

• 方法学 •

核心指标集报告规范：COS-STAR 声明



张明妍, 杨丰文, 李越, 张俊华, 译

天津中医药大学中医药研究院循证医学中心(天津 300193)

【摘要】 **背景** 通过确保对卫生服务使用者和其他特定领域卫生保健决策者关注的指标进行常规测量, 核心指标集 (core outcome set, COS) 可以提高研究的相关性。到目前为止, 已经形成了 200 多个 COS, 但这些报告的透明度还不理想。如果 COS 的报告不够完整和透明, COS 研究将无法实现其目标。 **方法与结果** 鉴于这些问题, 由经验丰富的 COS 制定者、方法学家、期刊编辑、COS 潜在用户 (临床试验者、系统评价员和临床指南制定者) 以及患者代表组成的一个国际性工作组, 制定了核心指标集报告标准 (Core Outcome Set-STAndards for Reporting, COS-STAR) 声明, 作为 COS 研究报告的指南。其形成过程先后包括初始报告条目产生阶段、代表关键利益相关方群体的 200 名人员的两轮德尔菲问卷调查, 最后是一个共识会议。COS-STAR 声明由一个包含 18 个条目的清单组成, 这些条目是确保所有 COS 研究报告透明度和完整性的必要条目。条目清单重点规范了形成特定 COS 的研究报告需涉及的前言、方法、结果和讨论部分。COS-STAR 声明的一个局限性在于其形成过程中缺乏低、中收入国家的代表性意见。而 COS 对在这些国家和地区开展的研究具有同等的相关性, 故今后该指南可能需要不断完善, 以应对这些国家和地区 COS 形成的任何其他挑战。 **结论** 随着 COS 研究的不断开展, COS-STAR 声明将成为提高 COS 研究报告质量的有力工具, 并使所有 COS 使用者受益。

Core Outcome Set-STAndards for Reporting: The COS-STAR Statement

Translated by: ZHANG Mingyan, YANG Fengwen, LI Yue, ZHANG Junhua

Evidence-based Medicine Center, Institute of Traditional Chinese Medicine, Tianjin University of Traditional Chinese Medicine, Tianjin, 300193, P. R. China

1 前言

临床试验的结局指标没有得到足够的重视, 对这个问题的认识也不断加深。如果研究结果要影响实践和未来研究, 其结局指标需要与卫生服务使用者和其他卫生保健决策者紧密相关。此外, 结局指标报告偏倚, 即根据结果选择纳入试验报告的结局指标, 已被确定是影响发表数据解读的一个问题。

核心结局指标集 (core outcome set, COS) 的形成与实施, 目前引起了卫生研究所有领域越来越多的关注^[1], 并与全球范围内的研究实践密切相关。一项最新调查显示, 目前一些试验者、系统评价员和指南制定者 (COS 使用者) 将 COS 研究作为工作中结局指标选择的起点^[1]。COS 的使用可帮助提高

指标测量的一致性, 并能够减少结局指标报告偏倚。而报告偏倚会导致研究实施和报告过程中的大量不必要的浪费^[2]。

新近一篇系统评价确认已有 200 多个发表的 COS 研究^[1], 并且还有许多已知的在研 COS 研究^[3]。COS 形成的第一步应该是“测量什么”, 而“如何做”和“何时做”通常是下一步。COS 的价值在很大程度上取决于它“为什么做”和“如何做”。COS 制定工作组的可靠性, 与他们指标评价的经验和如何让主要利益相关群体参与制定过程相关, 这会影响到 COS 的后续采用。此外, COS 形成研究的报告质量也影响其推行。最近研究显示, 目前的报告质量参差不齐^[4], 限制了临床试验者、系统评价员和指南制定者等 COS 潜在用户评估与他们自己工作相关性的能力。

核心结局指标集之前被定义是健康或卫生保健特定领域中所有临床试验都必须测量和报告的, 最小的、公认的、标准化的指标集合^[2]。COS 正在研究用于临床试验以外的领域。尽管先前制定了

DOI: 10.7507/1672-2531.201702053

通信作者: 张俊华, Email: tjzyxy@126.com

译者申明: 本文翻译获得原文作者授权。

本文译自: PLoS Med, 2016,13(10): e1002148.

关于 COS 研究中德尔菲调查的报告建议^[5],但因在 COS 形成过程中包含了其他各种方法和内容,故也应及时广泛收集这一领域渐增的活动共识^[1]。在本文中,我们介绍了一个研究项目的结果,其目标是采用“提高卫生研究质量和透明度网络”(Enhancing the QUALity and Transparency Of health Research, EQUATOR)^[6]提供的方法制定 COS 形成研究的报告指南(Core Outcome Set-Standards for Reporting, COS-STAR)。报告清单不仅关乎 COS 形成中使用的共识方法,还可以应用于有效性试验、系统评价或常规治疗的 COS 形成^[7]。

2 术语

一个核心指标集描述了在特定研究或实践环境中应该测量的内容,下一步需要确定如何界定或测量每个结局指标。仅有 38% 的已发表研究中含有测量 COS 结局指标的推荐意见^[4]。制定这个报告指南是为了解决形成 COS 的第一阶段问题,即应该测量什么。

3 伦理审批

本研究已咨询利物浦大学伦理委员会并通过伦理批准(参考 RETH000841)。如果参与者回复了德尔菲调查或同意参加共识会议,则被认为是知情同意。

4 COS-STAR 声明的形成

一份概述了 COS-STAR 研究中德尔菲过程的完整方案已经发表^[8],包括提供一份配套的解释和阐释的文档(Explanation and Elaboration, E+E)的意向,其中给出了每个清单条目的含义和合理性。

初始清单包括 48 个报告条目,产生于以前一个关于 COS 研究的系统评价,包括德尔菲调查^[5]、制定 COS 的个人经验以及项目管理工作组报告(COS-STAR 工作组)。这份包括潜在报告条目的初始清单被纳入到一个国际性的两轮德尔菲调查中,以确定这些报告条目的重要性。

德尔菲调查涉及四个主要利益相关者群体,他们将参与到 COS 形成、报告和应用等方面。邀请函通过电子邮件发送到:(1)在核心指标集有效性试验数据库(Core Outcome Measures in Effectiveness Trials, COMET, www.comet-initiative.org)中已发表的 196 个 COS 研究的主要作者,并请他们将邀请函转发给参与的所有方法学家;(2)250 种发表过 COS 研究和 70 种涉及女性健康(CoRe Outcomes in

Women's health, CROWN, <http://www.crown-initiative.org/>)核心指标的期刊编辑,主动支持采用 COS;(3)COS 潜在用户:①商业或非商业、在 clinicaltrials.gov 数据库注册(来自 8 449 项注册试验的 20% 随机样本)的开放性 III/IV 期试验主要研究者,②来自 53 个 Cochrane 评价小组的 76 位协调编辑;(4)33 名患者代表,先前参加过 COMET 专题会和 COMET 公众与患者参与、介入或协同工作组(Public and Patient Participation, Involvement and Engagement, PoPPIE)。邀请在指南国际协作网网站(<http://www.gin2015.net/about/>)上列出的临床指南制定者参与的可行性不大,因为他们是成员组织。因此调查问卷中增加了涉及临床指南制定的问题作为替代办法。目的是从每个利益相关者群体中招聘尽可能多的人员来进行德尔菲调查。给参与者发送一封个性化电子邮件,概述项目情况-类似于研究计划中的描述和 COS 研究的第一个系统评价^[4]。

德尔菲参与者评价每个报告条目的重要性,方案中界定尺度从 1 分(不重要)到 9 分(至关重要)^[8]。在第一轮德尔菲调查中,参与者可建议将新的报告条目纳入到第二轮调查中。在第二轮调查中,给参与第一轮的所有参与者展示每个条目的回复数量和分数分布情况,不同相关方分别提供,连同自己在内的第一轮的评分。第一轮德尔菲调查中建议增加的 9 个报告条目,也用于第二轮评分。根据前期规定^[8],每个相关方至少 70% 的投票参与者评分在 7~9 分之间,则被认为达成共识。25 名 COS 制定者、107 名 COS 使用者、40 名医学期刊编辑和 11 名患者代表参与了两轮德尔菲研究,其中 13 人曾参与过临床指南制定。因第二轮的反馈是按组表述的,则每个利益相关群体反馈人数变化不会影响结果。德尔菲过程运用 COMET(<http://www.comet-initiative.org/delphimanager/>)开发的 DelphiManager 软件进行控制和管理。两轮德尔菲研究的匿名数据,按照利益相关组逐项列出,可在“S1 Delphi Data”中获取。

共识会议已于 2016 年 1 月在英国伦敦举行,为期一天,共有 17 名国际代表参加,包括 6 名 COS 制定者、4 名医学期刊编辑、5 名 COS 使用者[1 名试验者、2 名 Cochrane 系统评价协作编辑和 2 名临床指南制定者]和 2 名患者代表。参会个人的遴选和邀请使用以下宽泛的原则:①完成两轮德尔菲过程的参与者;②四个利益相关者群体之间保持平衡;③使用德尔菲调查方法与使用其他方法的

COS 制定者之间保持平衡；④ 试验者、系统评价员及指南制定者等 COS 使用者之间保持平衡；⑤ 合理的地理分布。最初的邀请是由作者根据他们个人的专业知识确定的。如果某个人不能参加，将尽可能地让来自相同利益相关群体的其他人来代替。

每个参与者的作用不存在冲突，综合了临床与方法学经验。会议的目的是对 57 个潜在重要的可能纳入 COS-STAR 清单的报告条目进行讨论并投票表决。3 名额外的参与者（一名推广员和两名记录员）出席了会议，但没有参加讨论或投票。

所有 57 个报告条目的德尔菲结果在会前和会议期间均提供给共识会议代表（S1 Consensus Matrix Delphi）。共识会议的幻灯片展示了每个利益相关者群体的应答率、地理分布和第二轮德尔菲结果，可在 S1 Consensus Meeting Presentation 中获取。虽然调查的回复率低，但超过 86%（183/214）的第一轮参与者完成了第二轮，并且两轮间没有失访偏倚的证据（S1 Consensus Meeting Presentation）。

不能参加会议的 COS-STAR 工作组成员（研究者）需要提前联系与他们讨论结果，他们的意见会被记录并纳入会议讨论。每个可能报告条目的德尔菲结果介绍后，接着参会代表进行讨论并投票，如果超过 70% 的投票者（17 个投票者中至少 12 个）打分在 7~9 分之间，则保留该条目。投票使用 OMBEA 应答系统（<http://www.ombea.com/>）进行。在讨论共识过程第二个条目时，“介绍起初呈递给参与者的共识过程信息”，一位代表提出缺少一个伦理审批相关条目。工作组同意就此进行表决，并达成一致意见，故将此条目列入报告指南（S1 Consensus Meeting Critical Scores）。因在 COS 形成过程中纳入患者参与不断增加，所以这个问题变得更具实质意义。

管理工作组考虑报告清单格式的讨论。确认了与 COS 报告相关的必要条目，不论是为有效性试验、系统评价还是常规治疗形成的 COS 都要满足。鉴于以前报告指南形成的经验，会上代表提出了一个警示：避免第一个指南太过严格导致制定者不使用的风险增加。一些报告条目需要合并，增加一些子条目，以及建议另列一份解释性文件来提高。最终 COS-STAR 清单的有效性都得到考虑。具体来说，会后 COS-STAR 工作组做了以下修改：

（1）参与者（方法）：这个主题下的四个条目存在交叉。建议被提出并被工作组采纳，用一个包括“谁、如何以及为什么”的条目更有价值。

（2）共识过程：普遍认为，虽然主题下的五个条目都可以提供有用的信息，但有些条目太具体而无需纳入，承认没有金标准的途径。工作组提出并接受了这个建议，认为一个条目包括共识过程的描述更容易被参考使用。

（3）指标打分：建议将两方面都包括到单个条目中。

（4）参与者（结果部分）：建议将一些问题都合并到单个条目中。

（5）指标结果：普遍认为，因许多条目过于详细，应该合并到一起。这些问题在解释性文件中详细阐述（S1 Explanation and Elaboration）。

（6）局限性：一致认为，纳入一个与局限性有关的条目并在解释性文件中提供例子的做法较好。

会后，COS-STAR 清单草案被发给 COS-STAR 工作组和其他共识会议参会者。考虑了所有评论和修改意见，清单也作了相应地修订。经过多次反馈收集和清单完善的过程，直到不需要进一步改变。

最终清单条目的合理性得到两位指南制定专家（Douglas G. Altman, David Moher）的审查。检测由两个正撰写 COS 研究报告的 COS 制定者执行。正在制定 COS 的两名研究人员也审阅了该声明。测试人员独立于 COS-STAR 制定过程。这次测试反馈支持 COS-STAR 清单的最终版本。

5 COS-STAR 声明

表 1 中所提供的 18 个条目的 COS-STAR 清单适用于 COS 制定研究，其目的是决定 COS 应纳入哪些结局指标，但不包括这些指标如何定义或测量等工作的报告。这份清单旨在涵盖此类研究最少的报告要求，包括背景、领域、方法、结果、结论和局限性。在配套的解释性文件中（S1 Explanation and Elaboration），为每个清单条目的含义和合理性提供了说明。这份清单设计注重实用性，对在之前的系统评价中已经确定的问题不做限制^[4]，如 COS 形成共识的方法学（包括混合方法）和结局指标选择可能涉及的各参与群体（包括患者代表）。COS-STAR 清单为 COS 研究最小的报告提供指导，但在某些情况下，研究者可酌情增加额外的报告条目。例如，如果研究者认为重要，可能希望描述确定如何测量核心结局指标的步骤^[9]。

6 讨论

COS-STAR 清单的形成运用了制定医学报告指南推荐的方法^[6]。COS-STAR 清单的目的是提高

表 1 核心指标集报告规范: COS-STAR 声明

节/主题	条目编号	清单条目
文题/摘要		
文题	1a	从文题能识别文章报告的是 COS 制定
摘要	1b	提供结构式摘要
引言		
背景和目的	2a	介绍研究背景并解释制定 COS 的合理性
	2b	介绍制定 COS 的确切目的
范围	3a	描述 COS 涉及的健康问题和人群
	3b	描述 COS 涉及的干预措施
	3c	描述 COS 适用的条件
方法		
方案/注册登记	4	指出在哪里可以获取 COS 制定方案和(或)研究注册信息, 如果有这些信息
参与者	5	描述各利益相关组参与 COS 形成过程的合理性, 各组参与者的合格标准, 并描述相关的参与者是如何产生的
信息源	6a	描述用于产生初始结局指标清单的信息源
	6b	描述指标被排除或合并的方法及原因(如果适用的话)
共识过程	7	描述共识过程如何执行
指标评分	8	描述指标如何评分以及如何总结评分
共识定义	9a	描述共识定义
	9b	描述共识过程中如何考虑纳入或排除结局指标的确定程序
伦理和知情同意	10	提供关于研究伦理和知情同意问题的说明
结果		
方案偏离	11	描述对方案所作的任何改变(如果有)并说明原因, 并描述这些更改对结果的影响
参与者	12	提供 COS 制定所有阶段涉及人员的数量和相关特点等数据
指标	13a	列出所有在共识会议开始时考虑的结局指标
	13b	描述共识过程中引入的任何新指标和被排除掉的指标, 并说明原因
核心指标集	14	列出最终 COS 包括的指标
讨论		
局限性	15	讨论 COS 形成过程中的任何不足
结论	16	结合其他证据, 提供对最终 COS 的解释, 以及对今后研究的影响
其他信息		
资助	17	描述资助的来源/资助者的作用
利益冲突	18	描述研究团队内的任何利益冲突, 以及如何控制这些冲突

COS 研究报告的透明性和完整性, 以便 COS 使用者能够判断这个推荐的集合是否与他们的工作相关。例如, 尽管 COS 的范围描述日益充分^[1], 但是仍需改进以便最大限度推广应用并不断评估。

制定 COS-STAR 配套解释文件 (S1 Explanation and Elaboration), 提供每个 COS-STAR 清单条目的解释和良好报告实践框架, 并配以实例为有兴趣开展 COS 研究和报告形成工作的研究者服务。格式与其他解释性文件^[10, 11]中使用的格式类似。解释性文件还描述了计划好的 COS-STAR 声明的认可和推广应用策略。

COS-STAR 不是质量评估工具, 也不应用于这方面, 例如, 比较类似 COS 相关性研究的正确性。同样, 该清单也不推荐 COS 研究项目中达成共识应该使用的方法学或纳入哪些利益相关群体。这

些问题的相关指导可以在其他地方找到^[2, 12]。例如, 有几个研究已经考虑制定儿童哮喘相关的 COS, 每个研究使用不同的方法且产生了具有轻微差异的核心结局指标^[13-16]。尽管 COS-STAR 声明可能有助于对已发表的 COS 进行严格的评价, 但其不能从不同 COS 中区分哪个应该被使用。

与经过多次修订的类似报告指南^[17]一样, COS-STAR 是一个不断发展的指南, 未来可能需要更好地完善。共识会议代表承认: 对考虑纳入或排除一些与报告条目相关的经验证据和方法学发展方面还有很多不足, 这些问题今后需要更好地关注。例如, 有证据表明, 反馈方法确实会影响人们如何为结局指标评分^[18], 这或许表明这是重要的报告信息。然而, 该条目被从报告清单中排除, 除非有更好的关于 COS 制定者应该如何表述反馈的指南。

随着具有相关利益和经验的新利益相关方出现,该指南可能也需要修正。例如,监管机构最近建议在哮喘患者的相关医疗产品的试验中使用COS^[9],这可能促进对声明提出反馈意见。德尔菲调查与共识会议的参与人员是COS研究形成者代表,他们的地理分布主要是在北美和欧洲^[4]。这是一个公认缺陷,因为COS研究和本报告指南都应该考虑到低收入与中等收入国家的关联性。邀请患者代表参加,而不是患者。由于这是一个相对复杂的方法领域,选择对COS形成参与有一定程度了解的人员,因此需要加强他们的理解。随着COS形成研究中患者关注和参与程度增加,他们对未来修订本报告指南的贡献将会被寻求。COMET倡议的一个重要目标是促进更广泛的参与。

虽然德尔菲邀请的接受率可能较低,但每个利益相关者群体第二轮的参与度高达85%以上,且没有失访偏倚证据(SIConsensus Meeting Presentation)。邀请读者通过COMET网站(<http://www.comet-initiative.org/contactus>)提出评论、批评、经验和见解,这些会被考虑用于将来完善COS-STAR声明。

参考文献

- Gorst SL, Gargon E, Clarke M, *et al.* Choosing important health outcomes for comparative effectiveness research: an updated review and user survey. *PLoS One*, 2016, 11(1): e0146444.
- Williamson PR, Altman DG, Blazeby JM, *et al.* Developing core outcome sets for clinical trials: issues to consider. *Trials*, 2012, 13: 132.
- COMET Initiative 2016. Available at: <http://www.comet-initiative.org/>.
- Gargon E, Gurung B, Medley N, *et al.* Choosing Important Health Outcomes for Comparative Effectiveness Research: A Systematic Review. *PLoS One*, 2014, 9(6): e99111.
- Sinha I, Smyth RL, Williamson PR. Using the Delphi technique to determine which outcomes to measure in clinical trials: recommendations for the future based on a systematic review of existing studies. *PLoS Med*, 2011, 8(1): e1000393.
- Moher D, Schulz KF, Simera I, *et al.* Guidance for developers of health research reporting guidelines. *PLoS Med*, 2010, 7(2): e1000217.
- Clarke M, Williamson PR. Core outcome sets and systematic reviews. *Syst Rev*, 2016, 5: 11.
- Kirkham JJ, Gorst S, Altman DG, *et al.* COS-STAR: a reporting guideline for studies developing core outcome sets (protocol). *Trials*, 2015, 16: 373.
- Schmitt J, Spuls PI, Thomas KS, *et al.* The Harmonising Outcome Measures for Eczema (HOME) statement to assess clinical signs of atopic eczema in trials. *J Allergy Clin Immunol*, 2014, 134(4): 800-807.
- Moher D, Hopewell S, Schulz KF, *et al.* CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trial. *BMJ*, 2010, 340: c869.
- Liberati A, Altman DG, Tetzlaff J, *et al.* The PRISMA Statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and Elaboration. *PLoS Med*, 2009, 6(7): e1000100.
- Schmitt J, Apfelbacher C, Spuls PI, *et al.* The Harmonizing Outcome Measures for Eczema (HOME) roadmap: A methodological framework to develop core sets of outcome measurements in dermatology. *J Invest Dermatol*, 2015, 135(1): 24-30.
- Smith MA, Leeder SR, Jalaludin B, *et al.* The asthma health outcome indicators study. *Aust N Z J Public Health*, 1996, 20(1): 69-75.
- Sinha I, Gallagher R, Williamson PR, *et al.* Development of a core outcome set for clinical trials in childhood asthma: a survey of clinicians, parents, and young people. *Trials*, 2012, 13: 103.
- Reddel HK, Taylor DR, Bateman ED, *et al.* An official American Thoracic Society/European Respiratory Society statement: asthma control and exacerbations: standardizing endpoints for clinical asthma trials and clinical practice. *Am J Respir Crit Care Med*, 2009, 180(1): 59-99.
- Busse WW, Morgan WJ, Taggart V, *et al.* Asthma outcomes workshop: overview. *J Allergy Clin Immunol*, 2012, 129(3 Suppl): 1-8.
- Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *PLoS Med*, 2010, 7(3): e1000251.
- Brookes S, Macefield R, Williamson P, *et al.* Three nested RCTs of full or single stakeholder feedback within Delphi surveys during core outcome and information set development. *Trials*, 2015, 16(Suppl 2): 51.
- Committee for medicinal products for human use (CHMP) (2015). Guideline on the clinical investigation of medicinal products for the treatment of asthma. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/12/WC500198877.pdf.

收稿日期: 2017-02-22 修回日期: 2017-03-07
本文编辑: 蔡羽嘉