

随机对照试验中常见误区辨识与应对

李 静，孙 欢，李永弘

四川大学华西医院循证医学与临床流行病学研究中心，成都 610041

通信作者：李 静 电话：028-85422253，E-mail：lijing68@hotmail.com

【摘要】随机对照试验是评估不同干预措施疗效和常见不良反应的标准设计方案，但我国随机对照试验在研究设计、实施、效应指标选择、数据分析和结果解释等方面仍有提升空间。本文从随机对照试验实施的关键技术、结果展示和结果解读3方面分析了6点常见误区，期望帮助临床研究者明确随机化概念，正确实施随机分配方法和分配方案隐藏、正确理解随机对照试验中的基线平衡问题、合理选择效应指标、正确解读研究结果的统计学意义和临床意义，以生产更多高质量的随机对照试验证据，同时避免被误导。

【关键词】随机对照试验；分配方案隐藏；比值比；风险比；亚组分析

【中图分类号】R-1 **【文献标志码】**A **【文章编号】**1674-9081(2019)02-0166-06

DOI：10.3969/j.issn.1674-9081.2019.02.016

Common Misconceptions in Randomized Controlled Trials

LI Jing, SUN Huan, LI Yong-hong

Research Center of Evidence Based Medicine and Clinical Epidemiology, West China Hospital, Sichuan University,
Chengdu 610041, China

Corresponding author: LI Jing Tel: 86-28-85422253, E-mail: lijing68@hotmail.com

【Abstract】A well-conducted randomized controlled trial (RCT) is considered the gold standard for comparing the efficacy and common adverse events of different interventions. However, for China's RCTs, there is still room for the improvement in research design, implementation, selection of effective measures, data analysis, and interpretation of results. In this article, we present six misunderstandings or errors to help clinicians in clarifying the concept of randomization, correctly implementing the process of randomization and concealment of allocation, understanding the baseline imbalance in RCTs, reasonably selecting effective measurers, and correctly interpreting the statistical significance and clinical significance in an effort to produce more RCTs of high-quality and protect clinicians from potentially misleading presentations and interpretations of research findings.

【Key words】randomized controlled trial; allocation concealment; odds ratio; risk ratio; subgroup analysis

Med J PUMCH, 2019,10(2):166-171

临床研究是临床指南和医疗决策的重要依据^[1]。我国临床病例资源丰富，疾病种类繁多，具有开展临床研究得天独厚的条件。分析2006至2015年发表的临床研究论文发现，我国临床研究类论文呈正增长趋

势，但数量仍低于美国，临床研究类论文所占比重低于基础研究，研究质量亦亟待提高^[2-3]。随机对照试验（randomized controlled trial, RCT）是临床研究的重要组成部分，采用最严谨、最可靠的科学方法，很好

地解决了观察性临床研究中因组间不可比导致的混杂影响，成为评估上市前新医疗产品、比较不同干预措施疗效和常见不良反应的标准设计方法。本文将对 RCT 实施的关键技术如随机化原则、结果展示和结果解读中存在的部分常见问题进行分析，澄清一些基本概念和认识，希望有助于 RCT 质量的提高。

1 随机化原则的误区辨识与应对

尽管 RCT 已在临床医学领域应用 70 年，但其研究质量和研究报告质量仍有很大提升空间，其中随机化原则（randomization）是关键技术^[4-5]，如何正确理解、实施和在研究报告中准确阐述这一原则值得探讨。

1.1 随机对照试验中的“随机”是指什么？

案例 1：① 62 例女性患者，入院时随机抽样分为 A、B 两组；② 自 2000 年 1 月至 2010 年 4 月，在 210 例正常产妇中随机抽样，按随机分配表将产妇分为服药组 98 例，对照组 105 例；③ 143 例患者随机分为结核感染组和非结核感染组；④ 100 例小儿手术随机分成<5 岁和≥5 岁两组。

随机化原则是临床研究的重要方法和基本原则之一，包括随机抽样（random sampling）和随机分配（random allocation）两种形式。

随机抽样也称为概率抽样，在知道目标人群总体数量的前提下，采用随机抽样方法从总体中抽取一定数量的观察单位组成样本，使符合标准的观察单位均具有相同的机会被选择进入研究，提高样本对总体的代表性。当研究对象分散就诊如在门诊或住院部开展临床研究，研究者无法预测一定时间内因某病就诊的目标人群总数，也不可能在一定时间内当就诊人数超过估算的样本量后，再随机抽取一定数量患者开展研究。因此，案例 1 中的①和②描述的“随机抽样”是不可能实现的，说明研究者并未真正理解随机抽样的含义。

随机分配是指纳入研究的合格对象有相同的机会被分配入试验组或对照组，使已知和未知因素、能被测量和不能被测量的因素在组间达到基本相似，同时避免受到研究者或受试者主观意愿的干扰。RCT 只能用来检验根据现有医学知识认为对健康可能有益（至少利大于弊）的因素或措施，因此，案例③中患者是结核感染还是非结核感染是本身疾病的情况，不可能随机将患者分为结核感染组与非结核感染组。而案例④中小儿的年龄也是固有特征，不可能进行随机分配，类似情况还有如根据实验室检查结果、干预措施的可获得性等进行随机分配。

临床研究中往往是发现患者符合入选标准，在获得知情同意后即纳入研究，因此随机抽样多数情况下是不能实现的。RCT 中的“随机”是指随机分配研究对象，当然，若 RCT 能同时做到随机抽样和随机分配，结果既能保证对总体的代表性，又能有效避免选择性偏倚。

1.2 如何成功实施随机对照试验中的“随机分配”？

案例 2：① 26 个新生儿重症监护病房被随机分配提供家庭整合护理（n=14）或标准护理（n=12）；② 采用 SAS 9.4 生成随机数字，采用均衡随机化分组方式（区组长度在 4、6、8 间随机选取）进行随机分组，将分组的方案依次保存至顺序编号的密封不透明信封中，信封由研究设计者进行保管，当确定患者符合纳入排除标准并签署知情同意书，依据信封编号顺序分别开启密封信封，依据信封内分组方式，将患者分配至干预组或对照组。

RCT 中的“随机分配”是指采用真正随机分配的方法如简单随机法、区组随机法和分层随机法产生不可预测的分配序列，将符合入选标准的受试对象分配进入不同的研究组。随机分配序列产生后，受试对象入组的情况即已确定。如果产生分配序列与选择、分配合格受试对象入组的研究人员是同一人，或者产生的分配序列表保存在选择和分配受试对象入组的研究人员手中，甚至公开分配序列表，则研究人员就会预先知道下一个合格受试对象的入组情况。若该研究人员对要比较的干预措施如 A 和 B 有倾向性，认为 A 药应比 B 药好，为了让具有某种特征的受试对象接受 A 药以获得有益于 A 药的结果，就有可能改变随机分配序列，不按照事先产生的分配序列分配受试对象，导致选择性偏倚。如何保证患者按照预先设置的分配序列入组呢？为此引入了随机分配方案隐藏（allocation concealment）的概念。

随机分配方案隐藏是指产生和保存随机分配序列的人与参与试验并确定受试对象合格性的人不应是同一人，以确保患者、研究人员和其他参与试验的人员不会预先知道分配序列，避免选择性偏倚。多数研究发现^[6-8]，与分配方案隐藏完善的试验相比，未隐藏分配方案或分配方案隐藏不完善的试验，常常夸大治疗效果 30%~41%。完善的分配方案隐藏方法包括中央随机系统（通过电话或短信与试验办公室联系）、药房控制的随机（由药房准备药物或研究产品）、编号/编码的容器和按顺序编号、不透明、密封的信封。而不完善的分配方案隐藏方法包括根据住院号、生日等末尾数字的奇数或偶数分配研究对象，交替分配受

试对象，未密封、透明的信封^[9]。分配方案隐藏已逐渐被广泛应用，Cochrane 协作网也对此进行了精确的定义^[10]，但在应用中对其理解、描述仍不准确，且与盲法的概念常常混淆。

因此，RCT 中“随机分配”的成功实施应包括 2 个同等重要的步骤，即产生不可预测的随机分配序列和分配方案的隐藏，2 个步骤必须同时正确、无偏倚实施，方能达到真正随机分配的目的。案例 2 中的①既无具体随机方法的描述，也不清楚是否进行了分配方案隐藏，这种描述方式在国内杂志发表的 RCT 中比较常见，根据作者报告的信息无法判断随机分配方法的实施是否正确、成功。而案例②描述了随机分配序列产生方法和分配方案隐藏实施过程，唯一的遗憾是不清楚保存信封的研究设计者是否参与了试验对象的选择。正确描述随机分配过程应包括 4 个要素：(1) 如何产生随机分配序列（具体方法）？(2) 谁产生随机分配序列？(3) 谁保存随机分配序列？(4) 如何获取随机分配序列（如电话、短信、信封保存者在确定患者符合入选标准并记录其基本信息后，按照就诊顺序将相应的信封交给选择患者的研究人员拆封等）？

随机分配方案隐藏不应与盲法混淆，两者的目的、作用阶段和可行性是不一样的。随机分配方案隐藏是为了避免选择性偏倚，作用在受试对象分配入组前，在任何 RCT 中均能实施；而盲法是为了避免干预措施实施和结果测量时来自受试对象和研究人员的偏倚（实施和测量偏倚），作用于受试对象分配入组接受相应干预措施后，并非在任何 RCT 中均能实施^[8]。如比较外科手术和内科药物治疗某种疾病的疗效，随机分配方案隐藏是可行的，而对患者和手术医生采用盲法却难以实施。但在某些情况下，分配方案隐藏与盲法也可能为连续的过程，如在药物临床试验中，将试验药和对照药做成相同的形状、颜色、气味和味道，根据产生的随机分配序列，将药物放入外形、大小相同并按顺序编码的容器中。

1.3 随机分配一定能保证组间基线可比吗？

案例 3：2004 年 JAMA 发表的 CAMELOT 研究“抗高血压药物对血压正常的冠状动脉硬化性心脏病患者心血管事件的影响”一文^[11]，比较了氨氯地平或依那普利与安慰剂对心血管事件的影响。该研究为随机、双盲、多中心临床试验，共列出 30 个可能影响研究结果的基线因素，包括患者的基线特征、治疗和合并用药情况，其中 6 个因素在 3 个组间有统计学差异。

理论上，随机分配的目的是为了均衡组间已知和未知、能测量和不能测量的影响研究结果的重要因

素，保证组间结果的差异为所研究因素所致。但实际上，由于样本量和随机方法的不同，即使随机分配方法实施完全正确，也不可能让所有影响结果的重要因素在组间达到绝对均衡^[12]，认为 RCT 就可以均衡所有预后因素是自欺欺人。样本量大小与组间影响研究结果的重要因素的均衡性呈负相关，样本量越小，发生不均衡的可能性越大^[13]。传统的随机方法如简单随机法（simple randomization）和区组随机法（block randomization），虽然均能减少选择性偏倚，前者可能导致组间例数和重要预后因素的不均衡，后者可能均衡组间的样本例数，但也不能保证组间所有重要预后因素的均衡。目前可调整组间混杂因素均衡性的随机方法有分层区组随机（stratified block randomization）和最小化随机（minimization）。分层区组随机最常用，可降低随机化的不可预测性，但因每个区组的最后 1 例患者是没有选择的，当未完成的区组数较多或分层因素多而样本量有限时，也会造成严重的组间预后因素不均衡。最小化随机能起到均衡各组例数和重要预后因素的作用，在相同样本量情况下，可比分层区组随机均衡更多预后因素，被认为是临床试验的“铂金标准”。但也有专家认为最小化随机缺乏分配方案的隐藏，可能导致选择性偏倚^[14]。而且因其操作相对复杂，结果分析是否能采用常规统计方法及采用何种统计方法尚存在争议，在国内应用并不广泛^[14-15]。

当影响研究结果的重要预后因素不均衡时，如何处理呢？首先，可初步判断不均衡的预后因素是否会影响研究结论。如案例 3 中 6 个因素在 3 个组间不均衡，其中氨氯地平组和依那普利组患者年龄、低密度脂蛋白水平高于安慰剂组，2 个因素均可能增加氨氯地平组的心血管事件发生率；而利尿剂、血管紧张素转换酶抑制剂和钙通道阻滞剂的合并使用均为安慰剂组使用人数比例高于氨氯地平组，会缩小氨氯地平组与安慰剂组的疗效差异，在这些不均衡因素不利于氨氯地平组的情况下，研究仍然发现氨氯地平组与安慰剂组比较可减少心血管不良事件的发生，并有统计学差异，说明这些不均衡因素并不影响研究结论。其次，处理组间预后因素不均衡常采用统计方法如回归分析，校正不均衡因素对结果的影响。

2 比值比、风险比/相对危险度应用误区与应对

案例 4：假设采用 RCT 比较抗凝剂 A 和 B 治疗房颤患者发生胃肠出血的风险，结果见表 1。

表 1 抗凝剂 A 和 B 治疗房颤患者胃肠出血结果比较

治疗	出血 (n)	未出血 (n)	比值	风险
高风险人群				
抗凝剂 A	50	50	$50 \div 50 = 1$	$50 \div 100 = 0.5$
抗凝剂 B	20	80	$20 \div 80 = 0.25$	$20 \div 100 = 0.2$
关联强度			$OR = 1 \div 0.25 = 4$	$RR = 0.5 \div 0.2 = 2.5$
低风险人群				
抗凝剂 A	5	95	$5 \div 95 \approx 0.05$	$5 \div 100 = 0.05$
抗凝剂 B	2	98	$2 \div 98 \approx 0.02$	$2 \div 100 = 0.02$
关联强度			$OR = 0.05 \div 0.02 = 2.5$	$RR = 0.05 \div 0.02 = 2.5$

OR: 比值比; RR: 风险比/相对危险度

当随机对照试验的结局指标为二分类变量时, 表述治疗效应的指标可采用危险差 (risk difference)、风险比/相对危险度 (risk ratio/relative risk, *RR*)、机会比/比值比 (odds ratio, *OR*) 和针对时间事件数据的危害比 (hazard ratio, *HR*)^[16]。*OR* 和 *RR* 可用来量化干预措施效应, 两者有关联但并不相等, 也不能随意互换使用。目前 RCT 中 *OR* 和 *RR* 的使用存在 2 方面问题: (1) 误用: 一般情况下, *OR* 应用在回顾性研究中 (如病例对照研究和断面研究), 而 *RR* 用在有随访的前瞻性研究 (如 RCT 和队列研究)。若 RCT 中使用 *OR* 会产生什么样的后果呢? 一项研究^[17] 分析了 *N Engl J Med* 2004 至 2014 年发表的 RCT, 对 107 篇报告了未校正 *OR* (31 篇) 和 *RR* (76 篇) 的 RCT 分别计算了相应的 *RR* 和 *OR*, 结果发现 *OR* 平均夸大了 *RR* 62%, 其中 28 篇夸大 > 50%, 13 篇夸大 > 100%, 夸大的比例与事件发生率呈正相关。当事件发生率 > 10% 时, *OR* 倾向于夸大 *RR*, 而当事件发生率 < 10%, *OR* 和 *RR* 非常相近^[18]。从案例 4 中可见, 当两组患者的胃肠出血风险 > 10% 时 (高风险人群), *OR* 明显大于 *RR* 值, 当两组患者的胃肠出血风险 < 10% 时 (低风险人群), *OR* 与 *RR* 几乎相等。(2) 误解: *OR* 是两组的比值比 (案例 4 的 *OR* 是指抗凝剂 A 组发生与未发生胃肠出血人数之比除以抗凝剂 B 组发生与未发生胃肠出血人数之比, 即使用抗凝剂 A 发生胃肠出血的比值是抗凝剂 B 的 4 倍), 而 *RR* 是两组的事件发生率之比 (案例 4 的 *RR* 是指抗凝剂 A 组的胃肠出血发生率除以抗凝剂 B 组的胃肠出血发生率, 即使用抗凝剂 A 发生胃肠出血的风险是抗凝剂 B 的 2.5 倍), 在解释指标含义时 *RR* 比 *OR* 更直观易懂。但若事件发生率高于 10% 时, 将 *OR* 错误解释为 *RR*, 往往会夸大关联强度^[19]。针对案例 4 高风险人群的 RCT 应选择 *RR*, 若计算的是 *OR* 且将 *OR* 错误解释为 *RR*, 必然夸大胃肠出血的风险。

尽管当事件发生率较高时 *OR* 会夸大干预与事件间的关联强度, 但 *OR* 与 *RR* 的关联强度方向总是一致的 (均 > 1 或均 < 1)。

3 结果解读误区与应对

3.1 统计学意义与临床意义

案例 5: 采用 RCT 比较降压药 A 和 B 的降压效果, 结果是 A 药比 B 药多降低收缩压 2 mm Hg (1 mm Hg = 0.133 kPa), *P* = 0.02, 作者结论为 A 药降压效果优于 B 药。

显而易见, 临床专家不可能因 A 药比 B 药多降低收缩压 2 mm Hg 就认为 A 药降压效果优于 B 药。然而研究人员往往非常关注 *P* 值的大小, 若检验水平 α 设定为 0.05, 当主要结果指标的假设检验达到 *P* < 0.05 时, 认为该研究结果为“阳性”。如何正确解释此“阳性”结果呢? *P* = 0.02 仅代表降压药 A 和 B 的降压效果差异有统计学意义, 但是是否有临床意义需要结合专业知识, 从疗效、安全性和成本多方面综合考虑。*P* 值的大小和统计学意义既不能代表疗效的大小, 也不能代表临床重要性。事实上, 任何微小的疗效差异, 在样本量足够大的情况下均能获得 *P* < 0.05; 反之, 较大的疗效差异 (如 A 药比 B 药多降低收缩压 20 mm Hg), 在样本量较小情况下也可能出现 *P* > 0.05。因此, 不能将统计学意义与临床意义混淆^[20]。

3.2 亚组分析

案例 6: RITA-3 试验为一项多中心 RCT, 比较早期介入与保守药物治疗不稳定型心绞痛和非 ST 段抬高心肌梗死患者的疗效与安全性。主要结局指标之一为由死亡、心肌梗死组成的复合终点指标, 随访 1 年, 并根据性别进行了亚组分析^[21-22]。整合结果见表 2。

亚组分析是在结果分析阶段考虑到某些因素对结果的影响, 将观察对象按临床相似的特点 (如疾病严重程度、性别、年龄和种族等), 分成若干亚组比较, 常用于控制和纠正混杂因素的影响。RCT 中经常采用亚组分析, 可发现干预措施疗效和安全性在不同特征患者中的差异, 实现精准治疗。但亚组分析可能存在如下潜在问题^[23-24]: (1) 在 RCT 设计时若未考虑亚组因素而采用亚组分析, 破坏预先设定的随机分组, 影响组间基线可比性; (2) 当总体研究样本量足够但结果为阴性情况下进行多个亚组分析, 因多重比较, 亚组分析结果出现假阳性的概率随亚组分析次数增多而明显增加; (3) 当总体研究样本量足够但结果为阳性情况下进行亚组分析, 因样本估算时未考虑亚组因

表 2 RITA-3 试验随访 1 年及亚组分析结果 [n (%)]

结局	介入治疗 (n=895)	保守治疗 (n=915)	RR/OR (95% CI)	P 值
随访 1 年主要结果				
死亡	41 (4.6)	36 (3.9)	1.16 (0.75~1.80)	0.50
死亡和心肌梗死	68 (7.6)	76 (8.3)	0.91 (0.67~1.25)	0.58
亚组分析结果*				
死亡				
男性	23 (4.2)	28 (4.8)	0.78 (0.44~1.41)	0.031
女性	18 (5.1)	8 (2.4)	2.43 (1.01~5.84)	
死亡或心肌梗死				
男性	38 (7.0)	59 (10.1)	0.63 (0.41~0.98)	0.007
女性	30 (8.6)	17 (5.1)	1.79 (0.95~3.35)	
死亡、心肌梗死或顽固性心绞痛				
男性	67 (12.3)	125 (21.4)	0.48 (0.34~0.67)	0.002
女性	55 (15.7)	46 (13.9)	1.14 (0.74~1.76)	

* 亚组分析结果为校正 OR 和交互 P 值; RR、OR: 同表 1

素，则亚组分析时每组样本量不够，降低检验效能，增加假阴性结果的风险；(4) 事后分析 (post-hoc analysis) 可能导致选择性报告结果，当结果达到预期时就报告，反之则隐瞒不报。因此，应谨慎判断亚组分析结果的可信度，避免被误导。

为此，相关专家制定了 5 条适合临床医生判断亚组分析结果的标准^[25]：(1) 机遇能否解释观察到的亚组效应？可通过统计分析如交互作用的检验判断机遇的影响，但要考虑亚组分析时样本量不够，检验效能低而不能发现存在的亚组效应的可能性。(2) 亚组效应在不同研究间一致吗？若在不同研究间发现相似的亚组效应，可增加亚组分析结果的可信度。(3) 亚组效应是否为众多亚组假设中的少数？亚组效应的方向是否预先准确设定？若亚组假设多、为事后假设、假设的亚组效应方向错误，会影响亚组分析结果的可信度。(4) 是否有很强的生物学依据支持亚组效应？生物学依据可来自不同人群的研究包括动物研究、类似干预措施研究也观察到亚组效应和其它相关中间结局的研究结果，没有生物学依据会影响亚组效应的可信度，但有生物学依据也不能保证亚组效应一定真实存在。(5) 系统评价和 Meta 分析中的亚组分析还要考虑亚组效应是来自同一研究内还是不同研究间。来自不同研究间的亚组分析结果受影响的因素更多，研究对象特征、合并使用的干预措施、结局指标测量方法和研究的偏倚风险都可能存在差异，影响亚组分析结果的可信度。前 4 条是同时针对单个临床试验和系统评价，第 5 条仅针对系统评价和 Meta 分析。

按照上述判断单个临床试验亚组分析结果可信度的标准，案例 6 的亚组分析判断如下^[22,26]：(1) 存在交

互作用；(2) 亚组分析效应在不同研究间不完全一致，支持和不支持存在性别差异的研究均有；(3) RITA-3 的亚组分析为事后分析，但假设的亚组效应方向与预期一致，在 5 个亚组分析中 3 个有交互作用（表 2 仅列出有交互作用的亚组分析结果）；(4) 有一些生物学依据如男性和女性存在不同水平的生物标志物，男性有阳性生物标志物的比例高于女性；女性的冠状动脉较男性细，介入治疗容易出现更多并发症。当把顽固性心绞痛加入亚组分析因素时也有亚组效应，但该亚组分析是在主要研究结果无统计学意义、样本估算也未考虑到亚组因素情况下进行的。因此检验效能较低，多重比较也增加了假阳性结果的概率。故 RITA-3 的亚组分析结果有一定的提示作用（即不稳定型心绞痛和非 ST 段抬高的心肌梗死患者，介入治疗对男性可能更有益），但需谨慎解读，并进一步开展独立研究验证，明确介入治疗出现性别差异的机制。

4 小结

综上所述，本文从 RCT 实施的关键技术、结果展示和结果解读 3 方面分析了 6 点 RCT 研究中的常见误区，并阐述了相应判断方法，希望研究人员明确：(1) 随机对照试验中的随机是指随机分配；(2) 随机分配的成功实施包括产生不可预测的随机分配序列和分配方案隐藏 2 个重要过程；(3) 随机分配并不能保证组间基线的绝对可比；(4) 效应指标 OR 和 RR 的选择与事件发生率密切相关；(5) 有统计学意义的结果并不一定有临床意义；(6) 亚组分析结果要谨慎解读。期

待临床能够生产更多高质量的 RCT，为医疗决策提供可靠依据。

参 考 文 献

- [1] 陈耀龙, 杨克虎. 正确理解、制订和使用临床实践指南 [J]. 协和医学杂志, 2018, 9: 367-373.
- [2] Hu Y, Huang Y, Ding J, et al. Status of clinical research in China [J]. Lancet, 2011, 377: 124-125.
- [3] 王长通, 郭启勇. 中、美、日三国医学基础研究与临床研究类论文发表现状及影响因素分析 [J]. 中华医学科研管理杂志, 2018, 31: 219-223.
- [4] Ahmed Ali U, Reiber BMM, Ten Hove JR, et al. Journal impact factor and methodological quality of surgical randomized controlled trials: an empirical study [J]. Langenbecks Arch Surg, 2017, 402: 1015-1022.
- [5] Gonzalez GZ, Moseley AM, Maher CG, et al. Methodologic Quality and Statistical Reporting of Physical Therapy Randomized Controlled Trials Relevant to Musculoskeletal Conditions [J]. Arch Phys Med Rehabil, 2018, 99: 129-136.
- [6] 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 [J]. JAMA, 1995, 273: 408-412.
- [7] Moher D, Pham B, Jones A, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? [J]. Lancet, 1998, 352: 609-613.
- [8] Schulz KF, Chalmers I, Altman DG, et al. Allocation concealment: the evolution and adoption of a methodological term [J]. J R Soc Med, 2018, 111: 216-224.
- [9] Schulz KF, Chalmers I, Grimes DA, et al. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals [J]. JAMA, 1994, 272: 125-128.
- [10] Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials [J]. BMJ, 2011, 343: d5928.
- [11] Nissen SE, Tuzcu EM, Libby P, et al. Effect of antihypertensive agents on cardiovascular events in patients with coronary disease and normal blood pressure: the CAMELOT study: a randomized controlled trial [J]. JAMA, 2004, 292: 2217-2225.
- [12] Zhao W, Berger V. Imbalance control in clinical trial subject randomization—from philosophy to strategy [J]. J Clin Epidemiol, 2018, 101: 116-118.
- [13] Chu R, Walter SD, Guyatt G, et al. Assessment and implication of prognostic imbalance in randomized controlled trials with a binary outcome—a simulation study [J]. PLoS One, 2012, 7: e36677.
- [14] Taves DR. The use of minimization in clinical trials [J]. Contemp Clin Trials, 2010, 31: 180-184.
- [15] 吴春芳, 许金芳, 陆健, 等. 临床试验最小随机化的方法概况和研究前景 [J]. 中国新药杂志, 2010, 19: 823-826, 831.
- [16] Knol MJ, Duijnhoven RG, Grobbee DE, et al. Potential misinterpretation of treatment effects due to use of odds ratios and logistic regression in randomized controlled trials [J]. PLoS One, 2011, 6: e21248.
- [17] Balasubramanian H, Ananthan A, Rao S, et al. Odds ratio vs risk ratio in randomized controlled trials [J]. Postgrad Med, 2015, 127: 359-367.
- [18] Shah S, Youngquist S. Part 20: on odds and risk ratios [J]. Air Med J, 2013, 32: 8-9.
- [19] Tajeu GS, Sen B, Allison DB, et al. Misuse of odds ratios in obesity literature: an empirical analysis of published studies [J]. Obesity (Silver Spring), 2012, 20: 1726-1731.
- [20] Wasserstein RL, Lazar NA. The ASA's Statement on p-Values: Context, Process, and Purpose [J]. Am Stat, 2016, 70: 129-131.
- [21] Fox KA, Poole-Wilson PA, Henderson RA, et al. Interventional versus conservative treatment for patients with unstable angina or non-ST-elevation myocardial infarction: the British Heart Foundation RITA 3 randomised trial. Randomized Intervention of unstable Angina [J]. Lancet, 2002, 360: 743-751.
- [22] Clayton TC, Pocock SJ, Henderson RA, et al. Do men benefit more than women from an interventional strategy in patients with unstable angina or non-ST-elevation myocardial infarction? The impact of gender in the RITA 3 trial [J]. Eur Heart J, 2004, 25: 1641-1650.
- [23] Schulz KF, Grimes DA. Multiplicity in randomised trials II: subgroup and interim analyses [J]. Lancet, 2005, 365: 1657-1661.
- [24] Rothwell PM. Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation [J]. Lancet, 2005, 365: 176-186.
- [25] Sun X, Ioannidis JP, Agoritsas T, et al. How to use a subgroup analysis: users' guide to the medical literature [J]. JAMA, 2014, 311: 405-411.
- [26] Guo Y, Yin F, Fan C, et al. Gender difference in clinical outcomes of the patients with coronary artery disease after percutaneous coronary intervention: A systematic review and meta-analysis [J]. Medicine (Baltimore), 2018, 97: e11644.

(收稿日期: 2018-09-17)