

以手法为主要干预措施的随机对照试验 在样本量估计方面的常见问题分析

冯天笑¹, 肖想玉¹, 秦晓宽¹, 孙传睿¹, 王旭¹, 卜寒梅¹, 朱立国¹, 胡晶², 魏戌¹

(1. 中国中医科学院望京医院, 北京 100102;

2. 首都医科大学附属北京中医医院, 北京 100191)

摘要 目的:分析以手法为主要干预措施的随机对照试验(randomized controlled trial, RCT)在样本量估计方面存在的问题。**方法:**应用计算机检索 PubMed、Embase 和 Cochrane Library 中以手法为主要干预措施的 RCT, 检索时限均为 2017 年 1 月至 2024 年 6 月。依据文献检索及筛选方案筛选出符合要求的文献后, 由 2 名经过培训的研究人员独立提取是否报告样本量估计、假设检验类型、I 类错误概率 α 、把握度 $(1 - \beta)$ 、预期效应量、单侧或双侧检验、最小临床重要性差值(minimal clinically important difference, MCID)、脱落率、样本量计算工具、主要结局指标及多重检验等信息, 以频数和百分比的形式进行描述性分析。**结果:**共检索到 48 688 篇文献, 去除重复文献后随机抽取 20% 的文献, 通过筛选最终纳入 274 篇文献。在纳入的 274 项研究中, 192 项(70.07%) 报告了样本量估计。在报告了样本量估计的 192 项研究中, 16 项(8.33%) 报告了假设检验类型, 146 项(76.04%) 报告了 I 类错误概率 α , 164 项(85.42%) 报告了把握度 $(1 - \beta)$, 98 项(51.04%) 报告了预期效应量, 88 项(45.83%) 报告了单侧或双侧检验, 81 项(42.19%) 报告了 MCID, 156 项(81.25%) 报告了脱落率, 167 项(86.98%) 报告了样本量计算工具, 131 项(68.23%) 报告了主要结局指标。16 项报告了假设检验类型的研究均为优效性试验, 其中 9 项采用单侧检验, 7 项采取双侧检验; 37 项报告了 2 个及以上主要结局指标的研究中, 8 项(21.62%) 调整了 I 类错误概率 α ; 23 项报告了多个组别比较的研究中, 1 项(4.35%) 调整了 I 类错误概率 α 。**结论:**以手法为主要干预措施的 RCT 在样本量估计方面, 存在未报告样本量估计、样本量估计参数缺失或不明确、样本量估计方法与假设检验类型及主要结局指标不对应、多重检验时未调整 I 类错误概率 α 等问题。

关键词 按摩(中医); 正骨手法; 舒筋整复手法; 推拿疗法; 随机对照试验专题; 样本量

Common problems in sample size estimation of randomized controlled trials with manipulation as the primary intervention measure

FENG Tianxiao¹, XIAO Xiangyu¹, QIN Xiaokuan¹, SUN Chuanrui¹, WANG Xu¹, BU Hanmei¹, ZHU Ligu¹, HU Jing², WEI Xu¹

1. Wangjing Hospital of CACMS, Beijing 100102, China

2. Beijing Hospital of Traditional Chinese Medicine Affiliated to the Capital Medical University, Beijing 100191, China

ABSTRACT Objective: To analyze the common problems in sample size estimation of randomized controlled trial(RCT) with manipulation as the primary intervention measure. **Methods:** All the RCT articles with manipulation as the primary intervention measure included from January 2017 to June 2024 were retrieved from the PubMed, Embase and Cochrane Library through computer. The eligible articles were screened according to the retrieval and screening scheme, and the information, including whether reporting sample size estimation, hypothesis test type, type I error probability α , power of test $(1 - \beta)$, expected effect size, one-sided or two-sided test, minimal clinically important difference(MCID), dropout rate, sample size calculation tool, primary outcome indicators, and multiple tests, was extracted independently by two trained researchers for a descriptive analysis in the form of frequency or percentage. **Results:** Forty-eight thousand six hundred and eighty-eight RCT articles were searched out, after removing the duplicate articles, 20% of the remained ones were randomly selected for further screening, and 274 RCTs were included in the final analysis, among which 192 ones(70.07%) reported the sample size estimation. Furthermore, among the 192 RCTs, 16 ones(8.33%), 146 ones(76.04%), 164 ones(85.42%), 98 ones(51.04%), 88 ones(45.83%),

基金项目: 国家重点研发计划项目(2021YFC1712800, 2021YFC1712802); 北京市科技新星计划交叉合作课题(20220484228); 中国中医科学院望京医院高水平中医医院建设项目——中医药临床循证研究专项(WJYY-XZKT-2023-05)

通讯作者: 魏戌 E-mail: weixu.007@163.com

81 ones (42.19%), 156 ones (81.25%), 167 ones (86.98%), and 131 ones (68.23%) reported the hypothesis test type, type I error probability α , power of test $(1 - \beta)$, expected effect size, one-sided or two-sided test, MCID, dropout rate, sample size calculation tool, and primary outcome indicators, respectively. In addition, 16 RCTs that reported the type of hypothesis test were all superiority trials, with 9 ones using the one-sided test and 7 ones employing the two-sided test. Among the 37 RCTs that reported 2 or more primary outcome indicators, 8 ones (21.62%) mentioned the adjustment of type I error probability α . Among the 23 RCTs that reported multiple comparisons, one RCT (4.35%) referred to the adjustment of type I error probability α . **Conclusion:** Problems, such as failure to report sample size estimation, missing or unclear sample size estimation parameters, mismatch of sample size estimation method with hypothesis test types and primary outcome indicators, and failure to adjust type I error probability α in multiple test, exist in the sample size estimation of RCTs with manipulation as the primary intervention measure.

Keywords massage (TCM); bone setting manipulation; muscle tendon relaxing reduction; Tui Na therapy; randomized controlled trials as topic; sample size

手法包括中医理论指导下的推拿、正骨、按摩等手法和以现代医学理论、系统解剖学、生理学等作为基础的关节松动、肌肉放松等手法^[1-2]。手法治疗是重要的非手术疗法,国外手法治疗已形成了规范化、统一化的标准操作流程,国内手法治疗也形成了独具特色的学术流派,且在实践中不断发展和完善^[3-4]。然而,在进行手法的疗效评价时,仍面临循证医学证据不足的问题,这也限制了其在临床疾病诊治中的应用和发展。随机对照试验(randomized controlled trial, RCT)是临床疗效评价的金标准^[5],可以提供高等级的循证医学证据。然而,由于缺乏规范的试验设计和实施流程,部分 RCT 所提供的循证医学证据效用欠佳。样本量估计是 RCT 的关键环节,也是 RCT 获得可靠研究结果的前提条件。为此,我们分析了以手法为主要干预措施的 RCT 中样本量估计存在的问题,以期开展规范的临床试验提供参考。

1 资料和方法

1.1 文献检索

应用计算机检索 PubMed、Embase 和 Cochrane Library 中以手法为主要干预措施的 RCT,检索时限均为 2017 年 1 月至 2024 年 6 月。检索词包括 manip、osteopath、chiropract、manual therapy、mobilization、mobi-

lize、massage、tui na、randomized controlled trial、controlled clinical trial、random、trial。采用自由词与主题词相结合的方式检索,以 PubMed 为例,其具体检索策略见表 1。

1.2 文献筛选

去除重复文献后,采用 SAS9.4 软件随机抽取其中 20% 的文献。由 2 名经过培训的研究人员根据文献纳入和排除标准对检索到的文献进行筛选,遇到分歧,通过讨论或由第 3 名研究人员协助裁定。

1.2.1 文献纳入标准 ①国内外公开发表的 RCT; ②以手法(包括扳动类手法、关节松动术、整脊疗法、推拿、正骨、按摩等)为主要干预措施;③语种为英语。

1.2.2 文献排除标准 ①无法获取全文或数据无法提取的文献;②重复发表的文献。

1.3 资料提取

由 2 名经过培训的研究人员独立提取是否报告样本量估计、假设检验类型、I 类错误概率 α 、把握度 $(1 - \beta)$ 、预期效应量、单侧或双侧检验、最小临床重要性差值(minimal clinically important difference, MCID)、脱落率、样本量计算工具、主要结局指标、多重检验等信息,数据提取后进行交叉核对,遇到分歧,通过讨论或由第 3 名研究人员协助裁定。

表 1 PubMed 检索策略

序号	检索式
#1	manip * [Title/Abstract] OR osteopath * [Title/Abstract] OR chiropract * [Title/Abstract] OR manual therapy [Title/Abstract] OR mobilization [Title/Abstract] OR mobilize [Title/Abstract] OR massage [Title/Abstract] OR tui na [Title/Abstract]
#2	randomized controlled trial [Publication Type] OR controlled clinical trial [Publication Type]
#3	random * [Title/Abstract] OR trial * [Title/Abstract]
#4	#2 OR #3
#5	#1 AND #4

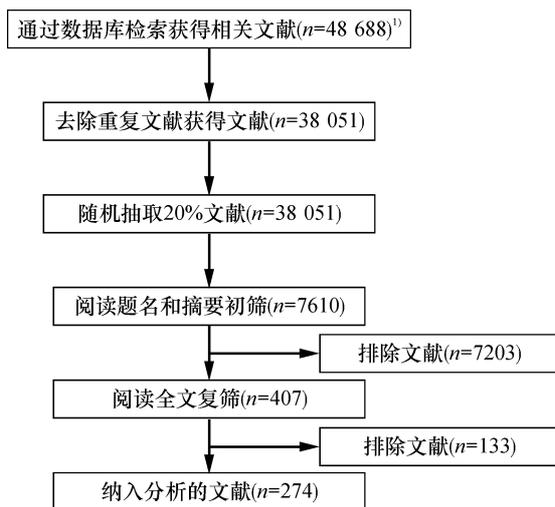
1.4 数据统计

采用 Excel 软件进行数据统计分析。以频数和百分比的形式对结果进行描述性分析。

2 结果

2.1 文献检索及筛选结果

共检索到 48 688 篇,最终纳入 274 篇文献。文献筛选流程及结果见图 1。



1)48 688 篇文献分别来源于 PubMed ($n = 11 513$)、Cochrane Library ($n = 28 774$) 和 Embase ($n = 8401$)。

图 1 文献筛选流程图

2.2 样本量估计报告率

274 项研究中,192 项研究报告了样本量估计。2017—2023 年,样本量估计报告率呈波动上升趋势(表 2)。

表 2 2017—2024 年以手法为主要干预措施的随机对照试验中样本量估计的报告率

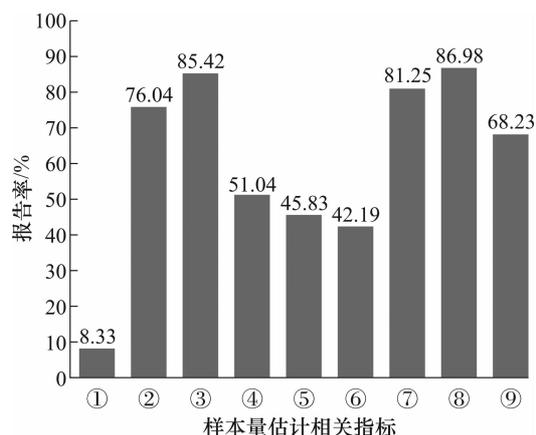
发表年份	文献数量/篇	报告样本量估计的文献数量/篇	样本量估计报告率/%
2017 年	30	18	60.00
2018 年	30	16	53.33
2019 年	31	24	77.42
2020 年	45	29	64.44
2021 年	29	22	75.86
2022 年	44	36	81.82
2023 年	47	37	78.72
2024 年	18	10	55.56

注:2024 年为上半年的样本量估计报告率。

2.3 样本量估计相关指标

16 项研究报告了假设检验类型,均为优效性试验;其中 9 项采用单侧检验, α 设定为单侧 0.025,7 项采取双侧检验, α 设定为双侧 0.05。146 项研究报告了 I 类错误概率 α ,其中 125 项设定 α 为 0.05,12 项

为 0.025,另外 9 项涉及多重检验,其中 4 项调整 α 为 0.0167,3 项为 0.0125,2 项为 0.01。164 项研究报告了把握度 $(1 - \beta)$,其中 128 项设定把握度 $(1 - \beta)$ 为 0.80,5 项为 0.85,21 项为 0.90,8 项为 0.95,2 项为 0.99。98 项研究报告了预期效应量,其中 85 项提及了预期效应量设定依据;在这 85 项研究中,53 项根据已发表的相关文献确定,32 项根据预试验或探索性试验确定。88 项研究报告了单侧或双侧检验,其中 65 项为双侧检验,23 项为单侧检验。81 项研究报告了 MCID,并将 MCID 用于样本量估计。156 项研究报告了脱落率,其中 62 项设定脱落率为 10%、21 项为 15%、70 项为 20%、3 项为 25%。167 项研究报告了样本量计算工具,其中 152 项采用 G * Power、PASS 等统计软件,6 项采用样本量计算公式,9 项采用统计手册或表格。131 项研究报告了主要结局指标,其中 94 项报告了 1 个主要结局指标,27 项报告了 2 个主要结局指标,8 项报告了 3 个主要结局指标,2 项报告了 4 个主要结局指标;37 项报告了 2 个及以上主要结局指标的研究中,8 项(21.62%)调整了 I 类错误概率 α ,其中 7 项采用 Bonferroni 法,1 项采用 Benjamini-Hochberg 法。23 项研究报告了多个组别的比较,其中 22 项报告了 3 个组别比较,1 项报告了 4 个组别比较;这 23 项研究中,1 项(4.35%)调整了 I 类错误概率 α 。样本量估计相关指标报告率见图 2。



①为假设检验类型,②为 I 类错误概率 α ,③为把握度 $(1 - \beta)$,④为预期效应量,⑤为单侧或双侧检验,⑥为最小临床重要性差值,⑦为脱落率,⑧为样本量计算工具,⑨为主要结局指标。

图 2 样本量估计相关指标报告率柱形图

3 讨论

临床试验可分为探索性试验和确证性试验。对于探索性试验,由于缺少必要的临床信息,无法通过

统计学方法确定样本量;但对于确证性试验,则需要正确的样本量估计方法进行准确的样本量估计。样本量过高会增加临床试验的难度,浪费医疗、人力、时间等资源;样本量过低则可能无法达到预期的检验效能,降低研究结果和结论的可信度^[6]。因此,样本量估计在国际上逐渐受到重视,一系列的临床试验方案撰写及结果报告标准中均对样本量估计提出了相应要求,如 *Methods and processes of the CONSORT Group: example of an extension for trials assessing non-pharmacologic treatments*^[7]、*CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials*^[8]、*SPIRIT 2013 statement: defining standard protocol items for clinical trials*^[9] 中均有关于样本量估计的条目。然而,目前 RCT 仍普遍存在着未报告样本量估计或样本量估计方法不合理的问题^[10-14]。我们在进行中医手法治疗颈型颈椎病 RCT 的质量评价研究中发现,81 项研究中仅 5 项研究报告了样本量估计^[10]。本研究结果显示,近些年来以手法为主要干预措施的 RCT 中,样本量估计报告率呈波动上升趋势,尤其 2021—2023 年,样本量估计报告率均超过 75%,提示样本量估计越来越受到研究者的重视。

然而,以手法为主要干预措施的 RCT 样本量估计仍存在较多问题:①部分研究未报告样本量估计。对于确证性试验,需要进行明确的样本量估计,以满足统计的准确性和可靠性(I类错误的控制和检验效能的保证),可靠地回答研究假设所提出的问题^[15-16]。本研究纳入 274 项研究,其中 82 项研究未报告样本量估计。②未明确假设检验类型,且样本量估计方法与假设检验类型不匹配。在临床试验中,样本量要根据研究目的、试验设计、主要结局指标、统计分析方法等进行估计^[17-18]。RCT 的假设检验类型有优效性试验、等效性试验和非劣效性试验。优效性试验是检验试验组干预措施的疗效是否优于对照组;优效性试验采用单侧检验, α 常取单侧 0.025。等效性试验是检验试验组干预措施的疗效是否与对照组“相等”(实际以差值不超过等效界值 Δ 进行评价);等效性试验需要进行两次单侧检验, α 常取单侧 0.05。非劣效性试验是检验试验组干预措施的疗效不劣于阳性对照组(实际以差值不低于事先给定的、临床可接受的非劣效界值 Δ 进行评价);非劣效性试验采用单

侧检验, α 常取单侧 0.025。本研究中明确假设检验类型的研究仅 16 项,且均为优效性试验;而这 16 项优效性试验中,7 项采取了双侧检验, α 设置为双侧 0.05。③未报告主要结局指标及未以主要结局指标进行样本量估计。临床试验的主要结局指标通常只有 1 个,样本量应基于主要结局指标来估计,以保证对主要结局的分析具有足够的把握度。本研究中,在提及样本量估计的 192 项研究中,61 项未明确主要结局指标,而在明确了主要结局指标的研究中,也存在采用次要结局指标进行样本量估计的问题。④样本量估计参数不完整。本研究结果显示,在报告样本量估计的 192 项研究中,146 项报告了 I 类错误概率 α , 164 项报告了把握度 $(1 - \beta)$, 98 项报告了预期效应量,81 项报告了 MCID, 156 项报告了脱落率。样本量估计参数不完整是临床试验样本量估计的共性问题,也是较为突出的问题。Abdulatif 等^[19] 系统分析了麻醉领域 10 本期刊发表的优效性 RCT 中样本量估计存在的问题,结果显示,在纳入的 193 项 RCT 中,19.7% 的 RCT 未报告样本量计算所需的必要参数。Dumbrigue 等^[20] 回顾性分析了 2008—2017 年口腔修复领域 2 本期刊发表的 RCT 中样本量估计存在的问题,结果显示,在纳入的 96 项 RCT 中,仅有 17% 的 RCT 在样本量估计中同时报告 I 类错误概率 α 、把握度 $(1 - \beta)$ 、MCID 和方差。⑤涉及多重检验的研究在进行样本量估计时未调整 I 类错误概率 α 。在样本量估计时,必须设定一个合理的 α 以控制 I 类错误发生,在多数情况下 $\alpha = 0.05$ ^[21]。I 类错误产生的原因有许多,多重检验是常见原因之一。对于报告多个主要结局指标或多个组别的 RCT,需要进行多重检验校正;而多重检验会增加 I 类错误发生的概率,导致 α 增大。因此,对于报告多个主要结局指标或多个组别的 RCT,在进行样本量估计时,需要调整 α 。目前,常用的调整 α 的方法为 Bonferroni 法,即通过公式 $\alpha' = \alpha/c$ (c 为组间两两比较的次数)调整 α 。采用 α' 计算各两组比较的样本量,并采用各两组比较中最大的样本量作为研究的最终样本量^[22-23]。本研究中,在 37 项报告了 2 个及以上主要结局指标的研究中,仅有 8 项调整了 I 类错误概率 α ,而在 23 项报告 3 个及以上组别的研究中,仅 1 项研究调整了 I 类错误概率 α 。

样本量估计是临床试验设计中的重要问题,但却

易被研究者忽略。上述问题的存在会影响研究结果和结论的可靠性,应引起手法研究者的关注和重视。在进行样本量估计时,完整的样本量估计参数是进行样本量估计的前提。我们认为,可以通过以下途径获得所需的样本量估计参数:①检索相似临床问题的相关研究(前期研究或相似研究);②开展小样本的预试验;③基于临床经验估计。面对以手法为主要干预措施的 RCT 样本量估计存在的未报告样本量估计或样本量估计方法不合理的问题,我们认为可以采取以下措施改善:①手法研究者应提高对样本量估计的重视程度,在临床试验设计阶段邀请方法学、统计学方面的专家论证研究方案,进而采用正确的样本量估计方法;②高校和科研机构应加强对手法研究者试验方案设计能力的培养,尤其在样本量估计的统计学原理、实施过程及报告方式等方面进行培训;③伦理委员会在审查临床试验方案时,将样本量估计作为着重审查的要素,进而督促研究者正确进行样本量估计。总之,严谨的 RCT 设计是高质量循证证据产生的前提条件,规范以手法为主要干预措施的 RCT 的样本量估计能为手法的推广应用提供高质量的证据支撑。

本研究结果表明,以手法为主要干预措施的 RCT 在样本量估计方面,存在未报告样本量估计、样本量估计参数缺失或不明、样本量估计方法与假设检验类型及主要结局指标不对应、多重检验时未调整 I 类错误概率 α 等问题。手法研究者应提高自身对样本量估计的重视程度以及正确进行样本量估计的能力,进而提高以手法为主要干预措施的 RCT 的质量。

参考文献

[1] 王晓宇,王虎城,刘蕾蕾,等. 手法治疗神经根型颈椎病疗效和安全性的系统评价再评价[J]. 中国循证医学杂志,2019,19(8):976-982.

[2] 魏戌,朱立国,李金学,等. 2014 年加拿大整脊疗法治疗成人颈痛循证指南解读[J]. 中国医学前沿杂志(电子版),2014,6(9):157-160.

[3] 王晓宇,李华南,张玮,等. 推拿国际化的关键因素——手法标准化建设[J]. 中华中医药杂志,2020,35(4):1658-1662.

[4] 严晓慧,严隽陶,龚利. 浅谈中医推拿手法标准化的重要性[J]. 河南中医,2009,29(3):242-243.

[5] 唐金陵, GLASZIOU P. 循证医学基础[M]. 2 版. 北京:北京大学医学出版社,2016:150.

[6] 刘萍,谢雁鸣. 中西医结合临床研究方法学[M]. 北京:

人民卫生出版社,2016:126.

[7] BOUTRON I, MOHER D, ALTMAN D G, et al. Methods and processes of the CONSORT Group; example of an extension for trials assessing nonpharmacologic treatments [J]. Ann Intern Med, 2008, 148(4): W60 - W66.

[8] MOHER D, HOPEWELL S, SCHULZ K F, et al. CONSORT 2010 explanation and elaboration; updated guidelines for reporting parallel group randomised trials [J]. J Clin Epidemiol, 2010, 63(8): e1 - e37.

[9] CHAN A W, TETZLAFF J M, ALTMAN D G, et al. SPIRIT 2013 statement; defining standard protocol items for clinical trials [J]. Ann Intern Med, 2013, 158(3): 200 - 207.

[10] 冯天笑,李康健,于大伟,等. 中医手法治疗颈型颈椎病随机对照试验的质量评价研究[J]. 中国全科医学, 2022, 25(29): 3664 - 3671.

[11] 冯天笑,李康健,于大伟,等. 中医手法治疗神经根型颈椎病随机对照试验结局指标的现状研究[J]. 中国全科医学, 2022, 25(20): 2525 - 2533.

[12] LEE P H, TSE A C Y. The quality of the reported sample size calculations in randomized controlled trials indexed in PubMed [J]. Eur J Intern Med, 2017, 40: 16 - 21.

[13] TAM W, LO K, WOO B. Reporting sample size calculations for randomized controlled trials published in nursing journals: a cross-sectional study [J]. Int J Nurs Stud, 2020, 102: 103450.

[14] IHEOZOR-EJIOFOR Z, LAKUNINA S, GORDON M, et al. Sample-size estimation is not reported in 24% of randomised controlled trials of inflammatory bowel disease: a systematic review [J]. United European Gastroenterol J, 2021, 9(1): 47 - 53.

[15] 胡晶,李博,张会娜,等. 针灸临床试验的样本量估计[J]. 中国针灸, 2021, 41(10): 1147 - 1152.

[16] 国家药品监督管理局. 总局关于发布药物临床试验的生物统计学指导原则的通告(2016 年第 93 号) [EB/OL]. (2016-06-01) [2023-06-13]. <https://www.nmpa.gov.cn/xxgk/ggtg/ypggtg/ypqgtgtg/20160603161201857.html>.

[17] 刘丹,周吉银. 临床科研项目样本量的要求[J]. 中国医学伦理学, 2019, 32(6): 716 - 718.

[18] 陈平雁. 临床试验中样本量确定的统计学考虑[J]. 中国卫生统计, 2015, 32(4): 727 - 731.

[19] ABDULATIF M, MUKHTAR A, OBAYAH G. Pitfalls in reporting sample size calculation in randomized controlled trials published in leading anaesthesia journals: a systematic review [J]. Br J Anaesth, 2015, 115(5): 699 - 707.

- [23] SUNG Y B, JUNG E Y, KIM K I, et al. Risk factors for neck shortening in patients with valgus impacted femoral neck fractures treated with three parallel screws: is bone density an affecting factor? [J]. *Hip Pelvis*, 2017, 29(4): 277-285.
- [24] 朴海旺, 祖思源, 陈申跳, 等. 股骨颈骨折空心钉内固定术后股骨颈短缩的危险因素分析[J]. *中国骨与关节损伤杂志*, 2023, 38(12): 1238-1241.
- [25] 中华医学会骨质疏松和骨矿盐疾病分会. 原发性骨质疏松症诊疗指南(2022)[J]. *中华骨质疏松和骨矿盐疾病杂志*, 2022, 15(6): 573-611.
- [26] HAJAR R. Risk factors for coronary artery disease: historical perspectives[J]. *Heart Views*, 2017, 18(3): 109-114.
- [27] YOU L, TEMIYASATHIT S, LEE P, et al. Osteocytes as mechanosensors in the inhibition of bone resorption due to mechanical loading[J]. *Bone*, 2008, 42(1): 172-179.
- [28] 秦威城, 何博文, 尚奇, 等. 从脂代谢指标探究骨质疏松症和骨质疏松性骨折的风险因素[J]. *中国骨质疏松杂志*, 2023, 29(12): 1791-1796.
- [29] 栗林, 徐毓菡, 周英杰, 等. 腰椎椎体间融合术后 cage 后移危险因素的 Meta 分析[J]. *中国脊柱脊髓杂志*, 2023, 33(10): 917-927.
- [30] VAN EMBDEN D, ROUKEMA G R, RHEMREV S J, et al. The Pauwels classification for intracapsular hip fractures: is it reliable? [J]. *Injury*, 2011, 42(11): 1238-1240.
- [31] VAN EMBDEN D, RHEMREV S J, GENELIN F, et al. The reliability of a simplified Garden classification for intracapsular hip fractures[J]. *Orthop Traumatol Surg Res*, 2012, 98(4): 405-408.
- [32] BEIMERS L, KREDER H J, BERRY G K, et al. Subcapital hip fractures: the Garden classification should be replaced, not collapsed[J]. *Can J Surg*, 2002, 45(6): 411-414.
- [33] POLAT A, MISIR A, BUYUKKUSCU M O, et al. Factors associated with femoral neck shortening after closed or open reduction and screw fixation [J]. *Indian J Orthop*, 2022, 56(2): 303-311.
- [34] 吾米提·塔西, 马信龙, 韩哲, 等. 断端皮质粉碎对青壮年股骨颈骨折股骨颈动力交叉钉系统固定术后疗效的影响研究[J]. *中华创伤骨科杂志*, 2021, 23(9): 775-781.
- [35] JAIN R, KOO M, KREDER H J, et al. Comparison of early and delayed fixation of subcapital hip fractures in patients sixty years of age or less[J]. *J Bone Joint Surg Am*, 2002, 84(9): 1605-1612.
- [36] ALKHATIB N, YOUNIS M H, HEGAZY A, et al. Early versus late treatment of paediatric femoral neck fractures: a systematic review and meta-analysis [J]. *Int Orthop*, 2019, 43(3): 677-685.
- [37] 中华医学会骨科学分会创伤骨科学组, 中国医师协会骨科医师分会创伤专家工作委员会. 成人股骨颈骨折诊治指南[J]. *中华创伤骨科杂志*, 2018, 20(11): 921-928.
- [38] QIU L, HUANG Y, LI G, et al. Essential role of reliable reduction quality in internal fixation of femoral neck fractures in the non-elderly patients—a propensity score matching analysis [J]. *BMC Musculoskelet Disord*, 2022, 23(1): 346.
- [39] WANG Y, MA J X, YIN T, et al. Correlation between reduction quality of femoral neck fracture and femoral head necrosis based on biomechanics [J]. *Orthop Surg*, 2019, 11(2): 318-324.
- [40] 魏文卿, 崔海坡, 陈舰, 等. 不同固定方式对股骨颈骨折固定效果的影响[J]. *医用生物力学*, 2020, 35(2): 216-222.

(收稿日期: 2024-06-13 本文编辑: 李晓乐)

(上接第 43 页)

- [20] DUMBRIGUE H B, DUMBRIGUE E C, DUMBRIGUE D C, et al. Reporting of sample size parameters in randomized controlled trials published in prosthodontic journals [J]. *J Prosthodont*, 2019, 28(2): 159-162.
- [21] 冯国双. 临床研究中样本量估算的影响因素[J]. *慢性病学杂志*, 2022, 23(5): 687-690.
- [22] HU J, FENG S, ZHANG X, et al. The use of multiple primary outcomes in randomized controlled trials of Chinese herbal medicine [J]. *Evid Based Complement Alternat Med*, 2021, 2021: 9975351.
- [23] ODUTAYO A, GRYAZNOV D, COPSEY B, et al. Design, analysis and reporting of multi-arm trials and strategies to address multiple testing [J]. *Int J Epidemiol*, 2020, 49(3): 968-978.

(收稿日期: 2024-06-24 本文编辑: 吕宁)