, Du YQ. Effect of warm acupuncture on gastrointestinal function and immune function in postoperative patients with colon cancer. *Chinese Acupuncture & Moxibustion*. 2011,31(06):513-7.
- [154] Xie FX, Tian YF. Effect of medicinal cake-partitioned moxibustion on SOD and MDA in monoerythrocytic cells of middle-aged and elderly people. *Journal of Clinical Acupuncture and Moxibustion*. 2010,26(10):43-44.
- [155] Lu DJ. Effect of moxibustion on peripheral blood T lymphocytes and NK cells in patients with chronic myeloid leukemia. *Yunnan Journal of Traditional Chinese Medicine and Materia Medica*. 2010,31(4):42-3.
- [156] Ding YH, Gu CY, Shen LR, et al. Effect of acupuncture combined with general anesthesia on perioperative T lymphocyte immunity in patients undergoing laparoscopic cholecystectomy. *Jilin Journal of Traditional Chinese Medicine*. 2011,31(01):80-8.
- [157] He CG, Cai CS. Clinical study on treatment of rheumatoid arthritis with acupuncture and moxibustion. *Journal of Clinical Acupuncture and Moxibustion*. 2010,26(10):43-4.

- [158] Lesnikar V. Epidemiology of hepatitis B and C in Croatia. *Acta Medica Croatica: Casopis Hrvatske Akademije Medicinskih Znanosti*. 2005,59(5):377-81.
- [159] Shin HR. Epidemiology of hepatitis C virus in Korea. *Intervirolgy*. 2006,49(1-2):18-22.
- [160] Nguyen VT, McLaws ML, Dore GJ. Highly endemic hepatitis B infection in rural Vietnam. *Journal of Gastroenterology and Hepatology*. 2007,22(12):2093-100.
- [161] Reynolds L, McKee M. Possible risks of transmission of bloodborne infection via acupuncture needles in Guizhou province, southwest China. *Journal of Alternative and Complementary Medicine: Research on Paradigm, Practice, and Policy*. 2008,14(10):1281-5.
- [162] Sulaiman HA, Julitasari, Sie A, et al. Prevalence of hepatitis B and C viruses in healthy Indonesian blood donors. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 1995,89(2):167-70.
- [163] Cheng J, Cao Y, Xia Y. From manual workshop to international standard maker: exploration on production standard of acupuncture needle by Chengjiang acupuncture school. *Chinese Acupuncture & Moxibustion*. 2015,35(2):189-93.
- [164] Yamashita H, Tsukayama H, White AR, et al. Systematic review of adverse events following acupuncture: the Japanese literature. *Complementary Therapies in Medicine*. 2001,9(2):98-104.
- [165] White A. A cumulative review of the range and incidence of significant adverse events associated with acupuncture. *Acupuncture in Medicine*. 2004,22(3):122-33.
- [166] Wang H, Du YH. *Acupuncture and Moxibustion*. Beijing: China Press of Traditional Chinese Medicine, 2012.
- [167] Bardy J, Finnegan JJ, Molassiotis A, Mackereth P. Providing acupuncture in a breast cancer and fatigue trial: the therapists' experience. *Complementary Therapies in Clinical Practice* 2015;21(4):217-22.
- [168] Litscher G. Ten years evidence-based high-tech acupuncture-a short review of peripherally measured effects. *Evidence-based Complementary and Alternative Medicine: eCAM*. 2009,6(2):153-8.
- [169] Jiang WL, Wei HJ, Guo ZY, et al. Changes in nitric oxide releases of the contralateral acupoint during and after laser acupuncture at bilateral same-name acupoints in human. *Evidence-based Complementary and Alternative Medicine: eCAM*. 2017,2017:5763458.
- [170] Yazdanpanahi Z, Ghaemmaghani M, Akbarzadeh M, et al. Comparison of the effects of dry cupping and acupressure at acupuncture point (BL23) on the women with postpartum low back pain (PLBP) based on short form McGill Pain Questionnaires in Iran: a randomized controlled trial. *Journal of Family & Reproductive Health*. 2017,11(2):82-9.

- [171] Ahmedov S, Filiz B. Effect of meridian acupuncture on aerobic performance of healthy young population: a randomized controlled study. *Journal of Alternative and Complementary Medicine* 2018;24(6):589-956. [DOI: 10.1089/acm.2017.0089].
- [172] Liang S, Christner D, Laux SD, et al. Significant neurological improvement in two patients with amyotrophic lateral sclerosis after 4 weeks of treatment with acupuncture injection point therapy using enercel. *Journal of Acupuncture and Meridian Studies*. 2011,4:257-61.
- [173] Lee SH, Kwon GS, Kang MS, et al. Comparative study on the effects of bee venom pharmacopuncture according to the treatment method for knee osteoarthritis. *Journal of Pharmacopuncture*. 2012,15(4):7-14.
- [174] Li K, Wei D, Li JK. Acupuncture plus acupoint-injection for treatment of cervical spondylosis of vertebral type. *Journal of Traditional Chinese Medicine*. 2008,28(4):243-4.
- [175] 张爱玲,吴莹,姜学亮.穴位注射治疗慢性乙型肝炎疗效分析. *中国针灸*.2005,(01):29-30.
- [176] Wang LQ, Wang WB, Fan JH. Evaluation and nursing of effect of Ginger moxibustion in the treatment of viral hepatitis. *Dang Dai Hu Shi*. 2017,3:147-8.
- [177] Kuge H, Mori H, Morisawa T, et al. Effect of different dosages of ST36 indirect moxibustion on the skin temperature of the lower legs and feet. *Medicines*. 2018,5(2):57.
- [178] Yun Y, Shin S, Kim KS, et al. Three cases of cutaneous warts treated with moxibustion. *Explore (NY)*. 2016,12(4):277-81.
- [179] Schlaeger JM, Stoffel CL, Bussell JL, et al. Moxibustion for cephalic version of breech presentation. *Journal of Midwifery & Women's Health*. 2018,63(3):309-22.
- [180] Zhang L. *Zhang Shi Yi Tong*. Shanghai: Shanghai Science Technology Publications, 1990.
- [181] Hsu WH, Ho TJ, Huang CY, et al. Chinese medicine acupoint herbal patching for allergic rhinitis: a randomized controlled clinical trial. *American Journal of Chinese Medicine*. 2010,38(4):661-73.
- [182] Lee SH, Chang GT, Zhang X, et al. Acupoint herbal patching for asthma: a systematic review and meta-analysis of randomized controlled trials. *Medicine (Baltimore)* 2016,95(2):e2439. [DOI: 10.1097/MD.0000000000002439].
- [183] ChiCTR-IOR-17011957. 穴位敷贴（疏肝健脾贴）联合抗病毒药物治疗肝郁脾虚型慢性乙型肝炎的疗效观察. www.chictr.org.cn/showproj.aspx?proj=20420 (first received 12 July 2017).

- [184] ChiCTR-IOR-17011962.耳穴磁疗联合中药塌渍治疗慢性乙型肝炎生存质量观察.
www.chictr.org.cn/showproj.aspx?proj=20433 (first received 12 July 2017).
- [185]黄晶晶,黄鸿娜,王振常.足三里穴位注射联合干扰素治疗慢性乙型肝炎临床观察.新中医,2011,43(11):89-91.
- [186]敬大成,钟森,欧阳海平,高爽,卢光亮,高崇勇.穴位注射联合聚乙二醇干扰素 α -2a治疗慢性乙型肝炎30例临床观察.湖南中医杂志.2016,32(6):97-99.
- [187]中华医学会肝病学会.中华医学会感染病学分会.慢性乙型肝炎防治指南(2010年版).临床肝胆病杂志.2011,16(1):113-128.
- [188]高月求,王灵召.慢性乙型肝炎中医诊疗专家共识(2012年1月).临床肝胆病杂志.2012,28(3):164-168.
- [189]廖芳莲.茵陈蒿贴治疗肝胆湿热型慢性乙型肝炎的临床研究.广州.广州中医药大学.2011.
- [190] 中华医学会肝病学会与中华医学会感染病学分会.慢性乙型肝炎防治指南.胃肠病学.2006,11(10):608-615.
- [191] 王志宏,岳蓓,林兰意,徐杰,于建武,吕淑兰. 针刺联合 α -干扰素治疗慢性乙型肝炎临床研究. 临床肝胆病杂志.2002,18(3):187-188.
- [192]王彬.脐部艾灸疗法对乙肝肝硬化的疗效观察及护理.光明中医. 2013,28(2):368-369.
- [193]中国中西医结合学会肝病专业委员会.肝纤维化中西医结合诊疗指南.中华肝脏病杂志.2006,14(11).
- [194]国家中医药管理局中医肝病重点专科写作组.中医肝病诊疗常规.2005,10:35-39.
- [195]张佩江,郭建中.针刺联合西药治疗慢性乙型肝炎临床观察. 辽宁中医药大学学报.2014,16(4):107-108.
- [196]中华医学会肝病学会,中华医学会感染病学分会.慢性乙型肝炎防治指南.中华传染病杂志.2005,23(6):421-429.
- [197]祝峻峰,王灵台.“水臌贴”敷脐联合托伐普坦等治疗慢性乙型肝炎肝硬化顽固性腹水的临床研究. 上海中医药杂志.2016,50(4):37-40.
- [198]王吉耀.现代肝病治疗学.上海:上海医科大学出版社,1999:244.
- [199] American Association for the Study of Liver Diseases. Practice Guidelines: Management of adult patients with ascites due to cirrhosis: an update. Hepatology (Baltimore, Md.).2009,49(6):2087-107.
- [200] Moore KP, Wong F, Gines P, et al. The management of ascites in cirrhosis: report on the consensus conference of the International Ascites Club. Hepatology (Baltimore, Md.). 2003,38(1):258-66.
- [201]邹逸天,华海清,郑清兴,等. 护肝拔毒巴布膏外敷联合常规疗法治疗慢性乙型肝炎50例临床研究.江苏中医药.2011,43(7):32-33.

- [202] 中华医学会传染病与寄生虫病学分会,肝病学会.病毒性肝炎的诊断标准.中西医结合肝病杂志.2001,11(1):56.
- [203] Jakobsen JC, Gluud C. The necessity of randomized clinical trials. *Br J Med Res.* 2013, 3(4):1453–68.
- [204] Silvio Garattini, Janus C. Jakobsen, Jørn Wetterslev, et al. Evidence-based clinical practice: Overview of threats to the validity of evidence and how to minimise them. *European Journal of Internal Medicine.* 2016, 32:13-21.
- [205] Kranz J, Schmidt S, Lebert C, et al. The 2017 Update of the German Clinical Guideline on Epidemiology, Diagnostics, Therapy, Prevention, and Management of Uncomplicated Urinary Tract Infections in Adult Patients. Part II: Therapy and Prevention. *Urol Int.* 2018,100:271–278.
- [206] Shore-Lesserson L, Baker RA, Ferraris V, et al. STS/SCA/AmSECT Clinical Practice Guidelines: Anticoagulation during Cardiopulmonary Bypass. *J Extra Corpor Technol.* 2018,50 (1): 5-18.
- [207] Rae-Grant A, Day GS, Marrie RA, et al. Comprehensive systematic review summary: Disease-modifying therapies for adults with multiple sclerosis: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. *Neurology.* 2018, 90 (17): 789-800.
- [208] Liu WM, Liu BY, Xie YM. The study of the evidence-based medicine application in the internationalization of Traditional Chinese Medicine. *Modernization of Traditional Chinese Medicine and Materia Medica.* 2006,(03): 42-46.
- [209] Song HK, Zhang X, Guo Y, et al. Formulation of a Methodological Guide to the Practice Guidelines for Diseases Prevention in Traditional Chinese Medicine. *Henan Traditional Chinese Medicine.* 2017,(12): 2066-2067.
- [210] The Cardiac Arrhythmia Suppression Trial (CAST) Investigators: Preliminary report: Effect of encainide and flecainide on mortality in a randomized trial of arrhythmia suppression after myocardial infarction. *N Engl J Med.* 1989, 321: 406-412.
- [211] Echt DS, Liebson PR, Mitchell LB, et al. Mortality and morbidity in patients receiving encainide, flecainide or placebo: The Cardiac Arrhythmia Suppression Trial. *N Engl J Med.* 1991,324:781-788.
- [212] Gluud C, Brok J, Gong Y, et al. Hepatology may have problems with putative surrogate outcome measures. *J Hepatol.* 2007,46(4):734–42.
- [213] Kent DM, Kitsios G. Against pragmatism: on efficacy, effectiveness and the real world. *Trials.* 2009,10: 48.
- [214] Higgins JPT, Green S. The Cochrane Handbook for Systematic Reviews of Interventions, Version 5.1.0. The Cochrane Collaboration. [Available from] <http://www.cochrane-handbook.org/>, 2011.

- [215] Dans AM, Dans L, Oxman AD, et al. Assessing equity in clinical practice guidelines. *J Clin Epidemiol.*2007,60(6): 540-546.
- [216] Evans T, Brown H. Road traffic crashes: Operationalizing equity in the context of health sector reform. *Inj Control Saf Promot.* 2003,10 (1-2): 11-12,.
- [217] Bornhöft G, Maxion-Bergemann S, Wolf U, et al. Checklist for the qualitative evaluation of clinical studies with particular focus on external validity and model validity. *BMC Med Res Methodol.* 2006,6: 56.
- [218] Atkins D, Eccles M, Flottorp S, et al. GRADE Working Group: Systems for grading the quality of evidence and the strength of recommendations I: Critical appraisal of existing approaches. The GRADE Working Group. *BMC Health Serv Res.* 2004,4:38.
- [219] Atkins D, Briss PA, Eccles M, et al. GRADE Working Group: Systems for grading the quality of evidence and the strength of recommendations II: Pilot study of a new system. *BMC Health Serv Res.* 2005,5:25.
- [220] Nagendran M, Pereira TV, Kiew G, et al. Very large treatment effects in randomised trials as an empirical marker to indicate whether subsequent trials are necessary: meta-epidemiological assessment. *BMJ.* 2016, 27 (355): i5432.
- [221] Deeks J, Higgins J, on behalf of the statistical methods group of the Cochrane collaboration. Statistical algorithms in Review Manager 5. [Available from] https://www.researchgate.net/publication/252093205_Statistical_Algorithms_in_Review_Manager_5, 2010.
- [222] Li G, Taljaard M, Van den Heuvel ER, et al. An introduction to multiplicity issues in clinical trials: the what, why, when and how. *Int J Epidemiol.* 2017, 46(2): 746-755.
- [223] Guyatt GH, Oxman AD, Kunz R, et al. What is “quality of evidence” and why is it important to clinicians? *BMJ.* 2008, 336: 995-998.
- [224] Qin XK, Li P, Han M, et al. Xiaohaihu Tang for treatment of chronic hepatitis B: A systematic review of randomized trials. *Journal of Chinese Integrative Medicine.* 2010,8(4): 312-320.
- [225] Hu XY, Zhong S, Chen J, et al. Effectiveness and safety of Xiao Chai Hu Tang plus interferon for chronic Hepatitis B: A systematic review. *Chin J Evid-based Med.* 2011, 11(3): 308-313.
- [226] 贾克明,彭晓君,邓燕林,等.小柴胡汤治疗慢性活动型肝炎的研究.解放军医学杂志.1990,15(2): 95-98.
- [227] 赵春宏.加味小柴胡汤联合阿德福韦酯治疗 HBeAg 阳性慢性乙型肝炎 40 例.河南中医.2014,34(10):1924-1925.
- [228] 张玉萍,蒋开平,何应伟.小柴胡冲剂合甘利欣治疗慢性乙型肝炎 44 例.实用医学杂志.2001,17(8): 778-779.
- [229] Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *Journal of Clinical Epidemiology.* 2009,62: 499-505.

- [230] Tajiri H, Kozaiwa K, Ozaki Y, et al. Effect of sho-saiko-to(xiao-chai-hu-tang) on HBeAg clearance in children with chronic hepatitis B virus infection and with sustained liver disease. *Am J Chin Med.* 1991,19(2): 121-129.
- [231] Sata M, et al. Therapeutic efficacy of IFN-Beta (Feron) in combination with Shosaikoto in patients with chronic active hepatitis B. *The Japanese Journal of Clinical and Experimental Medicine.* 1994,71: 814-820.
- [232] 廖芳莲.茵陈篙贴治疗肝胆湿热型慢性乙型肝炎的临床研究.广州:广州中医药大学.2011.
- [233] 陈果,骆建兴,扈晓宇.调神针法对慢性乙型肝炎患者腹泻型 IBS(肝郁脾虚证)疗效及血浆 SP、VIP 的影响. *世界华人消化杂志.* 2015,23(8): 1303-1307.
- [234] 张玲,王丽娟.猪苓多糖穴位注射配合恩替卡韦治疗慢性乙型肝炎 23 例疗效观察. *内蒙古中医药.*2016,35(08):103.
- [235] 张斌,徐乾.穴位注射华蟾素治疗活动性肝炎肝硬化的临床研究. *临床和实验医学杂志.*2006(08):1077-1078.
- [236] 刘士敬,朱倩.中药驱虫药联合干扰素治疗慢性乙肝 35 例疗效分析. *中国中医药科技.*1999(02):110-111.
- [237] 彭磐石.一贯煎加味配合针刺治疗慢性乙型病毒性肝炎的临床研究.湖南:湖南中医学院,2003.
- [238] 李晓杰,刘晓燕.督灸预防干扰素治疗慢性乙型肝炎副反应 60 例. *环球中医药.*2011,4(06):471-472.
- [239] Minozzi S, Ruggiero F, Capobussi M, et al. EBM, guidelines, protocols: knowledge, attitudes and utilization in the era of law on professional responsibility and safety of health care. *Recenti Prog Med.* 2018,109(5):294-306.
- [240] Goodarzi Z, Hanson HM, Jette N, et al. Barriers and Facilitators for Guidelines with Depression and Anxiety in Parkinson's Disease or Dementia. *Can J Aging.* 2018,37(2):185-199.
- [241] Zhao J, Han XJ, Wang LY, et al. Problems and countermeasures of development of clinical practice guideline of Chinese medicine. *Journal of Traditional Chinese Medicine.*2010,51(02): 119-121+141.
- [242] Savović J, Jones H, Altman D, et al. Influence of reported study design characteristics on intervention effect estimates from randomized controlled trials: combined analysis of meta-epidemiologic studies. *Health Technol Assess.* 2012,16: 1-82.
- [243] Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ.* 2009,338: b2393.
- [244] Little RJ, D'Agostino R, Cohen ML, et al. The prevention and treatment of missing data in clinical trials. *N Engl J Med.* 2012,367:1355- 60.
- [245] Chan AW, Hrobjartsson A, Haahr MT, et al. Empirical evidence for selective reporting

of outcomes in randomized trials: comparison of protocols to published articles. *JAMA*. 2004, 291: 2457-65.

[246] Dwan K, Altman DG, Arnaiz JA, et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS One*. 2008,3:e3081.

[247] Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Int Med*. 2010,152: 726-32.

[248] Chow SC, Wang H, Shao J. Sample size calculations in clinical research, second edition. Chapman & Hall/CRC Biostatistics Series: Chapman and Hall/CRC, 2007.

[249] Jennison C, Turnbull BW. Repeated confidence intervals for group sequential clinical trials. *Control Clin Trials*.1984, 5: 33-45.

[250] Jakobsen JC, Wetterslev J, Winkel P, et al. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. *BMC Med Res Methodol*. 2014,14:120.

[251] Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*. 2009, 339: b2535

[<http://www.prisma-statement.org>].

[252] Griffin NF, Melanie C, John W, et al. Composite outcomes in randomized trials - greater precision but with greater uncertainty? *JAMA*. 2003,289:2554-9.

[253] Phillips A, Haidiquet V. ICH E9 guideline 'statistical principles for clinical trials': a case study. *Stat Med*. 2003,22(1):1-11 [discussion 13-17].

[254] Isabelle Boutron, Douglas G. Altman, David Moher, et al. CONSORT Statement for Randomized Trials of Nonpharmacologic Treatments: A 2017 Update and a CONSORT Extension for Nonpharmacologic Trial Abstracts. *Annals of Internal Medicine*. 2017, 167 (1): 40-47.

[255] Kjaergard LL, Nikolova D, Gluud C. Randomised trials in Hepatology: predictors of quality. *Hepatology*. 1999, 30: 1134-1138.

[256] Moher D, Schulz KF, Altman D, for the CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *J Am Med Assoc*. 2001,285:1987-1991.

[257] Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med*. 2001,134:663-694.

[258] Garattini S, Bertele V. Non-inferiority trials are unethical because they disregard patients' interests. *Lancet*. 2007,370(9602):1875-7.

[259] Hirayama C, Okumura M, Tanikawa K, et al. A multicenter randomized controlled clinical trial of Shosaiko-to in chronic active hepatitis. *Gastroenterol Jpn*. 1989,24(6): 715-719.

[260] Tajiri H, Kozaiwa K, Ozaki Y, et al. Effect of sho-saiko-to(xiao-chai-hu-tang) on HBeAg clearance in children with chronic hepatitis B virus infection and with sustained liver

disease. *Am J Chin Med.*1991,19(2): 121-129.

[261] 王崇国,李旭红,黄水仙,等.自体 LAK 细胞回输加小柴胡汤或白细胞介素 II 治疗慢性乙型肝炎 173 例疗效分析. *临床肝胆病杂志.*1992,8(4):191-194.

[262] Sata M, et al. Therapeutic efficacy of IFN-Beta (Feron) in combination with Shosaikoto in patients with chronic active hepatitis B. *The Japanese Journal of Clinical and Experimental Medicine.* 1994,71: 814-820.

[263] 吴玮,李彦恒,彭真基,等.自体 LAK 细胞回输加小柴胡汤治疗慢性乙型肝炎 32 例. *中西医结合肝病杂志.*1994,(04):31.

[264] 余安清,王克奇,张文鼎,等.加味小柴胡汤与病毒唑联用治疗慢性乙型肝炎. *浙江中西医结合杂志.*2000,(03):10-12.

[265] 黄和冲.干扰素联合小柴胡冲剂治疗慢性乙肝 40 例疗效分析. *苏州医学院学报.*2001,(05):561-562.

[266] 李峥,廖红华,吴梅江. γ 干扰素与小柴胡汤联合抗肝纤维化的研究. *中西医结合肝病杂志.*2001,(S1):95.

[267] 李旭红,王崇国. α -1b 干扰素联合小柴胡汤治疗慢性乙型肝炎——附 56 例疗效分析. *临床肝胆病杂志.*2001,(03):187.

[268] 张玉萍,蒋开平,何应伟.小柴胡冲剂合甘利欣治疗慢性乙型肝炎 44 例. *实用医学杂志.*2001,17(8):778-779.

[269] 熊凡,孙静,熊伟.干扰素联合小柴胡汤治疗肝纤维化的临床观察. *湖北中医杂志.*2003,25(10):10-11.

[270] 孙晓娜,孙俊波.小柴胡汤联合甘利欣注射液治疗肝炎肝纤维化 60 例. *中国中医药信息杂志.*2004,11(6):534-535.

[271] Zhang SJ, Chen ZX, Lao SX, Huang BJ. Effect of Hejie decoction on T cell immune state of chronic hepatitis B patients. *World J Gastroenterol.* 2004,10(10):1436-1439.

[272] 洪新田.小柴胡汤与清开灵联用治疗慢性乙型病毒性肝炎的临床观察. *湖北中医杂志.*2005,(06):23-24.

[273] 孙阳.小柴胡汤治疗慢性乙型病毒性肝炎 32 例. *黑龙江中医药.*2005,(03):34.

[274] 薄其兰,杜雯雯.干扰素联合小柴胡汤治疗慢性乙型肝炎 50 例疗效观察. *山东医药.*2006,46(8): 82-83.

[275] 单秀民,于征.小柴胡汤对干扰素流感样反应的影响. *山东中医杂志.*2006,25(8):528-529.

[276] 史文英,张军勤.干扰素联合小柴胡汤治疗慢性乙型病毒性肝炎. *湖北中医杂志.*2006,28(12):35.

[277] 魏开效.小柴胡汤制剂对慢性乙型肝炎患者的临床试验研究. *中华实验和临床病毒学杂志.*2007,21(3):294-295.

[278] 陈小桃,文顺喜,黄坚灵,等.阿德福韦酯联合小柴胡汤治疗乙肝肝纤维化的研究. *基层医学论坛.*2008,12(9):786-788.

- [279]何炳福,周健康,沈华江.小柴胡片联合干扰素治疗慢性乙型肝炎肝纤维化的临床研究.中国中西医结合消化杂志.2008,16(6):384-387.
- [280]李正秋.小柴胡汤联合拉米夫定治疗慢性乙型病毒性肝炎的临床观察.河北中医.2009,31(5):709-710.
- [281]王礼凤.拉米夫定联合小柴胡汤治疗慢性乙肝临床观察[A].中华中医药学会仲景学说分会.仲景医学求真(续三).中华中医药学会仲景学说分会:中华中医药学会,2009:2.
- [282]杨艳娜.小柴胡汤治疗慢性乙型肝炎60例.内蒙古中医药.2009,(6):38-40.
- [283]邱冰,张沛怡,汤晓燕,等.小柴胡汤联合水飞蓟素治疗慢性乙型肝炎肝纤维化疗效分析.中国基层医药.2010,17(1):38-39.
- [284]田广俊,池晓玲,吴黎明,等.加减小柴胡汤对干扰素治疗慢性乙型肝炎常见不良反应的干预作用.现代医院.2010,10(9):49-50.
- [285]陈丹丹,龚作炯.小柴胡汤治疗慢性乙型肝炎肝纤维化的临床研究.现代中西医结合杂志.2011,20(31):3928-3930.
- [286]刘中景,王者令,孙晓慧,等.加味小柴胡汤对乙型肝炎病毒YMDD变异率的影响.中国中西医结合消化杂志.2011,19(4):218-220.
- [287]汪梦,汤雄,袁涛.小柴胡汤联合拉米夫定治疗HBeAg阳性慢性乙型肝炎疗效及其对乙肝病毒YMDD变异的影响.江西省第三次中西医结合肝病学术会议暨江西省中西医结合肝病新进展学习班.2012.
- [288]陈琳.小柴胡汤联合拉米夫定治疗慢性乙型肝炎35例.中国中医药现代远程教育.2013,11(23):79.
- [289]王成东.小柴胡汤联合抗病毒药物在乙肝治疗中的应用探讨.实用中西医结合临床.2013,13(7):11-12.
- [290]汪敏慧.小柴胡汤治疗慢性乙型肝炎肝纤维化的临床观察.医药前沿.2013,(33):357.
- [291]陈艳,周学东.小柴胡汤联合拉米夫定治疗慢性乙型肝炎合并肺结核60例.河南中医.2014,34(7):1291-1292.
- [292]毛果,蔡光先,陈斌.基于中医证候疗效评价的慢性乙型肝炎肝郁脾虚证最佳方药的探究.湖南中医杂志.2014,30(8):12-14.
- [293]王书明.中药治疗慢性乙型肝炎临床疗效分析.中国中医药远程现代教育.2014,12(13):32-33.
- [294]王美华,郑定容,程庆.加味小柴胡汤联合阿德福韦酯片治疗慢性乙型肝炎临床研究.社区医学杂志.2014,12(7):30-31.
- [295]程玲,秦文,武正菊.阿德福韦酯联合小柴胡汤治疗乙肝肝纤维化的临床分析.内蒙古中医药.2015,6:57.
- [296]姜方兴.小柴胡汤合升降散在慢性乙型肝炎中的临床价值.医学美学美容.2015,4:209-210.
- [297]李洁霞.小柴胡汤联合阿德福韦治疗慢性乙型肝炎的临床观察.大家健康.2015,9(6):325.
- [298]吴昊星.小柴胡汤联合拉米夫定治疗乙型肝炎肝硬化疗效观察.世界临床医

- 学.2015,9(6):113-114.
- [299]曾彬,胡阳黔,姜红梅,刘祖明.小柴胡汤加减治疗慢性乙型肝炎肝硬变 60 例.河南中医.2015,35(6):1284-1286.
- [300]周晓辉.小柴胡汤联合五灵肝复胶囊治疗慢性乙型肝炎的临床疗效及安全性.河北中医.2015,21(4):659-671.
- [301]周伟芳.中药小柴胡汤佐治慢性乙型肝炎的效果评价.内蒙古中医药.2015,8: 41-42.
- [302]康昭恒.柴胡汤联合恩替卡韦治疗慢性乙型肝炎的随机对照研究.实用中西医结合临床.2016,16(2):71-72.
- [303]王莹.小柴胡汤联合西药对慢性乙型肝炎患者肝功能及 HBV-DNA 拷贝数的影响.新中医.2016,48(7):75-76.
- [304]邹增城,陈鸿杰,杨宏志.小柴胡汤加减联合恩替卡韦治疗肝郁脾虚型慢性乙型病毒性肝炎的疗效观察.中医临床研究.2016,8(24):23-25.
- [305]陈小中.小柴胡汤加味治疗慢性乙型肝炎肝纤维化 32 例.湖南中医杂志.2017,33(5):57-58.
- [306]胡冬青,魏艳丽.加味小柴胡汤治疗慢性乙型肝炎及其组方机理的研究.2017,20:65-66.
- [307]林金宝,张一耕,林培坚.分析小柴胡汤与恩替卡韦联合用药对慢性乙型肝炎的临床疗效.中国医药指南.2017,15(20):204-205.
- [308]刘乐.小柴胡汤治疗慢性乙肝临床治疗的效果探讨.医药前沿.2017,7(23):351-352.
- [309]徐占英.耳穴压豆治疗乙肝 53 例.针灸临床杂志.1996.12(4):30.
- [310]陈家福,陈民,赵斌,王英.针刺对 HBV 携带者免疫功能影响的对比观察.湖南中医杂志.1998,14(5):20-21.
- [311]计沙,陈瑞玲.中西药配合针刺穴位治疗慢性乙型肝炎 46 例.淮海医药.1999. 17(2):33.
- [312]程启华,刘翔,朱华春.针刺配合中药治疗乙肝伴胁痛 32 例疗效观察.国际传统医药大会. 2000:468.
- [313]李萍,姜华静.三环穴位注射华蟾素对新肝功能的影响.全国首届慢性病毒性肝炎诊疗技术新进展研讨会暨第二届免疫辨证三环、三联疗法学术交流会.2001
- [314]孙茂彤,鞠传军,仇裕丰.穴位注射治疗无症状乙肝病毒携带者 84 例临床观察.南京中医药大学学报(自然科学版).2000,16(3):174-175.
- [315]马汉骞,邱美香,朱其守,等.经络穴位按摩治疗慢性乙型肝炎 50 例.中华实用中西医结合杂志. 2001, 14 (11):1.
- [316]唐建萍,钱铭.低频穴位电刺激治疗乙型肝炎的效果.中华国际医学杂志.2001, 1(1):74.
- [317]彭长林.灸药结合治疗慢性乙型肝炎 32 例.上海中医药杂志.2002,(1):27.
- [318]王志宏,岳蓓,林兰意,等.针刺联合 α -干扰素治疗慢性乙型肝炎临床研究.临床肝胆病杂志.2002,18(3):187-188.
- [319]杨莹洁,郑永志,胡巍.胸腺肽配合益肝 1 号治疗慢性乙型病毒性肝炎 100 例临床观察.

- 河北中医.2002.24(4):295-297.
- [320]李萍,于波,史玲,等.华蟾素穴位注射治疗慢性乙型肝炎临床研究.2003,3(11):800-801.
- [321]刘原龙,林华日.灯芯灸联合胸腺肽治疗抗 HBe 阳性慢性乙型肝炎疗效观察.针灸临床杂志.2003.19(9):39-40.
- [322]王坚,文睿,刘伟,刘新钰,等.中药复方 989 穴位注射抗慢性乙型肝炎纤维化作用的研究.辽宁药物与临床.2003,6(1):12-13.
- [323] Cheng JJ, Sun GJ. Observation on the effect of moxibustion on hyperbilirubinemia in hepatitis B type cirrhosis patients. World J.Acu-Moxi . 2001,14(4):12-15.
- [324]杜建霞,李芳,徐乾.苦参素穴位注射治疗慢性乙型重症肝炎.中国基层医药.2004,11(4):459-460.
- [325]沈建军.胸腺肽穴位注射治疗慢性乙型肝炎的疗效观察.中华实用中西医杂志.2004,4(17):1759-1760.
- [326]陈兰,钱玉平.针灸结合西医综合疗法治疗乙型肝炎合并肝内胆汁淤积 40 例—附西医综合疗法治疗 40 例对照.浙江中医杂志.2005,11:494.
- [327]杜建霞,李芳,徐乾.肝炎灵足三里穴位治疗慢性乙型肝炎血清 T 淋巴细胞亚群的变化.中华实用中西医杂志.2005,18(18):947.
- [328]黄加权,邓菊红,郑志想,等.穴位贴磁药对肝炎后肝硬化患者肝功能的保护.中国针灸.2005,25(9):613-615.
- [329]张武,李雄,苏海飞,等.干扰素 $\alpha - 2b$ 联合苦参素穴位注射治疗慢性乙型肝炎的临床观察.现代中西医结合杂志.2005,14(15):1998-1999.
- [330]刘文涛,陈海燕.针灸联合中药复方治疗慢性乙型肝炎临床观察.时珍国医国药.2006,17(8):1532-1533.
- [331]李宜川,黄瑾.足三里注射地塞米松治疗慢性重度乙型肝炎黄疸 28 例.中医外治杂志.2007,16(6):31-32.
- [332]苗爱军.拉米夫定联合胸腺肽穴位注射治疗慢性乙型肝炎 94 例疗效观察.山东医药.2007,47(26):81.
- [333]薛艳晖.联合疗法治疗慢性乙型肝炎肝纤维化 60 例临床观察.江苏中医药.2007,39(8):35-36.
- [334]张颖,俞海英.穴位注射胸腺肽 $\alpha 1$ 对慢性乙肝 T 细胞亚群的影响.中国微生态学杂志.2007,19(2):189-190.
- [335]孙浩.肝舒贴外敷治疗乙型肝炎肝硬化临床研究.南京:南京中医药大学.2008.
- [336]薛书奎.苦参素穴位注射治疗慢性乙型肝炎 37 例临床观察.山东医药.2008,48(25):24.
- [337]周敏,胡泰洪,丁虹,等.清肝解毒膏贴敷治疗慢性乙型肝炎 50 例.中医研究.2011,21(11):27-29.
- [338]邓良准,李锐.足三里注射苦参素治疗重度慢性乙型肝炎临床研究.中医外治杂志.2009,18(2):9-10.

- [339]阮永队,阮一帆,陈红梅,等.肝炎灵和黄茂注射液穴位注射治疗 YMDD 变异型慢性乙型肝炎的临床研究.新中医.2009.41(1):28-29.
- [340]邵志林,费新应,陈炎生,等.子午流注针法对减少慢性乙型肝炎 HBV-YMDD 变异的临床观察.上海针灸杂志.2009,28 (12):694-695.
- [341]谭善忠,徐乾,杨觉民,等.足三里穴位注射小剂量地塞米松小剂量地塞米松治疗肝衰竭早期疗效分析.中西医结合肝病杂志.2009.19(4):241-243.
- [342]张淑敏,刘敏,王萍.苦参素足三里穴位注射治疗慢性乙型肝炎的疗效观察与护理.中外健康文摘.2016.7(7):245-246.
- [343]韩秀华,王元敏,宋少军,等.艾灸肝俞、期门穴治疗慢性乙型肝炎的临床研究.中国现代医生.2011,49(19):86-88.
- [344]黄早秀,秦健,肖燕.不同方法注射乙型肝炎免疫球蛋白阻断乙型肝炎病毒母婴传播的临床研究.当代医学.2011,17(239):110-111.
- [345]李研.中药穴位贴敷治疗慢乙肝 108 例. 中华中医药学会内科分会.中华中医药学会第十三届内科肝胆病学术会议论文汇编.中华中医药学会内科分会:中华中医药学会.2008:3.
- [346]刘汶,王俊丽,戚团结,刘敏,齐京,李杰,徐洁,陈言.穴位贴敷配合清解毒方治疗乙型肝炎湿热证的临床研究.北京中医药.2011,30(10):726-730.
- [347]陆原,陈庆,邵铭,等.疏肝化瘀方穴位贴敷治疗慢性乙型肝炎胁痛 33 例总结.湖南中医杂志.2011,27(6):21-22.
- [348]杨智海,李旭英,黄学军. α - 2b 干扰素联合黄芪注射液治疗 HBeAg 阳性慢性乙肝 64 例.中医药信息.2011,28(6):94-95.
- [349]朱勤,杨育林,杨小乐,等.黄芪针穴位注射治疗慢乙肝临床研究与中医外治理论探讨.中华中医药学刊.2011.29(12):2695-2697.
- [350]邹逸天,华海清,郑清兴,等.护肝拔毒巴布膏外敷联合常规疗法治疗慢性乙型肝炎 50 例临床研究.江苏中医药.2011.43(7):32-33.
- [351]方红,高艳丽.隔药灸脐法配合穴位贴敷治疗慢性乙肝的疗效观察.光明中医.2012,27(6):1173-1174.
- [352]韩捷,顾亚娇.脐火疗法治疗阴黄(乙肝肝硬化) 15 例.中国针灸.2012,32 (6):490.
- [353]吴永斌,王修锋.苦参素穴位注射治疗慢性乙型肝炎 30 例.陕西中医.2012.33 (1):12-14.
- [354]张伟伟,韩正宇.穴位注射 α -干扰素治疗慢性乙型肝炎临床疗效观察.中国保健营养.2012,(07):1767-1768.
- [355]卓晓英,肖春玲,李淑惠,等.拉米夫定联合子午流注针法治疗慢性乙型肝炎的疗效观察.护士进修杂志.2012.27(11):1039-1041.
- [356]何玉霞,崔娜,李媛媛. α -干扰素穴位手法注射治疗慢性乙型肝炎的疗效观察.现代中西医结合杂志. 2013,22(21):2358-2360.
- [357]王彬.脐部艾灸疗法对乙肝肝硬化的疗效观察及护理.光明中医.2013.28(2): 368-369.

- [358]陈洁真,池晓玲,吴树铎,等. 中药穴位注射联合聚乙二醇干扰素 α -2a 治疗慢性乙型肝炎的临床研究. 中华中医药学刊. 2014, 32(8): 1965-1967.
- [359]胡洁华,钱雪梅. 苦参素穴位注射联合恩替卡韦治疗 HBeAg 阳性慢性乙型肝炎临床观察. 中国医药导刊. 2014, 16(5): 821-822.
- [360]雷任国,汤红丽,蓝日春,等. 针灸干预干扰素治疗慢性乙型肝炎 18 例疗效观察. 内蒙古中医药. 2014(30): 58-59.
- [361]李夏元. 用中药穴位贴敷法治疗肝病的临床疗效观察. 当代医药论丛. 2014, 12(4): 140-141.
- [362]汤红丽,雷任国,苏春雄,等. 民族外治法干预干扰素治疗慢性乙型肝炎副作用 36 例临床观察. 亚太传统医药. 2014, 10(24): 12-13.
- [363]王元敏,姚兰兰,韩秀华,等. 阿德福韦酯联合艾灸肝俞和期门穴治疗慢性乙型肝炎的临床研究. 中国实用医药. 2014, 9(15): 148.
- [364]魏文迪. 艾灸法联合拉米夫定治疗肝郁脾虚型慢性乙型肝炎的初步临床观察. 山东中医药大学. 山东. 2011.
- [365]张佩江,郭建中. 针刺联合西药治疗慢性乙型肝炎临床观察. 辽宁中医药大学学报. 2014, 16(4): 107-108.
- [366]张琴,郭卉. 以针灸为主中西医结合对乙肝后肝硬化细胞免疫功能的影响. 2014, (33): 29-30.
- [367]陈述,邹必英. 胸腺因子关元穴位注射联合恩替卡韦抗病毒的早期疗效观察. 江西省中西医结合学会、南昌市第九医院: 江西省中西医结合学会, 2015: 5.
- [368]胡先静. 生物信息红外肝病治疗仪合肝俞穴贴敷治疗慢性乙型肝炎的疗效观察与护理. 医学信息. 2015, 28(35): 173-174.
- [369]李世全,段小华,杨守平,等. 针刺联合甘利欣注射液治疗慢性乙型肝炎疗效观察. 中国公共卫生. 2015, 31: 277-278.
- [370]刘春光,陈少颖,农桂东,等. 中药穴位贴敷联合干扰素治疗慢性乙型肝炎的临床观察. 内蒙古中医药. 2015. 4: 92-93.
- [371]赵洁,贾建伟,雷金艳,等. 替比夫定序贯治疗对慢性乙型肝炎患者免疫应答及肝功能的影响. 中国中西医结合学会传染病专业委员会. 全国第 6 届中西医结合传染病学术会议论文汇编. 中国中西医结合学会传染病专业委员会: 中国中西医结合学会, 2015: 1.
- [372]王瑶,杨帆,邵冬姍. 中药穴位敷贴治疗乙型肝炎肝硬化水臌的疗效观察. 实用中西医结合临床. 2015. 15 (7): 72-73.
- [373]章淑萍,杨姝姚,邓志文,等. 香丹穴位注射对慢性乙型肝炎疗效的影响. 西部中医药. 2015, 28(11): 115-117.
- [374]蔡春江,梁燕,徐任. 香丹穴位注射对慢性乙型肝炎疗效的影响. 中华中医药学刊. 2016. 34(4): 1017-1020.
- [375]程毅凡. 艾灸然谷穴治疗肝硬化腹胀 50 例的临床疗效观察. 内蒙古中医药. 2016. 9: 123-124.
- [376]敬大成,钟森,欧阳海平,等. 穴位注射联合聚乙二醇干扰素 α -2a 治疗慢性乙型肝炎 30

- 例临床观察.湖南中医杂志.2016.32(6):97-99.
- [377]陆军平,李胜强,张赐敏.内科常规疗法配合针刺艾灸治疗顽固性肝硬化腹水的临床研究.四川中医.2016.34(8):160-161.
- [378]陆玲庆,樊松琴,张晶,等.涌泉穴贴敷中药对慢乙肝失眠症患者的干预效果研究.湖南中医药大学学报(综合二).2016.36(S2):1383.
- [379]张娟.针刺穴位治疗慢性肝病胃肠功能障碍的临床研究.湖北中医药大学.湖北.2016.
- [380]祝峻峰,王灵台.“水臌贴”敷脐联合托伐普坦等治疗慢性乙型肝炎肝硬化顽固性腹水的临床研究.上海中医药杂志.2016.50(4):37-40.
- [381]王黎茜,王维波,刘莉,等.隔姜艾灸辅助治疗病毒性肝炎的效果评价及护理.当代护士.2017(03):147-148.
- [382]仵倚,章新友,周小玲,等.中药穴位贴敷法治疗慢性乙肝有效性的系统评价.江西中医药.2017,48(04):64-67.

附表

附表 1 小柴胡汤干预慢性乙型肝炎检索策略

数据库	时间跨度	检索策略
Cochrane Hepato-Biliary Group Controlled Trials Register	2018.04	(Xiao-Chai-Hu or xiaochaihu or XCHT or chai*u or bupleur* or sho-sai-ko* or shosaiko* or minor bupleurum decoction*) AND ((hepatitis B or hep B or hbv) and chronic)
Cochrane Central Register of Controlled Trials (CENTRAL) in the Cochrane Library	2018.04	#1 (Xiao-Chai-Hu or xiaochaihu or XCHT or chai*u or bupleur* or sho-sai-ko* or shosaiko* or minor bupleurum decoction*) #2 MeSH descriptor: [Hepatitis B, Chronic] explode all trees #3 ((hepatitis B or hep B or hbv) and chronic) #4 #2 or #3 #5 #1 and #4
MEDLINE Ovid	1946-2018.04	1. (Xiao-Chai-Hu or xiaochaihu or XCHT or chai*u or bupleur* or sho-sai-ko* or shosaiko* or minor bupleurum decoction*).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] 2. exp Hepatitis B, Chronic/ 3. ((hepatitis B or hep B or hbv) and chronic).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] 4. 2 or 3 5. 1 and 4
Embase Ovid	1974-2018.04	1. (Xiao-Chai-Hu or xiaochaihu or XCHT or chai*u or bupleur* or sho-sai-ko* or shosaiko* or minor bupleurum decoction*).mp. [mp=title, abstract, heading word, drug trade name, original

title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word]
 2. exp chronic hepatitis B/
 3. ((hepatitis B or hep B or hbv) and chronic).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word]
 4. 2 or 3
 5. 1 and 4

LILACS (Latin American 1982-2018.04 (Xiao-Chai-Hu or xiaochaihu or XCHT or and Caribbean Health Science Information database) (Bireme) chai\$u or bupleur\$ or sho-sai-ko\$ or shosaiko\$ or minor bupleurum decoction) [Words] and ((hepatitis B or hep B or hbv) and chronic) [Words]

Science Citation Index 1900-2018.04 #3 #2 AND #1
 Expanded (Web of Science) #2 TS=((hepatitis B or hep B or hbv) and chronic)
 #1 TS=(Xiao-Chai-Hu or xiaochaihu or XCHT or chai*u or bupleur* or sho-sai-ko* or shosaiko* or minor bupleurum decoction*)

Conference Proceedings Citation Index - Science (Web of Science) 1990-2018.04 #3 #2 AND #1
 #2 TS=((hepatitis B or hep B or hbv) and chronic)
 #1 TS=(Xiao-Chai-Hu or xiaochaihu or XCHT or chai*u or bupleur* or sho-sai-ko* or shosaiko* or minor bupleurum decoction*)

中国知网(CNKI) 1979-2018.04 #1 ‘乙型肝炎’或‘乙肝’在摘要中
 #2 ‘小柴胡汤’或‘小柴胡颗粒’或‘小柴胡胶囊’或‘小柴胡片’在摘要中
 #3 ‘随机’在摘要中
 #4 ‘随机分组’在摘要中
 #5 #3 OR #4
 #6 #1 AND #2 AND #5

重庆维普(CQVIP) (VIP) 1989-2018.04 #1 ‘乙型肝炎’或‘乙肝’在摘要中
 #2 ‘小柴胡汤’或‘小柴胡颗粒’或‘小柴胡胶囊’或‘小柴胡片’在摘要中
 #3 ‘随机’在摘要中

#4 ‘随机分组’ 在摘要中
#5 #3 OR #4
#6 #1 AND #2 AND #5

万方 1990-2018.04 #1 ‘乙型肝炎’ 或 ‘乙肝’ 在摘要中
#2 ‘小柴胡汤’ 或 ‘小柴胡颗粒’ 或 ‘小柴胡胶
囊’ 或 ‘小柴胡片’ 在摘要中
#3 ‘随机’ 在摘要中
#4 ‘随机分组’ 在摘要中
#5 #3 OR #4
#6 #1 AND #2 AND #5

中国生物医学文献服务 1860-2018.04 #1 ‘乙型肝炎’ 或 ‘乙肝’ 在摘要中
系统 (Sinomed) #2 ‘小柴胡汤’ 或 ‘小柴胡颗粒’ 或 ‘小柴胡胶
囊’ 或 ‘小柴胡片’ 在摘要中
#3 ‘随机’ 在摘要中
#4 ‘随机分组’ 在摘要中
#5 #3 OR #4
#6 #1 AND #2 AND #5

附表 2 针灸疗法干预慢性乙型肝炎检索策略

数据库	时间跨度	检索策略
The Cochrane Hepato-Biliary Group Controlled Trials Register	2018.04	(acupunct* or acupress* or acupoint* or electroacupunct* or auriculotherap* or auriculoacupunct* or moxibust* or meridian* or patch*) AND (hepatitis B or hep B or HBV)
Cochrane Central Register of Controlled Trials (CENTRAL) in the Cochrane Library	2018.04	#1 MeSH descriptor: [Acupuncture] explode all trees #2 MeSH descriptor: [Acupuncture Therapy] explode all trees #3 acupunct* or acupress* or acupoint* or electroacupunct* or auriculotherap* or auriculoacupunct* or moxibust* or meridian* or patch* #4 #1 or #2 or #3 #5 MeSH descriptor: [Hepatitis B] explode all trees #6 hepatitis B or hep B or HBV #7 #5 or #6 #8 #4 and #7
MEDLINE Ovid	1946-2018.04	1. exp ACUPUNCTURE/ 2. exp ACUPUNCTURE THERAPY/ 3. (acupunct* or acupress* or acupoint* or electroacupunct* or auriculotherap* or auriculoacupunct* or moxibust* or meridian* or patch*).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] 4. 1 or 2 or 3 5. exp Hepatitis B/ 6. (hepatitis B or hep B or HBV).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

		7. 5 or 6
		8. 4 and 7
Embase Ovid	1974-2018.04	1. exp acupuncture/ 2. (acupunct* or acupress* or acupoint* or electroacupunct* or auriculotherap* or auriculoacupunct* or moxibust* or meridian* or patch*).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] 3. 1 or 2 4. exp hepatitis B/ 5. (hepatitis B or hep B or HBV).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] 6. 4 or 5 7. 3 and 6
LILACS (Bireme)	1982-2018.04	acupunct\$ or acupress\$ or acupoint\$ or electroacupunct\$ or auriculotherap\$ or auriculoacupunct\$ or moxibust\$ or meridian\$ or patch\$ [Words] and hepatitis B or hep B or HBV [Words]
Science Citation Index Expanded (Web of Science)	1900-2018.04	#3 #2 AND #1 #2 TS=(hepatitis B or hep B or HBV) #1 TS=(acupunct* or acupress* or acupoint* or electroacupunct* or auriculotherap* or auriculoacupunct* or moxibust* or meridian* or patch*)
Conference Proceedings Citation Index - Science (Web of Science)	1990-2018.04	#3 #2 AND #1 #2 TS=(hepatitis B or hep B or HBV) #1 TS=(acupunct* or acupress* or acupoint* or electroacupunct* or auriculotherap* or auriculoacupunct* or moxibust* or meridian* or patch*)
中国知网(CNKI)	1979-2018.04	#1 ‘乙型肝炎’ 或 ‘乙肝’ 在摘要中 #2 ‘针灸’ 或 ‘针刺’ 或 ‘刺激’ 或 ‘电针’

		或‘镭射’或‘穴位按压’或‘穴位注射’或‘艾灸’或‘贴敷’在摘要中
		#3 ‘随机’在摘要中
		#4 ‘随机分组’在摘要中
		#5 #3 OR #4
		#6 #1 AND #2 AND #5
重庆维普(CQVIP)	1989-2018.04	#1 ‘乙型肝炎’或‘乙肝’在摘要中
		#2 ‘针灸’或‘针刺’或‘刺激’或‘电针’或‘镭射’或‘穴位按压’或‘穴位注射’或‘艾灸’或‘贴敷’在摘要中
		#3 ‘随机’在摘要中
		#4 ‘随机分组’在摘要中
		#5 #3 OR #4
		#6 #1 AND #2 AND #5
万方	1990-2018.04	#1 ‘乙型肝炎’或‘乙肝’在摘要中
		#2 ‘针灸’或‘针刺’或‘刺激’或‘电针’或‘镭射’或‘穴位按压’或‘穴位注射’或‘艾灸’或‘贴敷’在摘要中
		#3 ‘随机’在摘要中
		#4 ‘随机分组’在摘要中
		#5 #3 OR #4
		#6 #1 AND #2 AND #5
中国生物医学文献服务系统 SinoMed	1860-2018.04	#1 ‘乙型肝炎’或‘乙肝’在摘要中
		#2 ‘针灸’或‘针刺’或‘刺激’或‘电针’或‘镭射’或‘穴位按压’或‘穴位注射’或‘艾灸’或‘贴敷’在摘要中
		#3 ‘随机’在摘要中
		#4 ‘随机分组’在摘要中
		#5 #3 OR #4
		#6 #1 AND #2 AND #5

综述一 高级别循证医学证据的评价与综合的重要性 和方法学进展

临床决策者必须运用最佳质量的临床研究证据为患者选择治疗措施。对临床随机对照试验进行的系统综述研究所得出的临床证据被公认为评价干预措施疗效的最高等级证据，并应用于临床实践指南的制定，指导临床决策者治疗方案的选择^[1-3]。

1. 临床随机对照试验的发展历史

西方医学之父希波克拉底首先提出，要将观察性研究引入医学^[4]。阿拉伯医生 Avicenna 提倡药物应进入人体试验，应要有比较和可重复性^[5]。1061 年中国的《本草图经》首次提出应用对照的思维来验证药物的疗效，其第一次提到对照试验“为评价人参的效果，需寻两人，令其中一人服食人参并奔跑，另一人未服人参也令其奔跑。未服人参者很快就气喘吁吁”^[6]。约 1740 年(清乾隆)《考证》中首次提出循证思维^[7]。英国医生 Lind 于 1753 年进行了一项早期临床试验，对 12 名患坏血病的海员进行对照试验，运用不同饮食干预坏血病海员 6 天后发现食用橙汁和柠檬汁的患者明显好转，能重新工作^[8]。1989 年丹麦医生 Fibiger 发表了著名的血清治疗白喉的半随机对照试验，运用交替分配对照试验的方法，比较皮下注射白喉血清与标准治疗，发现白喉血清能够明显降低患者的死亡率^[8]。1947 年英国医学研究委员会进行了世界上首次随机对照试验，其通过多中心随机试验，同时运用盲法，评估链霉素联合卧床休息对照单独卧床休息对肺结核患者的疗效，结果发现链霉素联合卧床休息可显著降低肺结核患者的病死率和 X 线片结果^[9]。该研究以其严格的计划、实施和报告被公认为临床试验史上的里程碑。

在中国第一个随机对照试验为苏德隆教授于 60 年代初期完成并发表的^[10]。中医药领域第一个随机双盲对照试验发表于 1983 年，为中药注射剂治疗心绞痛^[11]。尽管近十余年来随着临床试验发表的数量和比率在逐年增多，但试验质量仍然很低，极少有多中心的大样本随机试验^[12]。

随机对照试验的出现，是临床医学研究新纪元的里程碑。其出现可与显微镜的发明相媲美，也是循证医学证据的主要来源。

2. 随机对照试验的评价与综合的必要性

日常医疗工作需要以临床证据为依据作出临床决策。医生的医学知识和临床行为随着时间延长会逐渐过时，而传统的继续医学教育并不能提高临床医疗效能，因此在医疗

实践中需要提供新的证据以改进对病人的医疗质量,并进一步改进临床研究的质量^[13-14]。目前医学界面临着信息爆炸的问题,这导致临床医生在进行临床决策时,由于证据的不确定性,往往会导致医疗干预措施变异太大,医疗干预措施效果不确定等问题^[15]。因此,对于临床证据的评价与综合,可以提供尽可能免于偏倚的、对已知和未知信息的概括。临床决策者必须依据最高等级证据选择治疗方案,目前对临床随机对照试验进行评价,并对于高质量的随机对照试验的综合所得出的临床证据被公认为评价干预措施疗效的最高等级证据,并被应用于临床实践指南的制定^[1-3]。

3. 对随机对照试验进行评价与综合的方法学进展

1976年心理学家 Glass 首次提出了 meta 分析一词及其统计学分析方法^[16]。1982年 Thomas C Chalmers 提出了累计性 Meta 分析的概念,即将每一项新的随机试验结果累加到已知的针对某病某干预的随机临床试验 Meta 分析结果汇总^[17]。英国流行病学家 Archie Cochrane 在 1979 年提出并率先开展系统综述,对临床实践和研究产生了划时代的影响^[18]。系统综述是一种使用明确和清晰的方法对相关研究进行综合,使别人能够对其评论,批评或尝试对所获得的结论进行可重复的研究^[19]。在临床指南的制定过程中需要对证据的质量和推荐强度进行评估。运用系统综述的方法可以通过系统、明确的手段帮助人们减少或避免误差,解决意见分歧,促进信息的交流以满足需求,使指南制作过程公开、透明,还可以改变临床医生的医疗行为^[20]。目前在循证医学证据分级中,基于高质量的随机对照试验的系统综述或者 meta 分析被认为是最高等级证据^[2]。

20 世纪 70 年代末和 80 年代初, Sackett 和 Fletcher 等人将流行病学与医学统计学的原理和方法,有机地与临床医学结合起来,发展和丰富了临床研究方法学,创建了临床流行病学^[21]。Sackett 于 1984 年陆续在 JAMA 上发表系列文章,指导临床医生如何阅读医学文献,制定了评价有关病因、治疗、预后、诊断等文献的新标准,帮助临床医生得到有用的证据指导临床实践。1992 年, Sackett 领导的研究小组首次提出了循证医学的概念,1995 年其受英国政府的邀请到英国牛津,建立了英国循证医学中心,使循证医学的概念更加系统化,向全世界推广,以采用循证医学的方法和证据指导临床实践^[22]。于 1992 年首先在英国成立了 Cochrane 中心,1993 年成立了国际 Cochrane 协作网,帮助人们进行系统评价,并将系统评价结果通过电子光盘杂志和因特网分发给世界各地的临床医生、病人和卫生决策者,使循证医学的开展、证据的获得成为可能^[9]。

80 年代中后期,随着临床流行病学引入,国内开始对中华系列医学杂志上发表的临

床试验进行评价。90年代后期循证医学引入中国，对中国的临床医学杂志进行广泛的手工检索，并用循证医学标准对发表的大量临床试验文章进行了质量和方法学评价。国内外学者对中国进行的临床试验的方法学质量评价，结果发现：中国的临床试验存在发表偏倚，发表的临床试验几乎均报告了阳性结果；中国的临床试验尤其是中医药领域的试验得不到国际承认；最近发表的几篇有关中医药治疗肝炎的系统评价表明90年代末期发表的随机试验，质量仍未有显著提高。中国在随机临床试验的设计、进行、管理和监测、报告等方面存在问题，临床试验的技术方法与质量有待提高^[12]。

2000年后行业内学者逐渐引入系统综述方法研究中医药疗法治疗疾病的有效性和安全性^[23]。并采用Cochrane协作组的‘Risk of bias tool’对随机对照试验的偏倚风险进行评估。2000年GRADE分级理念由GRADE工作组提出，其建议临床干预措施的实施应基于对干预措施获益和风险的权衡^[24]。GRADE分级从研究的偏倚风险、证据的间接性、结果的不一致性和异质性、结果的不精确性、研究的发表偏倚五个方面对临床证据进行等级评估^[25,26]。时至今日，采用Meta分析方法的系统综述联合GRADE分级被逐渐引入中医药领域研究，并被应用于中医疗法治疗疾病的临床诊疗指南制定工作^[27]。

由于系统综述是将目前检索到的相关临床研究结果进行数据累积性分析，因此该过程中采用的重复性统计检验会导致Meta分析得到假阳性或假阴性结果，从而影响系统综述的可信性^[28,29]。丹麦哥本哈根临床试验中心提出将试验序贯分析TSA理念引入Meta分析中以最小化Meta分析的随机误差，并开发了TSA软件以实现这一操作^[30]。运用试验序贯分析（TSA）探索要获得“充分证据”所需要的所需样本量大小（IS），并对获得的证据进行统计显著性检验和无效性检验，以判断是否需要进行更多的临床试验以进一步提供证据^[31]。由于该研究方法较新，在中医药领域的应用尚处于起步阶段。

参考文献：

- [1] Silvio Garattini, Janus C. Jakobsen, Jørn Wetterslev, et al. Evidence-based clinical practice: overview of threats to the validity of evidence and how to minimise them. *Medicina fluminensis*. 2017, Vol. 53, No. 4, p. 424-438.
- [2] Straus SE, Richardson WS, Glasziou P, et al. Evidence-based medicine. How to practice and teach EBM. Edinburgh, UK: Churchill Livingstone, 2005.
- [3] Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. The Cochrane Collaboration. 2008.
- [4] 檀慧玲.论欧洲中世纪大学对近代自然科学发展的贡献.河北师范大学学报(教育科学

版).2016,18(04):61-66.

[5]李文辉,刘颖,Wang Xi Zee.西方医学简史回顾及其哲学思考.医学争鸣.2018,9(03):4-8.

[6]王晓晖,陈静,李静,等.中国临床试验实施质量控制的发展与变革.中国循证医学杂志.2018,18(08):776-782.

[7]刘建平.随机对照试验方法及进展.中国中西医结合学会循证医学专业委员会.临床科研常用方法及进展培训班会前培训材料.中国中西医结合学会循证医学专业委员会:中国中西医结合学会,2017:15.

[8]刘雅莉,谢琪,刘保延,等.临床试验百年历程概述.中国循证医学杂志.2016,16(11):1241-1249.

[9]喻佳洁,李琰,陈雯雯,等.循证医学的产生与发展:社会需求、学科发展和人文反思共同推动.中国循证医学杂志.2019,19(01):108-113.

[10]任涛,詹思延,沈霞,等.流行病学研究中的偏倚与混杂[J].中华流行病学杂志.2004(09):83-85.

[11]刘玥,高铸焯,付长庚,徐浩,史大卓,陈可冀.活血化瘀药物防治冠心病:循证与展望.中国循证医学杂志.2018,18(11):1145-1150.

[12]刘建平,曹卉娟.从循证医学看国内发表的中医药临床研究现状及改进策略[J].中国中西医结合杂志.2010,30(01):5-8.

[13]郑木明,林新宏,耿庆山,王思博.循证医学,临床医学的新模式.循证医学.2005,(05):52-56.

[14]张宝琴,马淑琴,李玲君,贺大林,边永娜.考研影响临床实习的调查分析[J].西北医学教育.2002,10(04):218-219.

[15]Diane M Korngiebel,Kenneth E Thummel,Wylie Burke,张晶,吴一波.实施精准医学:伦理挑战.中国医学伦理学.2018,31(09):1224-1226.

[16] Smith ML, Glass GV. Meta-analysis of psychotherapy outcome studies. Am Psychol.1977,32(9):752-60.

[17] Lau J, Antman EM, Jimenez-Silva J, et al. Cumulative meta-analysis of therapeutic trials for myocardial infarction. N Engl J Med.1992,327(4):248-54.

[18] Cochrane AL. Archie Cochrane in his own words. Selections arranged from his 1972 introduction to "Effectiveness and Efficiency: Random Reflections on the Health Services" 1972. Control Clin Trials. 1989,10(4):428-33.

[19] Zwarenstein M, Volmink J, Irwig L, et al. Systematic review: 'state of the science' health care decision-making. S Afr Med J. 1995,85(12):1266-1267.

[20] Durieux P1, Ravaud P, Dosquet P, et al. Effectiveness of clinical guideline implementation strategies: systematic review of systematic reviews. Gastroenterol Clin Biol. 2000, 24(11):1018-1025.

[21]闵婕,冯英明,张贺龙,姬统理.循证医学的发展历史及其指导下的诊断学教学.西北医

学教育,2007,(02):354-357.

[22]任为民,朱熙,王丹,温世浩.英国医学教育者专业标准概况与启示.解放军医院管理杂志,2016,23(01):56-58.

[23]刘为民,刘保延,谢雁鸣.循证医学在中医药国际化中的应用研究.世界科学技术,2006,(03):42-46.

[24] Rothwell PM. External validity of randomized controlled trials: 'to whom do the results of this trial apply? '.Lancet. 2005, 365: 82-93.

[25]吴宇侠,康德英,洪旗,等.临床随机对照试验的外部真实性及其评价的探讨.中华流行病学杂志. 2011,32:514-518.

[26] Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med. 2013, 158(3):200-7.

[27]宋海阔,张玄,郭义,等.制定《中医治未病实践指南》的方法学探讨.河南中医. 2017,(12):2066-2067.

[28] Borm GF, Donders AR. Updating meta-analyses leads to larger type I errors than publication bias. J Clin Epidemiol. 2009, 62:825-830.

[29] Lan KK, DeMets D. Discrete sequential monitoring boundaries for clinical trials. Biometrika. 1983:659-663.

[30] Jørn Wetterslev, Kristian Thorlund, Jesper Brok, et al. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. Journal of Clinical Epidemiology. 2008, 61:64-75.

[31] Jørn Wetterslev, Kristian Thorlund, Jesper Brok, et al. Estimating required information size by quantifying diversity in random-effects model meta-analyses. BMC Medical Research Methodology.2009, 9:86.

综述二 证据评价的工具现状及研制新的多维质量评价工具的必要性

根据循证医学临床证据等级分级,位于证据等级金字塔最顶层的是基于高质量随机对照试验的系统综述,以下依次是随机对照试验、队列研究、病例对照研究、病例系列研究、专家意见^[1]。在临床实践过程中,在应用临床证据之前,必须对获得的临床证据进行评估,其决定着目前获得的临床证据是否有效可用,及其可信度。目前对于不同类型的临床证据,其质量评估的工具各不相同。

1. 证据评价的工具量表现状

(1) 系统综述和 meta 分析的质量评价工具

目前学术界有许多用于评价系统综述和 meta 分析质量的工具。关于系统综述和 meta 分析的质量评价主要包括两方面内容:即研究的方法学质量评价和研究的报告质量评价。对于系统综合和 meta 分析的质量评价工具,大多围绕以下方面内容制定评价标准:①研究的问题,②研究制定的纳入、排除、诊断标准,③相关研究的检索以及纳入研究的筛选,④纳入研究的质量评价,⑤研究的数据处理分析,⑥关于研究结果的解释,⑦关于研究结果的讨论。

关于系统综述和 meta 分析的方法学质量评估工具较多且庞杂:1984 年 Richard J. Light 团队研发出的 Light Checklist^[2],1986 年由美国 Goldschmidt PG 团队研发的 Goldschmidt checklist^[3],1987 年由加拿大多伦多大学预防医学和生物统计学院研发的 L'Abbé checklist^[4],1987 年由美国 Mullen PD 团队研发的 Mullen checklist^[5],1987 年由 Mulrow CD 团队研发的 Mulrow checklist^[6],1988 年由加拿大 Oxman AD, Guyatt G 等人研发的 Oxman checklis^[7],1989 年由美国 Meinert C L 团队研发的 Meinert checklist^[8],1989 年由美国 Smith MC 等人研发的 Smith checklist^[9],1991 年由加拿大 Oxman AD, Guyatt G 等人研发的 OQAQ^[10],1992 年由澳大利亚 Wilson A 等人研发的 Wilson checklist^[11],1993 年由美国 Neely JG 等人研发的 Neely checklist^[12],1994 年由加拿大 Ohlsson A 等人研发的 Ohlsson checklist^[13],1994 年由 Oxman AD 等人研发的 Oxman checklist^[14],1995 年由荷兰 Assendelft WJJ 等人研发的 Assendelft scale^[15],1995 年由加拿大 Cook DJ 等人研发的 Cook checklist^[16],1995 年由法国 Nony P 等人研制的 Nony checklist^[17],1996 年由美国 Geller NL 等人研制的 Geller

checklist^[18], 1996 年美国流行病学项目办公室和疾病预防健康促进中心研制的 Thacker checklist^[19], 1997 年由 Greenhalgh T 等人研制的 Greenhalgh checklist^[20], 1998 年 Pogue JM 等人研究的 Pogue checklist^[21], 1999 年由 Auperin A 等人研制的 Auperin scale^[22], 1999 年由德国 Blettner M 等人研制的 Blettner checklist^[23]。以上这些纳入工具由从 5-101 数目不等的条目构成。而关于系统综述和 meta 分析报告质量评价工具主要有两个: 2005 年由加拿大渥太华医院研究机构研制的 PRISMA 清单^[24]以及 2007 年由阿姆斯特丹自由大学研制的 AMSTAR 量表^[25,26]。系统综述和 meta 分析的质量评价工具由于需要使用不断变化的研究内容和研究方法而需要不断更新, 除 PRISMA 清单是由 10 年前的 QUROM 清单更新而来, AMSTAR 是在纽约城市大学西奈山医学院临床试验所研制的 SQAC 和 OQAQ 量表基础上更新形成, 其余工具均未进行更新。目前 PRISMA 清单应用较为广泛, 且已被诸多杂志用于系统综述投稿时作者自查清单之一。

2000 年 GRADE 分级理念由 GRADE 工作组提出, 其建议临床干预措施的实施应基于对干预措施获益和风险的权衡。GRADE 分级从研究的偏倚风险、证据的间接性、结果的不一致性和异质性、结果的不精确性、研究的发表偏倚五个方面对系统综述获得的临床证据进行等级评估^[27]。

(2) 随机对照试验质量评价工具:

1996 年, Jadad 等人研制出 Jadad 量表用以评估随机对照试验的质量, 该量表主要对随机序列的产生、随机化隐藏、盲法、撤出与退出四个方面进行质量评估, 得分 1-3 分视为低质量研究, 4-7 分视为高质量研究^[28]。David Moher 等人于 2010 年发布 Consort 声明 2010 版, 用以规范随机对照试验的报告, 亦有学者用此声明来评估临床随机对照试验的质量^[29]。该小组还在 CONSORT 声明的基础上进行了扩展, 针对中医药领域、针灸领域研究分别提出了提出了 CONSORT 声明扩展板 (for TCM) 和 CONSORT 声明扩展板 (STRICTA)^[30]以及草药干预临床试验的 CONSORT 声明和非药物治疗干预的 CONSORT 声明等, 并于 2017 年进行了更新。2011 年 CASP 项目组提出了针对临床随机对照试验的 CASP checklist, 用以评估临床随机对照试验的质量^[31]。2011 年 Cochrane 协作组提出 Risk of bias tool, 通过随机分配产生、分配方案隐藏、对参与者及主要研究者的施盲、对研究结果评估者的施盲、结局指标数据的不完整性、选择性报告、其他偏倚等几个方面, 用以评估随机临床试验的质量^[32]。2017 年 Cochrane 肝胆病组模板 the Cochrane Hepato-Biliary Group Module 对 Risk of bias tool 的具体评价标准进行修改协定,

并在原来领域的基础上补充了营利偏倚^[33]。

(3) 评估干预措施效果的非随机研究的质量评价工具

评估干预措施效果的非随机研究是指其不使用随机化将单位（个体或群体）分配给比较组的干预措施效果研究。这包括在通常的治疗决定或人们的选择过程中发生分配的研究，这种研究通常被称为“观察性”研究，可包括队列研究，病例对照研究，受控前后研究，中断时间序列研究和对照试验，其中也包含研究的干预组使用不完全随机化的方法进行分配的研究（亦称为“准随机”研究）^[34]。2016年由 Jonathan AC Sterne 等人研制出 ROBINS-I 工具，以评估干预措施效果的非随机研究结果中偏倚风险，以比较两种或多种干预措施对健康的影响，被 Cochrane 协作组在全球范围内进行推广。

ROBINS-I 工具涵盖七个领域，前两个领域评估在干预措施实施之前相关问题，即由于混杂引起的偏倚和患者的选择产生的偏倚，第三个领域为干预措施的分类偏倚，其他四个领域评估在干预开始后产生的偏倚：即由于偏离预期的干预措施而产生偏差、由于缺失数据产生的偏倚、由于结局指标测量产生的偏倚和由于选择性报告产生的偏倚。对于前三个领域，干预措施效果的非随机研究的偏倚评估风险不同于随机试验的评估，因为随机化可以防止在干预开始之前出现的偏倚。然而，随机化并不能防止干预开始后出现的偏见。干预措施效果的非随机研究和随机试验中偏倚评估之间的最后四个领域存在实质性重叠^[37]。

此外，Cochrane 还推荐运用 Newcastle-Ottawa-Scale (NOS) 工具对队列研究和病例对照研究进行质量评价。其内容主要涵盖对研究对象的选择、组间可比性和结果测量三个领域，以积分的形式进行评估，满分共 9 分，得分越高，表明所评估的研究质量越好^[35, 36]。

(4) 病例系列研究质量评价工具

病例系列是无对照的观察性研究，涉及对多个患者同一种干预-疾病或结局的描述。病例系列由于未设置对照也未应用随机化，可能存在多种潜在的偏倚。2012 年加拿大卫生经济研究所制定了病例系列研究较为系统全面的质量评价工具，即 IHE 病例系列方法学质量评价工具，主要对病例系列研究质量的四个领域进行评估：研究目的、研究人群、干预与联合干预措施、结局测量、统计分析、结果与结论、利益冲突和资金来源^[37]。

(5) 定性研究质量评价工具

CASP (Critical Appraisal Skills Programme) 项目组提出 CASP checklist 用于

评价定性研究质量^[38],但该方法具有方法学局限性。2003年,英国政府首席社会研究员办公室提出UKGC-SRO工具用以评估定性研究质量^[39]。2007年,Goldsmith等人通过对Grade分级进行修改,过评估单个纳入研究的方法学局限性、研究结果的一致性、相关性和数据充分性对定性研究的系统综述总体质量进行分级^[40]。2015年,由挪威知识转化中心的Claire Glenton、Simon Lewin教授联合Cochrane协作网、Campbell协作网、GRADE工作组和世界卫生组织(WHO)等国际相关机构制定CERQual,专门用于评价定性研究系统综述的质量^[41],该方法也是从单个纳入研究的方法学局限性、研究结果的一致性、相关性和数据充分性这四个方面进行质量评估,并综合四个部分的评价结果,对系统综述证据的可信度给予极低、低、中、高四个不同等级的评价。

2. 研制新的多维质量评价工具的必要性

临床研究的质量是决定临床证据质量和实施的关键。临床研究的真实性和可信度主要反映在其内部真实性和外部真实性两方面,而这两种属性对于证据的质量影响和临床实施至关重要^[42,43]。可靠证据应具有内部真实性和外部真实性。内部真实性是指一项研究能够真正回答一个临床问题的能力^[44]。外部真实性是研究结果可以向外推广的程度^[45]。没有内部真实性,研究就不会具有外部真实性。内部真实性及外部真实性都可以被系统误差(即偏倚)、随机误差(即机遇)以及设计误差(即运用错误的设计去回答提出的问题,或者错误的情景)影响^[43,45]。

系统误差(即偏倚)是指结果或推断距离真实情况的偏离。系统误差可导致过高或者过低的估计真实干预措施效应^[32]。系统误差的风险和证据的强度取决于临床试验设计类型(系统综述,随机对照试验,观察性研究等)^[46]。临床研究类型所在等级越高,存在系统误差的风险越低,证据强度越大^[46]。Cochrane的偏倚风险工具‘Risk of bias tool’是一个被广泛认可的评估系统误差的评估工具^[32]。

随机误差(即不精确性)是指多次重复相同的研究会导致不同的评估效应^[32]。临床研究中,随机误差通常由零散数据或者重复检验引起。小样本研究结果可常常被抽样错误影响,因此结果通常不够精确^[32]。标准误(SE)则被应用于评估随机误差^[43]。

设计误差指设计要素的错误选择导致临床研究的外部真实性出现问题^[32]。设计误差可由不同的设计要素如PICOTA(参与者、干预措施、对照、结局指标、时间、分析)引起^[32,43]。在众多需要考虑的设计要素中,结局指标可作为临床设计的核心要素^[43]。对于设计误差评估的方法,主要围绕设计要点(结局指标、参与者选择、干预措施方案等)

进行,常用的评估方法如下:INCLEN 的公平滤镜^[47],PROGRESS 评估方法^[48],以及 Bornhöft 检查清单^[49]。Keus 及同事提出了证据矩阵(即误差矩阵,一个可视化评估工具)以评估临床证据的三个维度误差(即系统误差、随机误差以及设计误差)^[43]。

但是目前常用的质量评价工具大多只从偏重于单一误差维度来评价临床证据质量,因此研发出一款能够同时评价多维度误差的质量评价工具的需求十分迫切,且有助于更加全面、系统的对临床证据进行评估,进步促进临床研究质量的提高,有助于中医药临床实践应用,促进中医药学术水平提高。

参考文献:

- [1]Silvio Garattini, Janus C. Jakobsen, Jørn Wetterslev, et al. Evidence-based clinical practice: Overview of threats to the validity of evidence and how to minimise them. *European Journal of Internal Medicine*.2016,32:13-21.
- [2]Light RJD. *The science of reviewing research*. Cambridge, MA: Harvard University Press. 1984: 160-186.
- [3]Goldschmidt PG. *Information syntheses: a practical guide*. *Health Serv Res*. 1986, 21(2): 215-237.
- [4]L'Abbé KA, Detsky AS, O'Rourke K. *Meta-analysis in clinical research*. *Ann Intern Med*. 1987, 107(2): 224-233.
- [5]Mullen PD, Ramirez G. *Information synthesis and meta-analysis*. *Advan Health Educat Promot*, 1987, 2(7): 201-239.
- [6]Mulrow CD. *The medical review article: state of the science*. *Ann Intern Med*. 1987, 106(3): 485-488.
- [7]Oxman AD, Guyatt G. *Guidelines for reading literature reviews*. *Can Med Assoc J*, 1988, 138(8): 697-703.
- [8]Meinert CL. *Meta-analysis: Science or religion? Controlled Clin Trials*. 1989, 10(4): 257S-263S.
- [9]Smith MC, Stullenbarger E. *Meta-analysis: an overview*. *Nursing Sci Q*. 1989, 2(3): 114-115.
- [10]Oxman AD. *Checklists for review articles*. *BMJ*. 1994, 309(6955): 648-651.
- [11]Wilson A, Henry DA. *Meta-analysis*. *Med J Aust*. 1992, 156(5): 173-187.
- [12]Neely JG. *Literature review articles as a research form*. *Otolaryngol-Head & Neck Surg*. 1993, 108(6): 743-748.
- [13]Ohlsson A. *Systematic reviews-theory and practice*. *Scand J Clin Lab Invest*. 1994, 219: 25-32.
- [14]Oxman AD. *Checklists for review articles*. *BMJ*. 1994, 309(6955): 648-651.
- [15]Assendelft WJ, Koes BW, Knipschild PG, et al. *The relationship between methodological*

- quality and conclusions in reviews of spinal manipulation. *JAMA*. 1995, 274(24): 1942-1948.
- [16]Cook DJ, Sackett DL, Spitzer WO. Methodologic guidelines for systematic reviews of randomized control trials in health care from the Potsdam consultation on meta-analysis. *J Clin Epidemiol*. 1995, 48(1): 167-171.
- [17]Nony P, Cucherat M, Haugh MC, et al. Critical reading of the meta-analysis of clinical trials. *Thérapie*. 1995, 50(4): 339-351.
- [18]Geller NL, Proschan M. Meta-analysis of clinical trials: a consumer's guide. *J Biopharmaceut Stat*. 1996, 6(4): 377-394.
- [19]Thacker SB, Peterson HB, Stroup DF. Meta-analysis for the obstetrician-gynecologist. *Am J Obstet Gynecol*. 1996, 174(5): 1403-1407.
- [20]Greenhalgh T. Papers that summarise other papers (systematic reviews and meta-analyses). *BMJ*. 1997, 315(7109): 672-675.
- [21]Pogue J, Yusuf S. Overcoming the limitations of current meta-analysis of randomized controlled trials. *Lancet*, 1998, 351(9095): 47-52.
- [22]Auperin A, Pignon JP, Poynard T. Critical review of meta-analyses of randomized clinical trials in hepatogastroenterology. *Aliment Pharmacol Therap*. 1999, 11(2): 215-225.
- [23]Blettner M, Sauerbrei W, Schlehofer B, et al. Traditional reviews, meta-analyses and pooled analyses in epidemiology. *Int J Epidemiol*, 1999, 28(1): 1-9.
- [24]Linde K, ter Riet G, Hondras M, et al. Characteristics and quality of systematic reviews of acupuncture, herbal medicines, and homeopathy. *Forschende Komplementarmedizin und Klassische Naturheilkunde*. 2003, 10(2): 88-94.
- [25]Shea BJ, Hamel C, Wells GA, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol*, 2009, 62(10): 1013-1020.
- [26]Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol*. 2007, 7: 10.
- [27]Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction - GRADE evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology*. 2011, 64(4):383-94.
- [28]Jadad A R, Moore R A, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials*. 1996,17(1):1-12.
- [29]David Moher,Sally Hopewell,Kenneth F Schulz,et al.CONSORT 2010 说明与详述:报告平行对照随机临床试验指南的更新. *中西医结合学报*.2010,8(8):701-741.
- [30]Hugh MacPherson, Douglas G A, Richard H, et al. 针刺临床试验干预措施报告标准的修订:CONSORT 声明的扩展. *中西医结合学报*,2010,8(9):804-808.
- [31]Critical Appraisal Skills Programme (CASP).RCT checklist_14.10.10.[2017-05-12].
http://www.casp-uk.net/wpcontent/uploads/2011/11/casp_RCT_

Appraisal_Checklist_14oct.pdf.

[32] Higgins JPT. Cochrane handbook for systematic reviews of interventions version 5.1.0. The Cochrane Collaboration, 2011 [2017-05-12]. <http://www.cochrane-handbook.org>.

[33] <https://hbg.cochrane.org/>.

[34] Jonathan AC Sterne, Julian PT Higgins, Roy G Elbers, et al. Risk Of Bias In Non-randomized Studies of Interventions (ROBINS-I): detailed guidance. Cochrane Library.

[35] 刘鸣. 系统评价、Meta-分析设计与实施方法. 北京: 人民卫生出版社. 2010: 72-73.

[36] Rajasekhar A, Lottenberg R, Lottenberg L, et al. Pulmonary embolism prophylaxis with inferior vena cava filters in trauma patients: a systematic review using the meta-analysis of observational studies in epidemiology (MOOSE) guidelines. *J Thromb Thrombolysis*. 2011, 3,2(1): 40-46.

[37] Moga C, Guo B, Schopflocher D, et al. Development of a Quality Appraisal Tool for Case Series Studies Using a Modified Delphi Technique. Edmonton AB: Institute of Health Economics, 2012.

[38] Critical Appraisal Skills Programme (CASP). Qualitative Appraisal Checklist for Qualitative Research. Available at: <http://www.casp-uk.net/>.

[39] Government Chief Social Researcher's Office. Quality in Qualitative Evaluation: A framework for assessing research evidence. United Kingdom: Cabinet Office, 2003.

[40] Goldsmith MR, Bankhead CR, Austoker J. Synthesising quantitative and qualitative research in evidence-based patient information. *J Epidemiol Community Health*. 2007, 61(3): 262-270.

[41] Lewin S, Glenton C, Munthe-Kaas H, et al. Using Qualitative Evidence in Decision Making for Health and Social Interventions: An Approach to Assess Confidence in Findings from Qualitative Evidence Syntheses (GRADE-CERQual). *PLoS Med*. 2015, 12(10): e1001895.

[42] David MK, Georgios K. Against pragmatism: on efficacy, effectiveness and the real world. *Trials*. 2009, 10: 48.

[43] Frederik Keus, Jørn Wetterslev, Christian Gluud, et al. Evidence at a glance: error matrix approach for overviewing available evidence. *BMC Medical Research Methodology*. 2010, 10:90.

[44] Kent DM, Kitsios G. Against pragmatism: on efficacy, effectiveness and the real world. *Trials*. 2009, 10: 48.

[45] Rothwell PM. External validity of randomized controlled trials: 'to whom do the results of this trial apply?'. *Lancet*. 2005, 365: 82-93.

[46] Atkins D, Eccles M, Flottorp S, et al. GRADE Working Group: Systems for grading the quality of evidence and the strength of recommendations I: Critical appraisal of existing approaches. The GRADE Working Group. *BMC Health Serv Res*. 2004, 4:38.

- [47]Dans AM, Dans L, Oxman AD, et al. Assessing equity in clinical practice guidelines. *J Clin Epidemiol.*2007,60(6): 540-546.
- [48]Evans T, Brown H. Road traffic crashes: Operationalizing equity in the context of health sector reform. *Inj Control Saf Promot.* 2003,10 (1-2): 11-12.
- [49]Bornhöft G, Maxion-Bergemann S, Wolf U, et al. Checklist for the qualitative evaluation of clinical studies with particular focus on external validity and model validity. *BMC Med Res Methodol.* 2006,6: 56.

个人简介

1. 一般情况

姓名 孔德昭 性别 女 民族 汉
出生日期 1987年11月2日 籍贯：辽宁沈阳

2. 教育经历

2006.09-2013.06 辽宁中医药大学中医学专业，获硕士学位（本硕连读）
2016.09-2019.06 辽宁中医药大学中西医结合基础专业攻读博士学位

3. 工作经历

2013/09-2016/04, 辽宁中医药大学, 附属医院, 血脉病研究室, 医师
2016/05 至今, 辽宁中医药大学, 附属医院, 血脉病研究室, 主治医师

4. 发表论著

发表论著

- (1) **Kong DZ**[#], Wang Y, Liu Y, Zhang Z, Liu GH, Qi WC, Xiao L, Yuan DC, Yang GL*. *The association between blood lipid and phlegm turbidity syndrome of angina pectoris: A systematic review and meta-analysis.* Complementary Therapies in Medicine, 2014, 22: 801-813. (IF:2.216)
- (2) **Kong DZ**[#], Xia W, Zhang Z, Xiao L, Yuan DC, Liu Y, Yang GL*. *Safflower yellow injection combined with conventional therapy in treating unstable angina pectoris: a meta-analysis.* J Tradit Chin Med. 2013,33(5):53-61. (IF:0.589)
- (3) Zhang Z[#], Zhang F, Wang Y, Du Y, Zhang HY, **Kong DZ**, Liu Y, Yang GL*. *Traditional Chinese medicine for stable angina pectoris via TCM pattern differentiation and TCM mechanism: study protocol of a randomized controlled trial.* Trials. 2014, 15:422. (IF: 2.216)
- (4) **孔德昭**[#], 张哲, 王洋, 刘悦, 孟繁丽, 高晓宇, 潘家祥, 唐晶, 杨关林. “从脾论治”法对冠心病稳定型心绞痛患者生存质量的影响[J]. 辽宁中医杂志, 2018, 45(07):1345-1350.
- (5) **孔德昭**[#], 吕美君, 张哲, 王洋, 刘悦, 刘光辉, 陈雪娇, 杨关林*. 运用德尔菲法对《冠心病稳定型心绞痛脾虚痰浊证·中医证候计分表》条目筛选的研究[J]. 辽宁中医杂

志, 2014, (05):841-844.

(6)孔德昭#, 张哲, 王建华, 袁东超, 杨关林*. 心脑合病与心病、脑病的常见证型及分布特点[J]. 中华中医药学刊, 2013, (12):2614-2619.

(7)孔德昭#, 张哲, 袁东超, 夏伟, 吴瑾, 陈丽娟, 杨关林*. 冠心病稳定型心绞痛理化检查与量表间关系[J]. 中华中医药学刊, 2013, (06):1255-1260.

参编著作两部

(1)中医脾脏象理论研究, 人民卫生出版社, 304千字, 2018.

(2)中医心脑合病证治录, 辽宁科学技术出版社, 400千字, 2014.

5. 参研课题情况

1. 辽宁省教育厅一般项目, L2014372, 中药从脾论治冠心病随机对照试验的系统综述, 2014/01-2016/12, 3万, 已结题, 参加。

2. 国家重点基础研究发展计划(973计划), 2013CB531705, “脾虚生痰”所致冠心病心绞痛“从脾论治”疗效机制及规律研究, 2013/01-2017/12, 561万元, 已结题, 参加。

3. 辽宁省科学技术计划项目, 2009ZX09103, 和血通脉颗粒治疗冠心病临床前成药性评价, 2010/01-2012/12, 40万, 已结题, 参加。

4. 国家中医药管理局行业科研专项, 200807061, 心脑合病辨证规律的研究, 2009/04-2012/12, 90万, 已结题, 参加。

在学期间科研成绩

1. 发表论著

以下三篇为博士期间完成的 SCI 论文：前两篇为 Cochrane 系统综述，已进入最后编辑审校阶段，即将发表于 Cochrane Database of Systematic Reviews (IF: 6.754) 上；第三篇为方法学研究，已投稿至 BMC Complementary and Alternative Medicine (IF: 2.109) 上，已进入同行评议阶段。

(1) **De Zhao Kong**, Ning Liang, Dimitrinka Nikolova, Janus C Jakobsen, Christian Gluud, Jian Ping Liu*. *Xiao Chai Hu Tang, a Chinese herbal medicine formula, for chronic hepatitis B. Cochrane Database of Systematic Reviews*. 2018. Available from: <https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD013090/full>.

(2) **De Zhao Kong**, Ning Liang, Guan Lin Yang, Zhe Zhang, Yue Liu, Ye Yang, Yu Xi Liu, Qi Ge Wang, Fan Zhang, Hui Yong Zhang, Dimitrinka Nikolova, Janus C Jakobsen, Christian Gluud, Jian Ping Liu*. *Acupuncture for chronic hepatitis B. Cochrane Database of Systematic Reviews*. 2018. Available from: <https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD013107/full>.

(3) **De Zhao Kong**, Ning Liang, Guan Lin Yang, Zhe Zhang, Jun Wen Jiang, Yue Liu, Shi Bing Liang, Jing Li, Xue Han Liu, Yang Wang, Fan Zhang, Janus C Jakobsen, Christian Gluud, Jian Ping Liu*. *A qualitative and visual assessment approach of evidence - exemplified by Xiao Chai Hu Tang formula for chronic hepatitis B*. 2018. Submitted to BMC Complementary and Alternative Medicine.

2. 获奖情况

(1) **孔德昭** (7/8), 中医辨证规范的方法研究与应用示范, 中华中医药学会, 中华中医药学会科技进步奖, 三等奖, 2016

(2) **孔德昭** (3/11), 中医疗效评价技术临床探索与应用研究, 辽宁省中西医结合学会, 辽宁省中西医结合学会科学技术奖励, 一等奖, 2016

3. 承担或主研课题情况

(1) **国家自然科学基金项目(青年基金)**, 81803978, 基于“调和脾胃, 心无凝滞”探讨健脾祛痰化瘀方对Nox5-ERK1/2通路调控冠脉内皮氧化应激所致炎症反应的作用机制, 2019/01-2021/12, 21万元, 在研, **主持**。

(2) **国家教育部重大科研平台研究项目**, zyzx1702, 基于线粒体自噬探讨健脾祛痰法抗冠状动脉粥样硬化的作用机制, 2018. 01-2019. 12, 5万元, 在研, **主持**。

(3) **辽宁省教育厅科学技术研究项目**, L201713, 基于线粒体自噬探讨益脉颗粒治疗冠状动脉粥样硬化的疗效机制, 2017/10-2020/09, 2万元, 在研, **主持**。

4. 参加国际学术交流

(1) 于2018年9月14日-18日前往英国爱丁堡参加Cochrane爱丁堡学术年会“Cochrane Colloquium 2018”, 进行题目为“A survey of patient-centered outcome for coronary heart disease”的壁报展示。

致 谢

值此课题完成之际，在此向所有在学习期间给予生活和学习上帮助、指导、关心过我的老师、同学和朋友们致以诚挚的谢意！

本论文是在博士指导教师刘建平教授的严谨指导下研究和撰写的，从项目工作的具体实施到文章思路的形成以及撰写蕴含着指导教师的殷切期望和汗水，对恩师三载的无私培养致以诚挚的谢意！导师高尚的医德医风、严谨的治学态度、求实创新的科研作风、博采众长的学识风度使我终生受益，三载来中学生的每一步、每一个研究成果，都蕴含着指导教师的殷切期望。

由衷的感谢刘建平教授，在科学研究方面的教导使我受益非浅，在为人处事方面的教导令我终生难忘，在人生转折点给予的方向使我步履稳健，我将始终谨记恩师的谆谆教诲，在今后的人生道路中砥砺前行。

感谢丹麦哥本哈根试验中心（CTU）Christian Gluud 教授, Dimitrinka Nikolova 教授, Janus C Jakobsen 教授等各位教授给予的指导和帮助。

感谢梁宁、李静、梁士兵、刘雪寒、杨晔、刘羽茜、王琪格等同门科研人员给予的帮助！

感谢博士研修阶段所有指导过我的老师和研究生学院的全体老师和工作人员。

特别感谢我的家人对我的默默支持和无私奉献，是我完成学业的最大支持。

感谢答辩阶段师弟师妹给予的大力支持！

向所有协助进行课题研究及参加论文审阅和答辩委员会的老师、专家、评委们致以由衷的谢意！