

偏倚风险评估系列:(二)平行设计随机对照试验偏倚评估工具2.0介绍

杨智荣 孙凤 詹思延

CB1 8RN 英国剑桥大学临床医学院初级医疗中心(杨智荣); 100191 北京大学公共卫生学院流行病学与卫生统计学系(孙凤、詹思延), 北京大学循证医学中心(孙凤、詹思延)

通信作者:孙凤, Email: sunfeng@bjmu.edu.cn

DOI: 10.3760/cma.j.issn.0254-6450.2017.09.028

【摘要】 本文针对平行设计随机对照试验(RCT)Cochrane 偏倚评估工具2.0版本(RoB2.0)的主要内容进行介绍,同时对RoB2.0与前一个版本RoB1.0的区别与联系进行阐述,并举例说明RoB2.0的使用方法和注意事项。RoB2.0内容丰富,覆盖全面,为RCT的证据整合与评价提供了更多的偏倚风险信息。RoB2.0仍处于开发与进一步完善中,使用者可持续关注其后续更新与进展。

【关键词】 偏倚风险; 评估工具; 随机对照试验; 平行设计; 系统综述

基金项目: 国家自然科学基金(71673003)

Risk on bias assessment: (2) Revised Cochrane risk of bias tool for individually randomized, parallel group trials (RoB2.0)

Yang Zhirong, Sun Feng, Zhan Siyan
Primary Care Unit, Department of Public Health and Primary Care, School of Clinical Medicine, University of Cambridge, Cambridgeshire CB1 8RN, UK (Yang ZR); Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing 100191, China (Sun F, Zhan SY); Center of Evidence-based Medicine and Clinical Research, Peking University, Beijing 100191, China (Sun F, Zhan SY)

Corresponding author: Sun Feng, Email: sunfeng@bjmu.edu.cn

【Abstract】 This paper introduces the revised cochrane risk of bias tool for individually randomized, parallel group trials (RoB2.0), compares RoB2.0 and the previous version (RoB1.0). And illustrates the application of RoB2.0 for a published clinical trial. As a comprehensive tool, RoB2.0 provides more information on the risk of bias for evidence synthesis. RoB2.0 is still under development and it is suggested that the users should follow the updates of the developers in the future.

【Key words】 Risk of bias; Tool for assessment; Randomized controlled trial; Parallel design; Systematic review

Fund program: National Natural Science Foundation of China (71673003)

评估医学干预效果最理想的研究设计是随机对照试验(RCT),通过随机化的实施,RCT使各种可能的混杂因素在基线时达到组间均衡,但如果设计或实施不合理,其结果的真实性仍会受到偏倚的影响。因此,需要利用评估工具对RCT的偏倚风险进行评价。

一、RoB2.0的制定背景

RCT的偏倚风险评估工具(工具)最早可追溯至1961年^[1],随后制定的这类工具达数十种之多^[2]。其中2008年公布和2011年更新的Cochrane Collaboration's tool for assessing risk of bias in randomized trials (RoB1.0)的影响最为深远^[3-4]。该

工具克服了既往评估工具项目不完整或过于复杂、量表评分权重不合理等问题,成为Cochrane系统综述、非Cochrane系统综述中常用的主流评估工具。

然而,该工具并没有对RCT研究设计类型加以明确区分,也没有对干预的分配效果和依从效果进行明确界定。此外,该工具没有充分考虑组间沾染的问题。鉴于此,Cochrane方法学工作组对该工具进行了更新,涵盖平行设计、交叉设计和整群设计,明确了RoB1.0中一些容易造成混淆的概念,并调整了部分评估项目。新版的工具统称“偏倚评估工具2.0版本”(RoB2.0),是一个领域评估式(Domain-based evaluation)的工具。该工具于2016年10月20日

在项目网站上 (<https://sites.google.com/site/riskofbiastool/welcome/rob-2-0-tool>) 发布^[5],并在同年的韩国首尔 Cochrane 年会上正式公布。

平行设计的工具是整个 RoB2.0 工具的基础,交叉设计和整群设计的工具是在此基础上扩展而来的。本文主要介绍平行设计的 RoB2.0 工具,有关其他两种设计的工具解读和实例分析请查阅本刊随后的系列讲座。

二、工具解读

相比起 RoB1.0,可以把 RoB2.0 的更新归纳为以下 4 个方面:

1. 领域设置上的调整(表 1):

表 1 RoB1.0 和 RoB2.0 在领域设置上的对比

RoB1.0	RoB2.0
随机序列的产生(选择偏倚)	随机化过程中的偏倚
分配隐藏(选择偏倚)	分配隐藏
研究对象和干预实施者的盲法(实施偏倚)	偏离既定干预的偏倚
结局评估者的盲法(检测偏倚)	结局测量的偏倚
不完整结局数据(失访偏倚)	结局数据缺失的偏倚
选择性报告(报告偏倚)	结果选择性报告的偏倚
其他偏倚	(无)
(无)	整体偏倚

(1)把 RoB1.0 中随机序列的产生和分配隐藏这两个领域合并成一个领域,即随机化过程中的偏倚。

(2)把 RoB1.0 中研究对象和干预实施者的盲法这一领域扩展成一个新的领域,即偏离既定干预的偏倚。

(3)调整 RoB1.0 中的结局评估者的盲法、不完整结局数据、选择性报告 3 个领域,把名称分别改为结局测量的偏倚、结局数据缺失的偏倚、结果选择性报告的偏倚,并对评估标准进行了调整。

(4)完全删除 RoB1.0 中的“其他偏倚”这一领域。在 RoB2.0 中不再考虑增删偏倚来源,但可以适当调整每个领域里的信号问题(Signalling question)。

(5)明确提出整体偏倚这一综合评价的领域。即根据其他各领域的评价结果对单个 RCT 的特定结局作出整体的偏倚风险评价。

(6)预测偏倚的方向。RoB2.0 在每个领域下面都设置了这一评价项目,这是过去所有主流评估工具不曾涉及的方面。如果能推断偏倚的方向,即高估还是低估真实的结果,这显然要优于单纯的偏倚风险评估。但很多时候并没有证据来推测偏倚的方向,因此在 RoB2.0 中此部分是备选的项目。

2. 信号问题的调整:RoB2.0 明确地为每个领域设置了数个信号问题,要求评估者先对信号问题作

出“**Yes(是)**”、“**Probably yes(可能是)**”、“**Probably no(可能不是)**”、“**No(不是)**”、“**Not applicable(不适用)**”或“**No information(不清楚)**”的回答。通过对信号问题的回答,作者可了解偏倚来源的信息,即相关“信号”,然后依据这些“信号”进一步对各领域的偏倚风险进行评估。在对比 RoB1.0 的基础上,重点对 RoB2.0 各领域的评估标准进行逐一解读,并在附录中提供了完整的评估工具和相应的条目解释(附录见网站 <http://cpse.bjmu.edu.cn/>)。

(1)随机化过程中的偏倚:该领域合并了 RoB1.0 中的随机序列产生和分配隐藏。RoB1.0 中这两个领域分别有一个信号问题,即“分配干预的方案是否随机产生”,“是否采用了分配隐藏”。这两个信号问题都保留在 RoB2.0 里,此外还添加了一个新的信号问题,即“是否存在由随机化过程中的问题造成的组间基线不均衡”,该信号问题在 RoB1.0 中常见于评估者自行设置在“其他偏倚”领域里。有观点认为,在随机化的前提下,组间基线不均衡是因为随机因素引起的,不应作为偏倚风险的考虑范畴。然而,如果出现重要预后因素的组间不均衡,很可能导致结果难以解释。而且,当多个因素基线不可比时,有理由对随机化的过程提出质疑,尤其是当作者没有明确报告是否采用了真正随机的方法和/或分配隐藏的方法的时候,组间基线可比性是衡量随机化是否真正实现的一个替代指标。

(2)偏离既定干预的偏倚:该领域对应 RoB1.0 中研究对象和干预实施者施盲的部分,是 RoB2.0 改动最大的一部分。在理想条件下,所有 RCT 中的研究对象应当严格按既定的方案接受研究者分配的干预,但在实际研究实施中,经常有各种主观和客观因素导致患者实际接受的干预偏离了既定的干预方案。在 RoB1.0 中,该领域仅有一个信号问题,即“是否对研究对象和干预实施者实施了盲法”。这个问题在 RoB2.0 中被拆分成两个信号问题,分别评估“研究过程中研究对象是否知道他们接受哪种干预”和“研究过程中医护人员和研究人員是否知道研究对象接受哪种干预”。在此基础上,RoB2.0 还增加了四个信号问题。这四个问题的设置按照系统综述作者(注意,这里指的是系统综述的作者,而非原始研究的作者)的研究目的分为两类,在评估时需要根据不同的目的选择不同问题进行回答。两类研究目的分别是评估干预分配效果和评估干预依从效果。前者主要关注基线分配干预后的效果,而相对地忽略研究中研究对象沾染、更换干预、依从性差等问

题,其理念类似于RCT数据分析中的“意向性分析(intention-to-treat analysis, ITT分析)”;后者主要关注研究对象真正按照分配方案接受干预后的效果,强调研究对象实际遵从分配干预的情况下显示出来的效果,其理念类似于RCT数据分析中“遵循研究方案分析(per protocol analysis, PP分析)”。

如综述作者评估的是干预分配效果,且在前两个信号问题评估中,发现原始研究并未对研究对象和干预实施者施盲,则需进一步回答第三个问题,“在患者接受的干预中,是否存在非常规的干预改变”。通过这个问题, RoB2.0对偏离既定干预的情况进行了区分。一种情况是,既定干预的偏离是一种“常规”的干预变化,即这种偏离情况与临床实践相符。例如,患者因为干预产生的不良反应而终止干预,虽然既定干预产生了变化,但因为常规临床实践中,患者出现此类不良反应时也需要终止服药,故此类变化属于“常规”的干预变化。这种变化带来偏倚的可能性较小。另一种情况是,干预变化并不是“常规”实践。例如,分配到安慰剂组的患者去寻求其他治疗,产生了与既定研究方案的变化,这种变化可能对研究结果带来偏倚。如第三个问题评估中,存在“常规”临床实践变化外的干预偏离,则需进一步回答第四个问题,“干预变化是否在组间不平衡,并可能对研究结局造成影响”。并非所有与既定干预的偏离都会带来偏倚,只有在同时满足干预变化组间不平衡和干预变化影响研究结局这两个条件才可能带来偏倚。例如,某项RCT中,对照组部分患者改用了一种新药,而在干预组中,有相似比例和类似特征的患者也改用了该新药。尽管此例中存在与既定干预的偏离,且这种偏离可能影响结局,但因为这种变异在组间大致均衡,故带来偏倚的风险相对较低。本领域第五个信号问题“是否有研究对象没有被按照其分配的干预措施进行分析”是针对ITT分析的。因为综述作者的目的是评估干预分配的效果,在数据分析中,患者的分组应当依据随机化分配的方案确定,而不能通过患者实际接受的干预确定,否则会破坏组间随机化。如第五个信号问题评估中原始研究没有依据ITT分析原则进行数据分析,需进一步回答第六个信号问题,“错误的分组是否会对分析结果产生影响”。在数据分析中,若有患者偏离了既定干预,即便没有采用ITT分析,也不一定给研究结果带来偏倚。偏倚风险大小还取决于出现未遵从、沾染等情况的比例。RoB2.0认为,若超过5%的研究对象在组别的划分中出现了错误,则可能对结

果产生影响。但对于罕见结局,更小的比例即可对结果产生影响。

如综述作者评估的是干预依从效果,且在前两个信号问题评估中,发现原始研究并未对研究对象和干预实施者施盲,则需进一步回答第三个问题,“重要的伴随干预措施在不同组间是否均衡”。RCT中,研究对象除了接受分配的干预措施外,可能还同时接受其他的伴随干预。并非所有伴随干预都会对研究结果带来偏倚,只有当伴随干预在组间不平衡和伴随干预会对结局产生影响这两个条件同时满足的情况下,伴随干预措施才可能造成偏倚。例如,在一项评估降血压药物对心血管事件保护效果的临床试验中,一部分研究对象除接受分配的降压药(研究干预)外,还接受了降脂药治疗(伴随干预)。尽管降脂药对心血管事件(结局)会产生影响,但研究中,干预组和对照组中服用了降脂药的患者在数量比例和临床特征上相似(组间均衡),则偏倚风险相对较低。第四个信号问题,“计划实施的干预措施是否成功执行”,关注的是干预实施者是否按照既定方案实施了干预。第五个信号问题,“研究对象是否依从了分配的干预措施”,关注的是研究对象是否接受并依从了干预。如果干预措施是一次性实施的,例如只需接种一次的疫苗,只要成功实施了干预,就代表研究对象依从了干预,因为这种情况下不存在不依从的可能;但如果不是一次性实施的干预,则可能因为组间交叉、依从性不佳、中断干预而造成偏倚。在第三至第五个信号问题的评估中,如存在伴随干预、未严格实施干预或未严格依从等情况,则需回答第六个信号问题,“研究是否采用了恰当的方法估计干预的依从效果”。与分配效果不同,在评估干预依从效果时不能忽视影响依从性的因素,需要用合理的分析手段对实际接受干预的偏离情况进行调整,如逆概率加权和边缘结构模型等。若未进行调整,不管是使用ITT分析、PP分析,还是实际接受干预分析(As-treated analysis)均是不恰当的,都可能给依从效应的估计带来偏倚。

(3)结局数据缺失的偏倚:该领域对应RoB1.0中不完整结局数据带来的偏倚(失访偏倚)。在RoB1.0的基础上,将信号问题扩展至3个。第一个信号问题是“是否可以获得全部或绝大多数研究对象的研究结局数据”,该问题基本与RoB1.0中保持一致,如没有失访或失访率很低(一般是5%以下),则发生偏倚的风险较小。如存在一定比例的失访,则需进一步回答第二个信号问题,“结局数据缺失的

比例和数据缺失的原因在组间是否均衡”。“均衡”指组间缺失数据的比例接近,或仅有少量由随机误差造成的区别,而且各组缺失数据的原因在组间没有系统性差异。如存在组间缺失数据比例或原因不均衡,则需进一步回答第三个信号问题,“是否有证据支持,即便存在缺失数据,分析结果仍然是稳健的”。对于结果是否稳健,需要根据原始研究处理缺失数据的方法、敏感性分析的结果等进行综合判断。

(4)结局测量的偏倚:该领域对应RoB1.0中结局评估者的盲法(检测偏倚)部分。该领域评估中,应注意对“结局评估者”的定义。若RCT采用的是患者自报结局(Participant reported outcome, PRO),那么结局评估者就是指研究对象自身。RoB2.0在RoB1.0的基础上,将信号问题扩展至2个。第一个信号问题,“结局评估者是否知道研究对象接受的干预”,该问题基本与RoB1.0中保持一致,如结局评估者不知道研究对象接受的干预(或对研究结局评估者施盲),则发生偏倚的风险相对较小。如结局评估者知晓研究对象接受的干预,则需进一步回答第二个信号问题,“知道研究对象接受的干预措施是否会对结局测量产生影响”。并非所有的研究结局都会因评估者知晓分配的干预而受到影响。典型的例子是全死因死亡,该结局评估中不涉及任何主观判断,故即便评估者知晓分配的干预,带来偏倚的风险也相对较低。此信号问题的回答中,须特别注意一些看似不受到主观影响的结局,但实际是结合了主观判断和客观测量(X线光片诊断某疾病、心肌梗死、体格检查结果等),或结合了主观决定和客观测量的结局(住院时间、入院、出院、终止治疗、转科、剖宫产等),评估这类结局时知晓分配方案仍可能带来偏倚风险。例如,某RCT结局为心血管病死亡。尽管对死亡的评估是客观的,但心血管病的判断,或某个具体病例是否可以归属为心血管病死亡,则需要医生进行主观判断、诊断,此时知晓分配方案仍可能会带来偏倚风险。又如,某RCT结局为患者住院时间,尽管对住院时间本身的判断是客观的,但因为某些情况下,结局评估者(如评估者是患者的主管医生)可以决定住院时间的长短,此时知晓分配方案仍可能带来偏倚风险。

(5)结果选择性报告的偏倚:该领域对应RoB1.0中选择性报告(报告偏倚)。RoB2.0中该领域的核心内容与RoB1.0基本保持一致,关注的是研究者是否有选择地报告了研究结局数据。RoB2.0在RoB1.0的基础上,将信号问题细化为两个方面。

第一,研究者是否对多重结局测量(Multiple outcome measurements)进行了选择性报告。这里的多重结局测量是指对同一个结局在不同时间或用不同工具或用不同定义进行的多次测量。例如,某RCT在计划书中,明确指出研究结局是某指标在患者术后7 d、14 d、30 d时的测量值。但在最终的研究发表文章或研究报告中,仅报告了术后7 d的结局测量值,并未给出选择7 d这个时间点进行报告的恰当理由,此时就存在选择性报告。第二,研究者是否对多重分析的结果进行了选择性报告。例如,某RCT在计划书中,明确在统计学分析部分指出,将对某连续型变量进行成组 t 检验分析,并按照某标准将此指标划分为阳性和阴性,进一步进行 χ^2 检验。但在最终的研究发表文章或研究报告中,仅报告了成组 t 检验分析结果,并未给出选择此结果进行报告的恰当理由,这种选择性报告也可能引入偏倚。选择性报告常见的原因是,报告的那些结果,在组间差异具有统计学意义,或较为新颖,或可以使文章更容易发表。对多重结局测量或多重分析的选择性报告可掩盖部分结果,若未报告的结果与所报告的结果存在系统性差异,则很可能给决策带来误导。

对结果选择性报告的判断,应依据研究计划书、统计分析计划书、试验注册平台等信息来源进行判断。研究计划的发表必须先于试验的开始,在试验已经开始实施后进行的补救注册或计划书发表等不能排除选择性报告存在的可能。

当不能获得事先发表的研究计划时,系统综述者仍可通过以下两种途径来判断选择性报告的可能性:一是比较同一篇文章的方法和结果这两部分的一致性,二是比较同一研究在不同文章里所描述内容的一致性。若结局的定义、测量或分析存在不一致,而且没有合理阐述不一致的原因,则不能排除选择性报告的存在。系统综述者可进一步与原始研究的作者联系以了解不一致的原因。

(6)整体偏倚评估:该领域是RoB2.0中新增的内容。虽然在Cochrane手册中有涉及整体偏倚风险评估的内容^[3],但在RoB1.0中,并未明确设置一个整体风险评估的领域。这一领域在RoB2.0中有明确的体现,以反映所评估的RCT的偏倚风险全貌。如果上述所有5个领域评估结果均为低风险,则整体偏倚风险低(Low risk of bias);如果其中任意一个领域的评估结果为高风险(High risk of bias),或多个领域内的评估结果是可能存在风险,则整体风险高;除上述两种情况的RCT则可能存在偏倚风险

(Some concerns)。RoB2.0整体偏倚这一领域是对单个研究单个结局水平的评价,即对同一个RCT,不同的结局可能会有不同的整体评估结果。

3. 评估标准的调整:在完成信号问题的回答后,评估者需要根据回答的情况对相应领域的偏倚风险作出“Low risk”、“High risk”或“Some concerns”的评估(RoB1.0中的Unclear risk改为Some concerns)。最后,根据所有单个领域的评估结果对“整体偏倚”这一综合领域作出评价^[4]。每个领域的评价标准均与RoB1.0的有所区别。RoB1.0中,上述3种偏倚风险的判断可以直接根据原始研究中报告的信息进行,而RoB2.0中,开发团队为每个领域内的偏倚风险评估设置了一套类似于“操作路径”的评估标准,在确定了领域内每个信号问题的回答后,即可根据此标准评估该领域的偏倚风险。每个领域内的“路径”不尽相同,且开发团队已明确表示后续会提供自动化程序供评估者使用,并嵌入到RevMan软件中,感兴趣的研究者可参阅RoB2.0官方指南文件了解^[5]。

4. 基本评估信息的添加:RoB2.0中含有一个表格(见附录),要求评估者在使用该工具前填写一些与工具使用的相关信息。例如,在表格中评估者需要回答所纳入试验的设计类型,不同的设计类型需要选用不同的RoB2.0工具。评估者需要确定所评估的结局,如上所述,不同的结局所遭受的偏倚风险程度可能不同。评估者还需要明确指出评估的信息来源,多种信息来源可能有助于提高评估的质量。

需要特别指出的是,评估者要清楚自己(而非RCT研究者)在系统综述中所要评价的是干预的分配效果还是依从效果。正如在上文所述,在系统综述作者区分两种不同目的后,将使用不同的信号问题进行偏倚评估。而系统综述作者对综述目的的判断则取决于要回答何种研究问题,或关注何种干预效果等。总之,此处对于综述目的判断与区分,应当基于系统综述作者自己的背景知识、实践经验、专业判断、研究目的等信息,而不应基于纳入的RCT中的信息进行判断。

三、实例分析

以下将使用1篇已经发表的研究举例说明RoB2.0的实际使用。该文是2016年6月13日发表在新英格兰医学杂志上的一项平行设计随机对照试验,评估高心血管病风险的2型糖尿病患者中,与安慰剂比较,利拉鲁肽是否可以降低首次心血管病死亡、非致死心肌梗死和非致死性卒中的综合结局的风险^[6]。本研究除最终正式发表的试验结果文章外,还在

Clinicaltrials.gov网站进行了注册(NCT01179048),并提前发表过研究计划书^[7],这些资料都可以用于RoB2.0的偏倚评估。为了覆盖该工具的全部内容,本实例分析中,既按照评估分配效果进行了评价,同时也按照评估依从效果进行了评价。在工具实际使用中,综述作者只需要根据自己的研究目的对上述其中一种效果进行偏倚评估即可。由于RoB2.0主要在偏离既定干预的偏倚领域进行了调整,对该领域进行偏倚评估和解读(表2)。由于篇幅所限,读者可参考附录获取完整的评估前准备和研究基本信息,以及全文偏倚评估具体内容。

四、讨论

整体上来说,RoB2.0内容丰富、翔实。在RoB1.0的基础上,进一步完善了各个偏倚评估重要组成部分的细节,强调了偏倚评估在证据整合与评价过程中的重要地位。相比于RoB1.0,RoB2.0有明显的优势。RoB2.0细化了具体领域的评估过程,将RoB1.0评估中易于产生困扰和分歧的项目,均使用信号问题进行明确指示,有助于提高评估结果的可重复性和一致性。此外,RoB2.0不仅提供了偏倚风险评估,还提供了可能的偏倚方向评估,后者将有助于综述读者判断干预的真实效果方向。

与相对成熟的RoB1.0工具不同,RoB2.0尚处于完善阶段,截至本文投稿时,RoB2.0工具尚未在正式刊物上发表,也未正式纳入Cochrane手册,可能仍存在问题有待开发团队后续改进。①RoB2.0的可操作性相比RoB1.0明显下降。根据测试,完成RoB2.0的时间相比RoB1.0显著增加,判断各领域的偏倚风险时流程复杂,且目前尚无规范的方案展示评估结果(可暂时参考QUADAS-2评估结果的展示方式)。相信这个问题在开发团队公布程序化工具后会有所改善。②RoB2.0许多条目涉及到大量方法学知识和专业词汇,非方法学专业的综述作者使用本工具时可能存在困难,需要开发团队后续制定手册进一步明确一些方法学词汇定义及工具的使用方法,并举例说明其含义。③工具中一些细节需要进一步统一,如,某些条目中,回答“是”或“可能是”对应的是“低风险”,而另一些条目中同样的回答却是“高风险”,这很容易给评估者造成困扰并增加出错的机会。这些不一致完全可以通过转变信号问题的提问方式进行统一。④现在还没有能够方便进行RoB2.0评估的软件工具,但RoB2.0的开发团队已表示,正将RoB2.0整合至下个版本的Review Manager软件中。⑤RoB2.0中一些条目的设置从方法学角

表 2 RoB2.0 偏倚评估实例

领域	信号问题	回答	支持信息/理由
Assess the effect 系统综述的目的是评估干预分配的效果。(实际使用中选择干预分配或干预依从其中一种目的即可)			
偏离既定干预的偏倚	2.1 研究过程中研究对象是否知道他们接受哪种干预?	可能不是	“The code for a particular subject may be broken in a medical emergency if knowing the identity of the treatment allocation would influence...” 研究计划书 59 页 10.2 部分。 尽管各文件中没有明确的直接说明对研究对象如何施盲,但计划书中详细介绍了意外揭盲的条件,可推测本研究已对研究对象施盲。
	2.2 研究过程中医护人员和研究人員是否知道研究对象接受哪种干预?	可能不是	“When code is broken the treatment allocation will be accessible to the Investigator and the department responsible for global product safety, Novo Nordisk, or any other relevant party [i.e., Data Monitoring Committee (DMC)].” 研究计划书 59 页 10.2 部分。 与上个问题回答类似。尽管没有明确说明对医护人员和研究人員的施盲方法,但明确说明了意外揭盲时需要研究人員向安全委员会等机构报告,可推测对这些人員也施盲。
	2.3 如在 2.1 或 2.2 问题中回答是/可能是/不清楚则回答:是否存在非常规的干预改变?	不适用	
	2.4 如在 2.3 问题中回答是/可能是则回答:干预变化是否在组间不平衡,并可能对研究结局造成影响?	不适用	
	2.5 是否有研究对象没有按照其分配的干预措施进行分析?	不是	“Full Analysis Set (FAS): includes all randomised subjects. The statistical evaluation of the FAS will follow the intention-to-treat (ITT) principle and subjects will contribute to the evaluation “as randomised”.” “Analysis of the primary and the secondary efficacy endpoint will be based on the FAS.” 研究计划书 76 页 17 部分,研究计划书 78 页 18.2.1 部分和发表文献表 1。主要结局的分析中符合 ITT 分析原则。
	2.6 如 2.5 问题回答是/可能是/不清楚则回答:错误的分组是否会对分析结果产生影响?	不适用	
	偏倚风险评估	低风险	
	可选:如存在偏倚,可能的方向是什么?	不适用	
Assess the effect 系统综述的目的是评估干预依从的效果。(实际使用中选择干预分配或干预依从其中一种目的即可)			
偏离既定干预的偏倚	2.1 研究过程中研究对象是否知道他们接受哪种干预?	可能不是	“The code for a particular subject may be broken in a medical emergency if knowing the identity of the treatment allocation would influence...” 研究计划书 59 页 10.2 部分。 尽管各文件中没有明确的直接说明对研究对象如何施盲,但计划书中详细介绍了意外揭盲的条件,可推测本研究已对研究对象施盲。
	2.2 研究过程中医护人员和研究人員是否知道研究对象接受哪种干预?	可能不是	“When code is broken the treatment allocation will be accessible to the Investigator and the department responsible for global product safety, Novo Nordisk, or any other relevant party [i.e., Data Monitoring Committee (DMC)].” 研究计划书 59 页 10.2 部分。 与上个问题回答类似。尽管没有明确说明对医护人员和研究人員的施盲方法,但明确说明了意外揭盲时需要研究人員向安全委员会等机构报告,可推测对这些人員也施盲。

续表

领域	信号问题	回答	支持信息/理由
偏离既定干预的偏倚	2.3 如在2.1或2.2问题中回答是/可能是/不清楚则回答:重要的伴随干预措施组间是否均衡?	不适用	
	2.4 计划实施的干预措施是否成功执行?	可能是	“The mean percentage of time that patients received the trial regimen was 84% for liraglutide and 83% for placebo.” 发表文献结果部分。 本研究欲评估的干预措施是利拉鲁肽和安慰剂治疗,从结果上看该干预成功在多数研究对象中实施。因利拉鲁肽为注射治疗药物,在已知多数患者依从了既定干预的情况下,可推测干预措施成功实施。
	2.5 研究对象是否依从了分配的干预措施?	是	“The mean percentage of time that patients received the trial regimen was 84% for liraglutide and 83% for placebo.” “The median daily dose of liraglutide was 1.78 mg (interquartile range, 1.54 to 1.79), including periods during which the patients did not receive liraglutide.” 发表文献结果部分。 本研究欲评估的干预措施是利拉鲁肽和安慰剂治疗,依从率较高;研究者计划实施的干预剂量为1.8 mg,试验结束后,患者实际接受的平均药物剂量是1.78 mg,非常接近预计的干预剂量。据此,可以推测该研究中分配干预依从性较好。
	2.6 如在2.3,2.4或2.5问题中回答不是/可能不是/不清楚则回答:研究是否采用了恰当的方法估计干预的依从效果? 偏倚风险评估 可选:如存在偏倚,可能的方向是什么?	不适用 低风险 不适用	

注:支持信息/理由中加“”的内容均为研究发表文献、计划书、注册网站等来源的原文摘录

度来说可能仍待商榷,特别是对于偏离既定干预的偏倚评估,在两种不同综述目的下分别有不同的信号问题,如何提高这两类问题的区分度,需要开发团队进一步说明并讨论。

综上所述,建议系统综述作者持续关注RoB2.0的更新,以及Cochrane手册和Review Manager的相应更改的内容。在有条件的情况下,可积极地与开发团队沟通,及时反馈在对工具的理解和使用等方面遇到的问题与建议,使得RoB2.0能够更好地在证据整合乃至循证医学实践中体现其意义和价值。

利益冲突 无

参 考 文 献

[1] Badgley RF. An assessment of research methods reported in 103 scientific articles from two Canadian medical journals [J]. Can Med Assoc J, 1961, 85(5): 246-250.
[2] Moher D, Jadad AR, Nichol G, et al. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists [J]. Control Clin Trials, 1995, 16(1): 62-73. DOI:

10.1016/0197-2456(94)00031-W.
[3] Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions: cochrane book series [M]. Chichester: Wiley, 2008. DOI: 10.1002/9780470712184.
[4] Higgins JPT, Altman DG, Gøtzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials [J]. BMJ, 2011, 343: d5928. DOI: 10.1136/bmj.d5928.
[5] Higgins JPT, Sterne JAC, Savović J, et al. A revised tool for assessing risk of bias in randomized trials [M]//Chandler J, McKenzie J, Boutron I, et al. Cochrane Methods. Cochrane Database Syst Rev, 2016, 10 Suppl 1: S29-31.
[6] Marso SP, Daniels GH, Brown-Frandsen K, et al. Liraglutide and cardiovascular outcomes in type 2 diabetes [J]. N Engl J Med, 2016, 375(4): 311-322. DOI: 10.1056/NEJMoa1603827.
[7] Marso SP, Poulter NR, Nissen SE, et al. Design of the liraglutide effect and action in diabetes: evaluation of cardiovascular outcome results (LEADER) trial [J]. Am Heart J, 2013, 166(5): 823-830.e5. DOI: 10.1016/j.ahj.2013.07.012.

(收稿日期:2017-03-10)

(本文编辑:王岚)