

加强流行病学中观察性研究报告质量 (STROBE) 声明: 观察性研究报告规范

The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies

Erik von Elm, Douglas G Altman, Matthias Egger, Stuart J Pocock, Peter C Gtzsche, Jan P Vandembroucke, for the STROBE initiative

很多生物医学研究都是观察性的。但对这类研究的报告却通常不完整,从而限制了对其优缺点及普适性的评价。“加强流行病学中观察性研究报告质量 (STROBE)”工作组针对一篇完整明确的观察性研究报告应包括的内容提供了一套建议。建议涵盖了流行病学研究的三种主要设计类型:队列设计、病例对照设计和横断面设计。2004年9月,我们召开了为期2天的研讨会,与会的方法学专家、研究人员和期刊编辑共同起草了一份观察性研究报告应纳入的条目清单草案。此后,又通过协调小组数次会议及与有关人员电子邮件讨论等对该草案进行了修改,纳入了经验性证据和方法学方面的意见。通过研讨会及之后反复的咨询和修订,我们制定出了一个包含22个条目的清单(STROBE声明),分为论文的题目、摘要、引言、方法、结果和讨论等部分。其中18个条目适用于所有三种主要的观察性研究设计,其余4个条目则专门用于队列、病例-对照或横断面设计。详细的解释和示范文件已另行发表,在《公共科学图书馆-医学》(PLoS Medicine)、《内科学年报》(Annals of Internal Medicine)和《流行病学》(Epidemiology)的网站可以免费获得。我们希望STROBE声明能有助于改善观察性研究报告的质量。

引言

许多医学研究中的问题是采用观察性研究来探讨的^[1]。大多数对致病因素的研究都依赖于队列、病例-对照或横断面设计研究。这类观察性研究对于医疗干预的利弊研究也同样有用^[2]。随机化试验不可能回答有关特定干预措施的所有重要问题。例如,观察性研究更适用于分析某种治疗的罕见或远期不良作用,更可能为常规治疗效果提供信息^[3]。

研究报告应该透彻明晰,以便读者能够了解研究计划是什么、研究过程怎样、发现了什么、得出了怎样的结论。研究的可信性取决于他人对研究设计、执行和分析的优缺点的批判性评价。一篇明确的研究报告还能帮助他人决定是否以及如何将其研究结果纳入系统评价中^[4-5]。可是,在已发表的观察性研究报告中,重要的信息往往缺失或含糊不清。一项对发表在综合医学期刊或专业期刊上的流行病学研究报告的分析显示,这些报告通常不解释为何会选择某些可能导致混杂效应的变量^[6]。在精神病学文献中,只有少数病例-对照设计的研究报告解释了确立病例和对照的方法^[7]。一项对卒中文献中的纵向研究报告的分析显示,49篇文章中有17篇(35%)没有详细说明受试者的纳入标准^[8]。有人认为如果研究报告不具有足够的明确性,研究的效益将不易实现^[9],所以有必要制定一个撰写观察性研究报告的规范^[10-11]。

研究报告规范能够改善报告的质量。强化临床试验报告(CONSORT)声明制定于1996年,于5年后修订^[12]。CONSORT声明得到了许多医学期刊的支持^[13],促进了随机试验报告质量的提高^[14-15]。其他研究领域也随之建立了相似的标准,如随机试验^[16]或诊断性研究^[17]的荟萃分析报告标准。我们组建了一个由方法学家、研究人员和期刊编辑组成的国际性合作小组,着手制定观察性研究报告规范,即加强流行病学中观察性研究报告质量(STROBE)声明。

STROBE 声明的目的和应用

STROBE声明是指流行病学中采用三种主要设计——队列设计、病例-对照设计和横断面设计——的观察性研究报告应纳入的条目清单。声明旨在为全面完善地报告观察性研究提供指导,而不是为研究的设计或实施本身提供建议。而且,尽管研究报告的明确性是评价其质量的前提条件,但STROBE清单并不能作为评价观

察性研究质量的工具。

本文介绍 STROBE 声明的内容,解释其制定过程。在另一篇详细的解释和示范文章(与本文相配套)中^[18-20],论证了不同清单条目的纳入,同时提供了方法学背景及已发表的我们认为符合明确性要求的报告范例。我们强烈推荐结合该说明文件使用 STROBE 声明,说明文件可在 PLoS Medicine(www.plosmedicine.org)、Annals of Internal Medicine(www.annals.org)和 Epidemiology(www.epidem.com)的网站免费获得。

STROBE 声明的制定过程

2004,我们组建了 STROBE 工作组,获得资助并召开了研讨会,同时建立了网站(www.strobe-statement.org)。我们通过检索教科书、图书目录数据库、参考文献列表和个人档案获得了相关资料,包括之前的推荐方案、报告撰写的经验性研究和描述相关方法学研究的论文。由于观察性研究可使用多种不同的设计方法,我们认为从一开始就应该明确规定 STROBE 清单的适用范围。我们决定主要针对分析性观察研究中最常用的三种研究设计,即队列设计、病例-对照设计和横断面设计。

2004年9月,我们在英国 Bristol 召开了为期2天的研讨会。参加会议的共有23人,包括 Annals of Internal Medicine、BMJ、Bulletin of the World Health Organization、International Journal of Epidemiology、JAMA、Preventive Medicine 和 The Lancet 的编辑人员以及来自欧洲和北美的流行病学专家、方法学专家、统计学专家和相关工作者。另外10位对 STROBE 有兴趣但未能到会的学者通过书面形式参与了讨论。与会人员被分成3个小组,分别针对三种研究设计确定非常重要而应纳入清单的条目。我们为各组分发了预先拟订的条目清单草案(在 STROBE 网站可获得)以供讨论。之后,所有与会者再对3个清单草案进行了集体讨论,并修订条目使其尽可能适用于所有的三种设计方法。最后,全体成员还制定了对 STROBE 声明最终定稿和宣传推广的策略。

会议之后,我们起草了一个包括三种设计方法的综合条目清单,并将其公布在 STROBE 网站上。我们邀请与会者及其他专家和期刊编辑对这一清单草案进行评价。随后,我们在网站上发表了三版的 STROBE 声明和两版的评价汇总及修订变化。在这个过程中,协调小组(即本文作者)进行了八次为期1~2天的会议以及数次电话会议,对清单进行了修订,并着手撰写本文及解

释和示范文件^[18-20]。协调小组邀请另外3位有方法学专长和编辑经验的作者共同撰写解释和示范文件,该文件撰写过程中还采纳了30多位学者的反馈意见。初稿完成后,我们留出数周时间供专家评价,并通过电子邮件提醒协作者截止日期。

STROBE 组成

STROBE 声明是由22个条目组成的清单,我们认为这些条目是优质的观察性研究报告必备的重要内容(附表)。这些条目分别针对论文的题目和摘要(条目1)、引言(条目2~3)、方法(条目4~12)、结果(条目13~17)和讨论(条目18~21)以及其他信息(条目22,关于研究资金)等。18个条目适用于所有三种研究设计,其余4个条目(条目6、12、14和15)则根据设计类型而定,并且对于全部或部分条目而言存在不同版本。对一些条目而言(用“*”表示),病例-对照研究中病例和对照的资料以及队列和横断面研究中暴露组和未暴露组的资料要分别描述。尽管本文只列出了一个清单,但在 STROBE 网站可以获得三种设计方案相应的条目清单。

意义和局限性

STROBE 声明旨在帮助作者撰写分析性观察研究报告,协助编辑和同行评审人决定该论文可否发表,以及帮助读者严格评价已发表论文。我们采用了一种开放的方式来制定 STROBE 清单,吸收了以往制定其他标准——尤其是 CONSORT 声明——的经验。我们回顾了相关的经验性证据和方法学研究结果,制定出草案后又各版本公之于众,进行了广泛的、反复的咨询和修改。因此这个清单是基于学科背景不同、视角也各不相同的大量专家学者反复论证后形成的。与清单同时使用的详细的解释性文件也非常得益于同一咨询过程。

观察性研究可以服务于从探索新发现到证实或推翻旧发现的多种不同性质的研究目的^[18-20]。一些研究是探索性的,目的是提出令人感兴趣的假设,另一些研究则旨在用已有数据验证假设。还有一类研究则是根据已有假设仔细设计新数据的收集方法。我们认为,本文发表的清单对上述所有研究都适用,因为读者总是需要了解研究计划是什么(和不包括什么)、研究过程怎样、发现了什么以及结果有什么意义。我们承认目前 STROBE 声明还局限于三种最常用的观察性研究设计,但我们希望能够对其进行扩展,使清单适用于其他设计

附表 STROBE 声明:观察性研究报告中应当纳入的条目清单

条目	建议
题目和摘要	
1	(a)在题目或摘要中用常用术语表明研究所采用的设计 (b)在摘要中对所做工作和获得的结果做一个简明的总结
引言	
背景/原理	2 解释研究的科学背景和原理
目的	3 阐明具体研究目的,包括任何预先设定的假设
方法	
研究设计	4 尽早陈述研究设计的关键内容
研究设置	5 描述研究机构,研究地点及相关资料,包括招募的时间范围、暴露、随访和数据收集等
参与者	6 (a)队列设计——描述纳入标准,参与者的来源和选择方法,随访方法 病例-对照设计——描述纳入标准,病例和对照的来源及确认病例和选择对照的方法,病例和对照选择的原理 横断面设计——描述纳入标准,参与者的来源和选择方法 (b)队列设计——对于配对设计,应说明配对标准及暴露和非暴露的人数 病例-对照设计——对于配对设计,应说明配对标准和每个病例配对的对照数
变量	7 明确定义结局、暴露、预测因子,可能的混杂因素及效应修饰因素,如果相关,给出诊断标准
数据来源/测量	8* 对每个有意义的变量,给出数据来源和详细的测量方法。如果有一个以上的组,描述各组之间测量方法的可比性
偏倚	9 描述解决潜在偏倚的方法
样本大小	10 描述样本量的确定方法
定量变量	11 解释定量变量是如何分析的。如果相关,描述分组的方法和原因
统计方法	12 (a)描述所用的所有统计方法,包括减少混杂因素的方法 (b)描述所有分析亚组和交互作用的方法 (c)解释如何解决数据缺失 (d)队列设计——如果相关,描述解决失访问题的方法 病例-对照设计——如果相关,描述如何对病例和对照进行配对 横断面设计——如果相关,描述考虑到抽样策略的分析方法 (e)描述所用的灵敏性分析方法
结果	
参与者	13* (a)报告研究各阶段参与者的人数,如可能合格的人数,参与合格性检查的人数,证实合格的人数,纳入研究的人数,完成随访的人数及完成分析的人数 (b)解释在各阶段参与者退出研究的原因 (c)考虑使用流程图
描述性数据	14* (a)描述参与者的特征(如人口统计学、临床和社会特征)以及暴露和潜在混杂因素的相关信息 (b)描述就每一个待测变量而言缺失数据的参与者人数 (c)队列设计——总结随访时间(如平均随访时间和全部随访时间)
结局数据	15* 队列设计——报告随时间变化的结局事件数或综合指标 病例-对照设计——报告各种暴露类别的人数或暴露综合指标 横断面设计——报告结局事件数或综合指标
主要结果	16 (a)报告未校正的估计值,如果相关,给出混杂因素校正后的估计值及其精确度(如95%可信区间)。阐明按照哪些混杂因素进行了校正以及选择这些因素进行校正的原因 (b)如对连续变量进行分组,要报告每组观察值的范围 (c)对有意义的危险因素,最好把相对危险转化成针对有意义的时间范围的绝对危险度
其他分析	17 报告进行过的其他分析,如亚组分析、交互作用分析和灵敏性分析

续表

	条目	建议
讨论		
关键结果	18	根据研究目标概括关键结果
局限性	19	讨论研究的局限性,包括潜在偏倚或不准确的来源。讨论任何潜在偏倚的方向和大小
解释	20	结合研究目标、研究局限性、多重分析、相似研究的结果和其他相关证据,谨慎给出一个总体的结果解释
可推广性	21	讨论研究结果的普适性(外推有效性)
其他信息		
资金来源	22	提供研究资金的来源和资助机构在研究中的作用。如果相关,提供资助机构在本文基于的初始研究中的作用

*在病例-对照研究中,分别给出病例和对照的信息,如果相关,在队列研究和横断面研究中给出暴露组和未暴露组的信息。在一篇详细的解释和示范文章中,讨论了清单中的每一个条目,提供了方法学背景及已发表的明确推荐的范例。STROBE 清单最好与这篇文章联合使用(在 *PLoS Medicine*, *Annals of Internal Medicine* 和 *Epidemiology* 的网站可免费获得)。在 STROBE 网站可获得队列研究、病例-对照研究和横断面研究各自相应的清单。

方法(如病例交叉研究或生态学研究)以及特定的领域。CONSORT 声明现在已有四个扩展版^[21-24]。STROBE 声明的第一个扩展版是针对基因-疾病相关性研究的,即基因相关性研究 STROBE 扩展版(STREGA),目前正在制定过程中^[25]。我们要求有意制定 STROBE 声明扩展版的个人或团体首先与协调小组联系,以免重复劳动。

STROBE 声明不应被解读为严格限定观察性研究报告撰写格式的规范。一篇报告应对清单中的条目有足够详细和清楚的描述,但条目的顺序和描述格式取决于作者的喜好、期刊类型以及各个研究领域的传统。例如,我们在讨论清单中结果的报告方式时,是按照一系列独立的条目逐一进行的,但是作者完全可以在文章中的一段或一个表格中同时陈述多个条目。同样,条目 22 有关研究资金来源和资助机构作用的内容可以在文章的附录或方法部分叙述。我们的目标并不是要使研究报告格式化。曾经有一位专业医学期刊的编辑要求随机临床试验研究报告的作者按照 CONSORT 声明格式撰写论文^[26]。我们认为就规范类型和使用术语而言,稿件不应该是 STROBE 式的。我们鼓励作者使用叙述性元素(包括对典型病例的描述),以补充有关研究的重要信息,使文章更具可读性^[27]。

我们强调,STROBE 声明并非为评价已发表的观察性研究报告的研究质量而制定。这种评价工具已经由其他小组制定,最近的一项系统回顾已经对其进行了评估^[28]。在解释和示范文件中,我们列举了一些对研究内容报告完好的论文范例,但这些研究的结果在进一步研究中并未获得证实——因为我们针对的是报告的完好性,而不是研究质量本身的优劣。但是,如果作者和期刊

都采用了 STROBE 清单,那么诸如混杂因素、偏倚和普适性等问题就会变得更为透彻明晰,这将有助于减少科学团体和大众传媒对新发现过分“热情”的报道^[29],从长远来看将改善研究的方法学问题。报告方式的改善还将有助于更明智地决定需要开展新研究的时间及所要达到的研究目标。

我们没有对清单中的每一个条目和子条目进行广泛的系统评价,也没有独立进行研究以便从实证角度进行补充。此外,尽管在清单制定过程中没有人被剔除,制定清单的人员组成仍受到已有关系网的影响,就地理位置而言也不具有代表性(主要由欧洲和北美的学者组成),就研究方向和学科而言也可能没有代表性。我们强调,STROBE 声明及其他用于研究报告撰写的推荐方案应该被视为需要长期评估、反复提炼、在必要时修订的成长性文件。我们非常乐于接受对 STROBE 声明进行更为广泛宣传的建议,如通过专业期刊和其他语种期刊再次发表本文。那些愿将清单翻译成其他语言的研究小组或个人应提前与协调小组沟通。我们将会根据评论、批评、新证据及使用经验对清单进行修订。恳请读者们在 STROBE 网站发表评论。

Lancet 2007; 370: 1453-57
(赵乐译 邹光勇 杜怀东 审)

致谢: 本文翻译后得到邹光勇教授和杜怀东博士的悉心审校,特此表示衷心感谢。

邹光勇

Assistant Professor

Dept of Epidemiology & Biostatistics, Schulich School of Medicine & Dentistry,
University of Western Ontario, London, ON, Canada N6A 5C1

杜怀东

PhD candidate, MD

Center for Nutrition and Health, National Institute of Public Health and the Environment (RIVM), the Netherlands

Acknowledgments

The workshop was funded by the European Science Foundation (ESF). Additional funding was received from the Medical Research Council Health Services Research Collaboration, and the National Health Services Research and Development Methodology Programme. We are grateful to Gerd Antes, Kay Dickersin, Shah Ebrahim, and Richard Lilford for supporting the STROBE initiative. We are grateful to the following institutions that have hosted working meetings of the coordinating group: Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland; Department of Social Medicine, University of Bristol, Bristol, UK; London School of Hygiene and Tropical Medicine, London, UK; Nordic Cochrane Centre, Copenhagen, Denmark; and Centre for Statistics in Medicine, Oxford, UK. We are grateful to six reviewers who provided helpful comments on a previous draft of this paper.

Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland (E von Elm MD, Prof M Egger MD); Centre for Statistics in Medicine, University of Oxford, Oxford, UK (Prof D G Altman DSc); Department of Social Medicine, University of Bristol, Bristol, UK (M Egger); London School of Hygiene and Tropical Medicine, University of London, London, UK (Prof S J Pocock PhD); Nordic Cochrane Centre, Copenhagen, Denmark (P C Gøtzsche MD); and Department of Clinical Epidemiology, Leiden University Hospital, Leiden, Netherlands (Prof J P Vandenbroucke MD)

Correspondence to: Dr Erik von Elm, Institute of Social and Preventive Medicine (ISPM), University of Bern, Finkenhubelweg 11, CH-3012 Bern, Switzerland (e-mail: strobe@ispm.unibe.ch)

参考文献

- [1] Glasziou P, Vandenbroucke JP, Chalmers I. Assessing the quality of research. *BMJ*, 2004, 328: 39-41
- [2] Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ*, 1996, 312: 1215-18
- [3] Papanikolaou PN, Christidi GD, Ioannidis JP. Comparison of evidence on harms of medical interventions in randomized and nonrandomized studies. *CMAJ*, 2006, 174: 635-41
- [4] Jüni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. *BMJ*, 2001, 323: 42-46
- [5] Egger M, Schneider M, Davey Smith G. Spurious precision? Meta-analysis of observational studies. *BMJ*, 1998, 316: 140-44
- [6] Pocock SJ, Collier TJ, Dandreo KJ, et al. Issues in the reporting of epidemiological studies: a survey of recent practice. *BMJ*, 2004, 329: 883
- [7] Lee W, Bindman J, Ford T, et al. Bias in psychiatric case-control studies: literature survey. *Br J Psychiatry*, 2007, 190: 204-09
- [8] Tooth L, Ware R, Bain C, Purdie DM, Dobson A. Quality of reporting of observational longitudinal research. *Am J Epidemiol*, 2005, 161: 280-88
- [9] Bogardus ST Jr, Concato J, Feinstein AR. Clinical epidemiological quality in molecular genetic research: the need for methodological standards. *JAMA*, 1999, 281: 1919-26
- [10] Anon. Guidelines for documentation of epidemiologic studies. Epidemiology Work Group of the Interagency Regulatory Liaison Group. *Am J Epidemiol*, 1981, 114: 609-13
- [11] Rennie D. CONSORT revised - improving the reporting of randomized trials. *JAMA*, 2001, 285: 2006-07
- [12] Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet*, 2001, 357: 1191-94
- [13] Moher D, Altman DG, Schulz KF, Elbourne DR. Opportunities and challenges for improving the quality of reporting clinical research: CONSORT and beyond. *CMAJ*, 2004, 171: 349-50
- [14] Plint AC, Moher D, Morrison A, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust*, 2006, 185: 263-67
- [15] Egger M, Jüni P, Bartlett C. Value of flow diagrams in reports of randomized controlled trials. *JAMA*, 2001, 285: 1996-99
- [16] Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Quality of Reporting of Meta-analyses*. *Lancet*, 1999, 354: 1896-900
- [17] Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *Ann Intern Med*, 2003, 138: 40-44
- [18] Vandenbroucke JP, von Elm E, Altman DG, et al, for the STROBE initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLoS Med*, 2007, 4: e297
- [19] Vandenbroucke JP, von Elm E, Altman DG, et al, for the STROBE initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Ann Intern Med*, (in press)
- [20] Vandenbroucke JP von Elm E, Altman DG, et al, for the STROBE initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Epidemiology*, (in press)
- [21] Ioannidis JP, Evans SJ, Gotzsche PC, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med*, 2004, 141: 781-88
- [22] Campbell MK, Elbourne DR, Altman DG. CONSORT statement: extension to cluster randomised trials. *BMJ*, 2004, 328: 702-08
- [23] Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. *JAMA*, 2006, 295: 1152-60
- [24] Gagnier JJ, Boon H, Rochon P, Moher D, Barnes J, Bombardier C. Reporting randomized, controlled trials of herbal interventions: an elaborated CONSORT statement. *Ann Intern Med*, 2006, 144: 364-67
- [25] Ioannidis JP, Gwinn M, Little J, et al. A road map for efficient and reliable human genome epidemiology. *Nat Genet*, 2006, 38: 3-5
- [26] Ormerod AD. CONSORT your submissions: an update for authors. *Br J Dermatol*, 2001, 145: 378-79
- [27] Schriger DL. Suggestions for improving the reporting of clinical research: the role of narrative. *Ann Emerg Med*, 2005, 45: 437-43
- [28] Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol*, 2007, 36: 666-76
- [29] Bartlett C, Sterne J, Egger M. What is newsworthy? Longitudinal study of the reporting of medical research in two British newspapers. *BMJ*, 2002, 325: 81-84

(责任编辑 张隆欣)