

使用常规收集卫生数据开展观察性研究的报告规范-RECORD 规范



聂晓璐, 彭晓霞, 译

首都医科大学附属北京儿童医院临床流行病学与循证医学中心(北京 100045)

【摘要】 常规收集卫生数据是指基于管理和临床目的且事先没有特定研究目标而收集的数据, 已被越来越多地用于研究。此类数据发展迅速, 可及性好, 但相关注意事项并未在现有的报告规范中被提及, 如加强观察性流行病学研究报告的声明(strengthening the reporting of observational studies in epidemiology statement, STROBE)。使用常规收集卫生数据开展观察性研究的报告规范(the report of studies conducted using observational routinely collected data, RECORD)可填补该空缺。RECORD 规范是 STROBE 规范扩展版, 其可用来提出针对使用常规收集卫生数据开展观察性研究有关报告的条件要求。RECORD 清单扩展了包括题目、摘要、前言、方法、结果、讨论和其他内容等需要在此类研究报告中包含的 13 个条目内容。该规范包括了清单、详尽的解释性信息以提高清单的使用。该规范还给出每条 RECORD 清单条目良好的报告实例。本文及其官网和留言板(<http://www.record-statement.org>)可提高 RECORD 规范的应用和理解。通过应用 RECORD, 作者、期刊编辑和同行评议者可促进研究报告的质量。

【关键词】 RECORD 规范; 指南

The reporting of studies conducted using observational routinely-collected health data (RECORD) statement

Translated by: NIE Xiaolu, PENG Xiaoxia

Clinical Epidemiology & Evidence-based Medicine Center, Beijing Children's Hospital, Capital Medical University, Beijing, 100045, P. R. China

1 前言

医疗卫生实践中产生数据的可用性在日益提高, 通过对疾病发病率和结局的监测, 改变了研究范畴。常规收集卫生数据的定义为: 不是针对开展研究前即确定的研究问题而收集的数据^[1]。这些数据来源包括广泛的研究资源供应(如疾病登记)、临床管理(如初级保健数据库)、卫生系统规划(如卫生管理数据)、临床护理文件(如电子健康记录数据仓库)或流行病学监测(如癌症登记和公共卫生报告数据)。在多种医疗保健机构和地理位置背景下产生的这些数据为临床医学、卫生服务规

划和公共卫生决策提供了创新、高效及成本-效益较高的研究可能^[2]。国际范围内, 政府和资助机构优先使用常规收集的健康数据作为改善患者护理、改革健康研究和提高卫生保健效率的工具^[3]。

虽然扩展的数据可用性为回答亟需解决的研究问题提供了重要机会, 但同时也给研究实施和评价及结果应用带来挑战。常规收集卫生数据来源的多样化及其数据结构的快速扩展给如何确定各个数据源的优势、局限和偏倚带来了挑战。常规收集数据的研究报告存在的不完整或不充分的特点可能会加重这些挑战。对基于使用常规数据源的研究进行系统分析发现诸多报告问题^[4]: 暴露和结局编码的有关信息报告不完整, 不同数据源的链接率的细节存在报告不足或缺失等。已有系统评价的结果提示对常规数据来源数据进行验证的研究报告质量不高^[5,6], 这可能掩盖偏倚, 妨碍开展 Meta 分析, 导致错误结论。

DOI: 10.7507/1672-2531.201702009

通信作者: 彭晓霞, Email: pengxiaoxia@bch.com.cn

译者申明: RECORD 规范中文版翻译得到 REDORD 指导委员会授权、认可发表, 且得到了加拿大卡尔加里大学 Hude Quan 教授和 Mingkai Peng 博士校对指导。

本文译自: PLoS Med, 2015, 12(10): e1001885.

框 1 人群术语的概念 (源人群、数据库人群和研究人群)

在RECORD规范中提到,在使用常规收集数据开展的研究中,共涉及3个不同层次的人群,即:①源人群:指数据库人群的来源,因此是研究者想要最终作出推断的人群;②数据库人群:即来源于源人群,且包含在数据库中具有数据记录的人;③研究人群:研究者使用代码和算法从数据库人群中识别得到的人群(图1)^[7]。例如,在临床实践研究数据链(Clinical Practice Research Datalink, CPRD)中,源人群为英国所有注册全科医生的人,数据库人群指包含在CPRD中的个体,而研究人群包括在具体研究中使用代码和算法从CPRD内选择的个体。



图1 使用常规收集数据开展的研究中涉及的人群层次

为改善报告质量,现已针对一系列研究设计和内容制定了许多报告指南^[8,9]。加强观察性流行病学研究报告的声明(Strengthening the reporting of observational studies in epidemiology statement, STROBE)是为提高观察性研究报告的透明度而制定的^[10,11],并被主流医学期刊广泛接受和认可。已有研究发现在文章编辑过程中使用STROBE确实提高了研究报告的质量^[12,13]。大多数使用常规收集数据开展的研究属于观察性研究,因此,STROBE声明同样适用。尽管STROBE声明可应用于所有观察性研究,但使用常规收集数据开展研究在报告中涉及的特有问題尚未提及。2012年,关注使用常规收集卫生数据开展研究的国际科学家小组和STROBE组织代表在伦敦的初级保健数据库研讨会上会面,对STROBE声明在常规收集卫生数据研究中的应用进行了讨论^[14,15]。会议认识到STROBE声明具体到使用此类数据开展研究报告时产生的差距,并对STROBE声明进行扩展达成共识。因此,依据该组织国际合作的流程发起倡议,起草使用常规收集卫生数据开展观察性研究的报告规范(the report of studies conducted using observational

routinely collected Data, RECORD)作为STROBE的扩展,旨在探索和提出有关使用常规收集卫生数据开展研究的特定报告事项。RECORD规范涉及了100余位国际利益相关者,包括研究人员、期刊编辑和数据使用者,即使用常规数据的研究结果为决策提供信息的利益相关者。制定RECORD规范的方法另有详述^[16],即根据报告指南的既定方法而制定^[17]。简言之,对利益相关者进行两轮调查,确定优先考虑纳入RECORD规范中的条目。随后工作委员会人员会面,撰写了该声明的文档。利益相关人员审核声明并提出反馈意见。最终版的清单及其解释性文件由指导委员会撰写,并经工作委员会审查和批准。STROBE指导委员会的成员均参与了RECORD规范的制定。

与STROBE方法一致,RECORD规范并非以推荐开展研究的方法为目的,而是为完善其报告,以确保读者、同行评审、期刊编辑和其他研究消费者能够评估研究的内部和外部真实性。通过改善使用日常收集卫生数据的研究报告质量,努力减少不清楚的研究报告并且遵守科学研究过程的原则:创新、透明和可重复^[18]。

2 RECORD清单及其解读

完整的RECORD清单见表1。由于RECORD清单是现有STROBE条目的扩展,因此将其补充在相应的STROBE条目旁。我们建议作者充分阐明清单中的每个条目,但不需要严格按照清单顺序进行描述。我们将RECORD清单条目的详细解释作为正文中的一部分提供。当不需要针对使用常规收集卫生数据的研究在STROBE之外额外增加清单条目时,需要的解释参见STROBE相应条目。

2.1 题目和摘要

RECORD 1.1: 在题目或摘要中明确使用的数据类型。可能的话,应包含所使用数据库的名称。RECORD 1.2: 如需要,应在题目或摘要中报告研究发生的地理位置和时间窗。RECORD 1.3: 如果研究实施了数据库的链接,应在题目或摘要中给予清晰阐述。

实例 ① “2010年结肠镜检查后穿孔和出血:基于法国综合健康保险数据(SNIIRAM)”^[19]。② “荷兰医院标准化死亡率(HSMR)方法和心脏外科手术:使用医院管理数据为基准的国家队列与临床数据库比较”^[20]。

解释 由于缺乏检索使用常规采集卫生数据这类研究的相关医学主题词(Medical Subject

表 1 由 STROBE 声明扩展的 RECORD 规范：在使用常规收集卫生数据开展观察性研究时应报告的条目清单

条目编码	STROBE 条目	RECORD 条目
题目和摘要		
1	① 在题目或摘要中使用常用术语体现研究设计； ② 在摘要中对所做研究及其发现进行详细且全面地总结。	RECORD 1.1：在题目或摘要中明确使用的数据类型，可能的话应包含所使用数据库的名称； RECORD 1.2：如需要，应在题目或摘要中报告研究发生的地理区域和时间窗；RECORD 1.3：如果研究实施了数据库的链接，应在题目或摘要中给予清晰阐述。
前言		
背景/原理	2 解释拟报告研究的科学背景和原理。	
目标	3 阐明具体目标，包括任何预先设定的假设。	
方法		
研究设计	4 在文章中尽早体现研究设计的要素。	
场景	5 描述数据收集的机构、地点和时间范围，包括招募研究对象的时期、暴露、随访和数据收集。	
研究对象	6 ① 队列研究：描述研究对象的选择标准、来源和方法，描述随访方法。病例-对照研究：描述研究对象的选择标准，来源以及确定病例和选择对照的方法，阐述选择病例与对照的原理。横断面研究：描述研究对象的选择标准、来源和方法。② 队列研究：对于匹配研究，报告匹配标准以及暴露与非暴露人数。病例-对照研究：对于匹配研究，报告匹配标准以及每个病例匹配的对照个数。	RECORD 6.1：应详细列出研究人群选择方法（例如用于识别研究对象的编码或算法）。如不可行，应给予解释；RECORD 6.2：任何针对选择研究人群编码或算法进行的验证研究均应标引参考文献。如果为本研究实施的验证没有在其他地方发表，需要提供验证方法的细节和结果；RECORD 6.3：如研究涉及数据库链接，考虑使用流程图或其他图表展示以说明数据链接过程，包括每一步骤中能实现数据链接的个体人数。
变量	7 清晰定义所有结局、暴露、预测因子、潜在混杂因素和效应修正因子。如可行，给出诊断标准。	RECORD 7.1：提供将暴露、结局、混杂因素和效应修饰因子进行分类的编码和算法的完整列表。如不能报告，应予以解释。
数据来源/测量	8 对每个研究变量，提供数据来源和详细的评价（测量）方法。如果存在两组或以上，描述组间测量方法的可比性。	
偏倚	9 描述用于减小潜在偏倚的一切措施。	
样本量	10 描述样本量是如何确定的。	
定量变量	11 解释在分析时如何处理定量变量。如需要，描述分组方法的选择及其原因。	
统计方法	12 ① 描述所有统计方法，包括控制混杂的方法；② 描述用于检测亚组和交互作用的所有方法；③ 描述如何处理缺失数据；④ 队列研究：如需要，解释如何处理失访。病例对照研究：如需要，解释如匹配设计病例和对照。横断面研究：如需要，描述基于抽样策略的分析方法；⑤ 描述所有敏感性分析	
数据获得和清洗方法	-	RECORD 12.1：作者应描述研究者从数据库人群中获取了多少来建立研究人群；RECORD 12.2：作者应提供研究所使用的数据清洗方法的有关信息。
链接	-	RECORD 12.3：说明研究是否在两个及以上数据库之间，使用了在个体水平、机构水平或其他数据进行数据链接。应提供链接的方法以及对其进行质量评价的方法。
结果		
研究对象	13 ① 报告研究每一阶段的研究个体数量（如潜在合格人数、参加合格性检查的人数、确认的合格人数、纳入研究的人数、完成随访的人数以及纳入分析的人数）；② 描述各阶段退出研究的原因；③ 考虑使用流程图	RECORD 13.1：详细描述研究纳入的个体选择方法（如研究人群选择），包括基线数据质量，数据可及性和链接等方面在内的筛选。可使用文本描述或研究流程图描述纳入个体选择方法。
描述性资料	14 ① 描述研究对象的特征（如人口学、临床和社会特征）、关于暴露和潜在混杂因素的信息；② 报告每个变量存在缺失数据的人数；③ 队列研究：汇总随访时间（如总随访时间和平均随访时间）	
结局资料	15 队列研究：报告结局事件数量或总结结局变量的测量值。病例-对照研究：报告每一类暴露的人数或总结暴露变量的测量值。横断面研究：报告结局事件数或总结结局变量的测量值。	

续表 1

	条目编码	STROBE 条目	RECORD 条目
主要结果	16	① 报告未调整的估计, 如需要, 报告调整混杂后的估计及其精确度 (如 95% 可信区间)。阐明对哪些混杂因素进行了调整, 以及选择这些混杂因素的原因; ② 将连续变量转化为分类变量时, 报告分类的界值; ③ 如需要, 考虑将相对危险度估计转化为一定时期内的绝对危险	
其他分析	17	报告其他分析, 如亚组分析、交互作用和敏感性分析。	
讨论			
关键结果	18	对照核心目标概括核心研究结果。	
局限性	19	讨论研究局限性, 应考虑潜在偏倚来源或不精确性。讨论潜在偏倚的方向和大小。	RECORD 19.1: 讨论将不是为回答特定问题而创建或收集的数据用于研究时的影响。包括对错分偏倚、未测量混杂、缺失数据以及随时间变化的适用性的讨论, 因为这些影响是此类研究所固有的。
解释	20	结合研究目标、局限性、多种分析方法、相似研究的结果和其他相关证据, 对结果进行谨慎、全面的解释。	
外推性	21	讨论研究结果的外推性 (外部真实性)。	
其他信息			
资助	22	描述本研究资助来源和资助者的角色, 如需要, 给出本篇论文所基于的原始研究的资助情况。	
方案、原始数据和程序编码的可获得性	-		RECORD 22.1: 作者需提供如何获得所有补充信息, 如研究方案、原始数据或程序编码的方法。

Heading, MeSH), 所以如何快速检索此类研究十分重要。然而, 考虑到各种各样的数据类型, 简单地说使用“常规数据”是不够的。应在题目和/或摘要中明确常规数据的类型。数据类型包括健康管理数据、其他管理数据 (如, 保险、出生/死亡登记和就业)、疾病登记、初级保健数据库、电子健康病历数据以及人口登记。命名所使用的数据库固然重要, 但不能替代在题目或摘要中提供数据源类型的重要性。

STROBE 清单已包括涉及地理区域和时间框架的项目。我们建议使用 RECORD 清单的论文也在题目或摘要部分将其作为必需条目。显然, 区域和时间框架的详细程度需符合字数限制并考虑保密问题。然而, 对研究区域的报告最低要求应使用定义研究人群 (如国家、州、省和地区) 的最高地理级别的名称。

此外, 应在题目或摘要中报告数据库之间的链接 (如果使用到)。例如, 可接受的措辞包括“使用多个链接的健康管理数据库”或“将某数据库 (名称) 与另一数据库 (名称) 链接”。在题目或摘要中使用“链接”一词可提供足够的信息, 关于链接方法的具体细节应在正文的方法学部分加以阐述。

2.2 前言

除 STROBE 中的条目外, 不需在 RECORD 规范中添加特定条目。STROBE 声明建议在“前言”

部分说明“具体目标, 包括任何预先设定的具体假设”。阐明具体的研究目标对于任何观察性研究的复制和翻译都是必要的。对于使用常规收集数据的研究, 作者应进一步澄清分析目的是出于数据探索并发现新关系 (如数据挖掘或旨在产生假说的研究^[21,22]) 的目的, 还是为了验证一个或更多的假设^[23]。作者也应该指出他们的假设是在数据分析之前还是之后形成的。同时, 应清楚地说明是否有研究方案, 如何获得, 以及是否在可公开获得的研究注册中心注册了该研究。因为使用常规收集数据开展研究的方法学优势和局限性可能还是有争议的, 所以清楚地描述研究目标是必要的^[23,24]。简单地将研究标记为描述性, 而没有澄清其目的是产生还是检验假设是不充分的。

2.3 方法

2.3.1 场景 RECORD没有额外扩展STROBE关于“描述数据收集的机构, 地点和时间范围, 包括征集研究对象、暴露、随访和数据收集的时间范围”的条目要求。应注意, 除在题目和/或摘要中已提及的数据库类型以外, 应提供相关信息以便读者理解数据库的内容及其真实性和收集数据的原始原因。例如, 记录电子病历的人可能是专科医师或初级保健医师, 也可能是急诊科医生或住院医师, 也可能是高年资医师或医学生。他们可能接受过专门培训, 以确保数据记录的完整和可重复性, 也可

能没有接受过培训^[25]。为让读者判断是否可以将研究发现应用到源人群中,作者还应该描述数据库人群与源人群的关系,包括选择标准。

2.3.2 研究对象 RECORD 6.1: 应详细列出研究人群选择方法(如用于识别研究对象的编码或算法)。如不可行,应给予解释。RECORD 6.2: 任何针对选择研究人群编码或算法进行的验证研究均应标引参考文献。如果为本研究实施的验证没有在其他地方发表,需要提供验证方法的细节和结果。RECORD 6.3: 如果研究涉及数据库链接,考虑使用流程图或其他图表展示以说明数据链接过程,包括每一步骤中能实现数据链接的个体人数。

实例 RECORD 6.1: 这部分报告较好的文章内容如下: OCCC (Ontario Crohn's 和 Colitis Cohort) 使用验证的算法基于年龄组来识别的 IBD 患者。这里的每个算法在安大略省应用于特定年龄组,多个队列,医疗实践类型和区域并已验证。对于 18 岁以下的儿童,根据他们是否进行结肠镜检查或乙状结肠镜检查来定义。如果他们进行了内镜检查,定义为在 3 年内需要 4 个门诊医师联系或 2 次因 IBD 住院。如他们未经内镜检查,则需要 7 个门诊医师联系或在 3 年内 3 次因 IBD 住院……该算法正确识别 IBD 儿童所具有灵敏度为……^[26]。本文参考了两个先前的算法验证研究来识别不同年龄的炎性肠病患者,包括诊断准确性的测量。

RECORD 6.2: ① Ducharme 和他的同事详细描述了识别儿童肠套叠的编码验证过程,然后使用已验证的编码描述流行病学。验证研究中涉及的列于文章的图 2^[27]。② Benchimol 和同事没有进行验证工作,然而,参考了先前进行的验证工作,详细描述了识别编码算法的诊断准确度信息^[26]。

RECORD 6.3: RECORD 网站上的示例图 2、图 3 和图 4 展示了一些可能的方式以说明链接过程: 图 2 说明链接过程的韦恩图(在 Herrett 等人的许可下复制^[28],见网站: <http://record-statement.org/images/figure2.jpg>)。图 3 混合流程图和韦恩图说明链接过程(转载得到 van Herk-Sukel 等人的许可^[29],见网站: <http://record-statement.org/images/figure3.jpg>)。图 4 与参与者流程图结合的链接图(在 Fosbøl 等人许可下^[30],见网站: <http://record-statement.org/images/figure4.jpg>)。

解释 RECORD 6.1 和 6.2: 报告用于获得研究人群的识别编码/算法的真实性对于使用常规收集卫生数据开展的观察性研究报告的透明性是必要的。此外,报告编码/算法可让其他研究者实施

外部或内部验证。

应明确和清晰地阐述用于识别研究对象的方法,包括是否基于唯一编码,算法(记录或编码的组合),数据库之间的链接,或自由文本字段。

正如许多其他流行病学研究一样,在使用常规卫生数据开展的研究中,错分偏倚的风险可能威胁研究结果的真实性^[31]。虽然在使用大样本人群数据库的研究中存在放大错分的风险,但这些研究提供了研究罕见疾病或非常见疾病的机会^[32]。作为使用常规收集卫生数据研究的必要部分,识别方法的验证日益受到重视,尤其是在使用以计费为目的收集的管理数据开展研究时使用的疾病编码^[33]。外部验证研究通常需要将用于识别研究人群的编码或算法与参考标准进行比较。最常用的参考标准是医疗记录,患者或医生的调查,以及临床登记^[5,34]。此外,数据库的内部验证可以通过比较单个数据库内交叉出现的数据源^[35]。准确性的测量类似于诊断试验研究的报告,包括灵敏度、特异度、阳性和阴性预测值,或 kappa 系数^[5,34]。

因此,对于使用常规收集卫生数据开展的观察性研究,我们建议在正文方法部分中提供详细的识别编码/算法的外部或内部验证信息。如果之前已开展过一项或多项验证研究,应标引参考文献。如未进行验证研究,则应明确说明。此外,应就识别方法的准确性(使用常见的诊断准确性术语)及它们在亚组人群中应用的功能进行简要讨论。如果作为观察性研究的一部分而开展的验证研究,我们建议作者使用已发表的验证研究的报告指南^[5]。重要的是要说明验证是否基于不同于本研究选择的数据来源或数据库人群开展的,因为编码可能在不同的人群或数据库中发挥不同的作用^[36]。另外,如果用于比较的参考标准存在已知问题,例如不完整或不准确,则应当报告这些问题并讨论其局限性。作者应讨论用于识别研究人群和结果所使用的选择编码/算法的含义、错分风险,及其对研究结果的潜在影响。对使用不同于当前研究人群的其他人群进行验证的意义进行讨论则更为重要。

RECORD 6.3: 流程图或其他图形展示可表达链接过程的相关信息,并可以简化冗长的描述。这样的图示可提供关键数据,如实现链接和未链接的个体所占比例和特征信息。读者据此可确定获得成功链接的数据库人群所占比例及研究人群的代表性。链接流程图可是独立的图(如,韦恩图或流程图),也可结合 STROBE 建议参与者流程图。图表展示可以有多种形式,我们不再给特定推荐。

2.3.3 变量 RECORD 7.1: 提供将暴露、结局、混杂因素和效应修饰因子进行分类的编码和算法的完整列表。如不能报告这些内容,需给予解释。

实例 ① Hardelid 和同事在他们的补充数据 2 的 S1 表中提供了所有编码^[37]。② Murray 及其同事在其附录 S1 中提供了有风险组的所有编码^[38]。

解释 正如用于识别研究群体的编码/算法一样,用于划分暴露、结局、混杂因子和效应修饰因子的编码/算法也会为研究带来潜在错分偏倚。为允许重复研究,对其进行评价并与其他研究进行比较,我们建议在正文、在线附录和/或外部网站上提供用于研究实施的所有诊断、程序、药物或其他编码的列表。对于通过调查方式收集的常规数据,针对研究对象的调查问题的措辞要清晰明确。考虑到所有研究(包括使用常规卫生数据开展的研究)均存在错分偏倚风险^[31],作者应提供足够的细节,使研究具有可重复性,并使偏倚风险易于识别。如可在文章正文中描述验证研究,或提供其他已出版或在线材料的参考文献。如上所述,作者应说明验证研究所使用的人群是否不同于当前研究所使用的人群来源或数据库。

我们认识到,在某些情况下,研究人员可能会避免在发表刊物中提供所使用的编码列表和算法,因该信息被视为专利或受版权、知识产权或其他法律的保护。例如,一些共病调整指数是由营利性公司创建,并销售给学术研究领域的研究者使用^[39,40]。在这些情况下,作者可能依赖数据提供者或受信任的第三方来收集、清洗和/或链接数据。作者应针对他们无法提供编码列表或其他用于识别研究个体或疾病信息的情况加以详细说明,并尽量提供这些列表所有权机构的联系信息。此外,作者应该阐明无法提供这些信息会如何影响其他研究者对研究的重复和评估。理想情况是,第三方应提供详细信息说明如何收集、清洗或链接数据。数据提供者 and 使用者之间的良好沟通可使双方获益。

一些学者认为编码列表代表了研究人员的知识产权。公开发表这些列表并允许其他研究人员将其应用到自己的研究中,可能会影响作者创建编码列表的知识产权和信用。我们认为这种观点不符合透明的科学标准,即允许研究的可重复。因此,除受法律或合同保护的编码列表,我们建议公开全部编码列表。

考虑到许多期刊的字数和页数限制以及编码列表/算法的长度,采用纸质期刊发表编码列表可行性较低。因而,可使用文本、表格、杂志网页版

将详细信息作为附录给予报告,也可由作者或其他个人永久在线存储,或者存放在第三方数据存储库(如, Dryad 或 Figshare)中。文章的正文和参考文献部分应提高如何获得编码列表的详细信息。如 ClinicalCodes.org 的编码仓库承诺公开基于健康数据开展研究中所使用的编码文档并透明化^[41]。如果将编码列表作为在线补充发表在期刊网站或作者提供的外部网站上,则应在期刊文章中公开链接。在杂志网站或 PubMed Central (<http://www.ncbi.nlm.nih.gov/pmc/>)上发表编码列表,会增加获得补充资料的可能性(只要杂志在发行)。如果在外部私人或机构网站上发表是唯一选择,建议至少在杂志文章发表后 10 年内,均可持续获得这些编码列表。如果更改了 URL 地址,则需要从原有网址自动的重新定位。这些措施将有助于未来的文章读者能够获得完整编码列表。

在文章中(或在线附件形式)除提供编码列表,作者还应提供分析指出研究所使用的编码/算法是否会引入偏倚。此类偏倚可能包括错分偏倚、确认偏倚以及缺失数据所致的偏倚。如果基于不同代码/算法集实施敏感性分析,则应对其进行描述和评估。潜在偏倚的讨论可结合 RECORD 和 STROBE 清单的其他部分,如研究对象选择和编码验证。

2.3.4 统计学方法 数据获得和清洗: RECORD 12.1: 作者应描述研究者从数据库人群中获取了多少来建立研究人群。RECORD 12.2: 作者应提供研究所使用的数据清洗方法的有关信息。RECORD 12.3: 说明研究是否在两个及以上数据库之间,使用了包含个体水平、机构水平或其他数据进行数据链接。应提供链接的方法以及对其进行质量评价的方法。

实例 RECORD 12.1: 以下文章描述对英国通用实践研究数据库(GPRD)子集的获取: ①“通过医学研究委员会许可协议, GPRD 将其数据集限制为通过其资助的项目的 100 000 名个体。该限制规定了使用病例对照而不是队列设计,以确保我们为每个特定症状确定足够的癌症病例……”^[42]。②“来自通用实践研究数据库的随机样本……经医学研究委员会学术机构许可获取^[43]”。

RECORD 12.2: 以下内容是数据清洗方法的描述示例^[44]: 用于链接的公共标识符的完整性在数据集和不同时间是不同的(近年来标识符更完整)。对于 LabBase2, 标识符的完整性有所不同(图 2)。对于 PICANet(儿科重症监护网),出生日期和医院

数量 100% 完整, 大多数其他标识符 >98% 完整, 除了 NHS (国家卫生服务) 号 (85% 完整)。对于这两个数据集, 进行数据清洗和数据准备: NHS 或医院编号例如 “未知” 或 “9999999999” 设置为空; 通用名称 (例如, “宝贝”、“双胞胎 1”、“婴儿”) 设置为空; 为多个姓创建了多个变量和名字; 开始 “ZZ” (表示没有英国邮政编码) 的邮政编码设置为空。

RECORD 12.3: 下面从文章摘录的是有关数据链接的级别, 所使用的链接技术和方法, 以及用于评估链接质量的方法的良好报告示例: ① “我们将活产和胎儿死亡证明链接到事件的时间链中, 排除人工流产和异位妊娠, 构成了个体女性的生殖经历”^[45]。② 两篇文章包含了对研究特定链接的详细描述报告^[44, 45]。在 Harron 及其同事^[44]的文章中, 有详细的解释提供了匹配过程的图形演示。另外, 计算链接概率的方法描述如下: “计算匹配概率 $P(M|协议模式)$ 以估计联合标识符集匹配的概率。这避免了标识符之间的独立性的假设。概率被导出为链接数除以每个协议模式对的总数 (基于在训练数据集中识别的可能链路)。例如, 如果有 378 个比较对出生日期和 Soundex 一致, 但性别不一致, 其中 312 个是可能的链接, 协议模式的匹配概率 $[1, 1, 0]$ 为 $312/378=0.825$ ”^[44]。Adams 和同事的文章也提供了解释链接的详细过程: “确定性链接包括 I 期, 其需要六个处理步骤, 在这六个处理步骤期间形成链并将单独的 (先前未链接的) 记录添加到链中。接下来是阶段 n, 其需要多次通过文件以组合属于同一母亲的链”^[45]。③ 相比之下, 如果一项研究涉及以前的链接数据, 则参考以前的文章即可, 如: “两个数据库的记录基于出生日期都与市政登记处联系在一起, 性别和邮政编码, 并随后相互链接。链接由荷兰统计局执行, 并在以前的出版物中描述”^[20]。④ 以下是良好报告链接和未链接个人特征的例子: “根据本文的目的, 不匹配的 ISC (住院统计数据收集) 记录将被称为 ISC 残差, 不匹配的 MDC (助产师数据收集) 记录作为 MDC 残差和链接对作为匹配记录……选择变量可用于两个数据集, 在三组间比较-ISC 残差、MDC 残差和匹配记录”^[46]。

解释 RECORD 12.1 和 12.2: 如果数据分析人员不熟悉队列建立的微小差别或建立研究队列的目标, 可能发生错误。因此, 应当报告作者获得数据库的范围。描述研究不同阶段的数据清理方法应包括用于筛选错误和缺失数据的方法, 包括范围

检查, 重复记录检查和重复测量的处理^[47, 48]。要报告的其他方法包括数据分布频率的评价、数据交叉表、图形探索或异常值检测所使用的统计方法^[49]。可提供关于错误诊断的其它细节, 包括合理性的定义和分析中的错误处理。清晰透明地描述数据清洗方法至关重要, 因为方法的选择可影响研究结果, 研究的可重复性和研究结果的再现^[50]。

RECORD 12.3: 对于链接研究, 我们建议报告成功链接的估计率, 确定性或概率性链接的使用, 用于链接的变量质量和类型, 以及所有链接验证的结果。如果数据库之间的记录链接是针对特定研究的, 则应报告链接方法及其质量评价, 包括由谁实施链接操作的信息。如果适用, 应提供以下细节: 区组变量、链接变量的完整性、关联规则、阈值和人工核查等^[44]。如在研究开始前已实施链接 (如, 前期研究或通用的链接), 或者由外部提供者 (如, 数据链接中心) 实施数据链接, 那么就需要通过参考文献描述数据资源和链接方法。

描述数据链接方法并评估其成功率是至关重要的, 可允许读者评估任何链接错误和有关偏倚产生的影响^[51]。具体地讲, 读者应知道使用的链接类型是确定性的还是概率性的, 以便确定链接是否可能受到错误匹配或缺失匹配的影响。当不同数据源之间存在唯一标识时, 可使用确定性链接。无法获得这种标识时, 描述应用的记录链接规则 (或统计链接主键) 便十分重要。相对而言, 概率链接会使用多个标识, 有时这些标识具有不同的权重, 并匹配高于特定阈值。可能也会用到混合方法。例如, 一些记录使用确定性链接, 当无法获得唯一标识时则可应用概率链接。当链接错误的概率 (如, 错误匹配和缺失匹配) 与结局变量存在关联时, 会发生链接偏倚。例如, 链接率会因患者特征, 如年龄、性别和健康状况而变化。即使链接过程中很小的错误也可能引入偏倚, 从而导致研究关联被高估或低估^[52]。作者应使用标准方法报告链接错误, 包括与金标准或参考数据集进行比较, 灵敏度分析, 以及链接和未链接数据特性的比较^[53]。报告链接错误允许读者确定链接的质量以及与链接错误有关的偏倚可能性。

2.4 结果 (研究对象)

RECORD 13.1: 详细描述研究纳入的个体选择方法 (如, 研究人群选择), 包括基于数据质量, 数据可及性和链接等方面的筛选。可使用文本描述或研究流程图描述纳入个体选择方法。

实例 以下所给出良好报告示例: 1998 年到

2007 年间,我们在 SEER (监测,流行病学和最终结果) 中根据诊断一处或多处肺癌和支气管癌病例,识别出 161 401 个医疗保险受益人。在这些患者中,我们总共识别出 163 379 例单独诊断事件性肺癌。(一些患者在研究期间间隔超过一年有两处原发性肺癌)。图 1 显示为最终队列 46 544 例患者共 46 935 例 NSCLC (非小细胞肺癌)^[54] 的来源。(关于示例流程图,请参见图 5,可从 <http://record-statement.org/images/figure5.jpg>) 获取。

解释 作者应详细描述由常规收集的卫生数据的原始数据获得研究人群过程,因为记录研究人群和数据库人群之间的差异可使研究结果得到更好的应用(参见 RECORD 6.1)。使用常规数据源的研究者经常基于可用数据的质量等因素来限制其研究人群。如,研究者可将研究周期限制到已知数据质量可接受的时间范围内,从而排除潜在参与者。研究可能会排除医疗实践与电子病历记录不符的记录,或者等到医疗实践与病历一致时再加入^[38,55]。研究人群也会由于数据可获得性而受到限制。如,在利用美国医疗保险数据库的研究中,近期在健康组织注册的受益人通常会因为记录缺少临床事件而被排除^[54,56]。当使用适用性随时间波动的数据源(如,保险数据库)时,研究人员需要明确说明如何定义适用性,以及如何在研究中管理适用性的更改。如果研究使用的是存在链接的常规数据,研究人群通常会减少或限制在那些可进行链接的个体中进行^[57]。特别严格的队列同样会因为方法学的原因减少一些混杂来源。

因此,获得最终研究人群的步骤、纳入排除标准和队列建立和分析过程中,不同阶段研究参与者的纳入和排除标准都应在正文中明确定义,可采用文字或使用合适的流程图。研究人群可使用不同的编码和/或算法(参见 RECORD 6.1)得到,并且随着时间变化使用不同的编码可能影响研究人群^[58,59]。一些研究也使用几种敏感/特异性的病例定义,这可能对随后的分析产生影响。这些步骤的描述在评估研究结果的外部有效性及在某些情况下评估可能的选择偏倚时至关重要。可报告敏感性分析以评估数据缺失和研究人群的代表性的潜在影响。提供关于从初始数据库选择研究人群的信息以便于研究的可重复。可对不同的研究人群进行辅助分析,并通过在线附录进行报告。

2.5 讨论(局限性)

RECORD 19.1: 讨论将不是为了回答特定问题而创建或收集的数据用于研究时的影响。包括对

错分偏倚、未测量混杂、缺失数据以及随时间变化的适用性的讨论,因这些影响是此类研究所固有的。

实例 以下文章描述使用管理数据库有关的局限性: ① “第三,这项研究是一个回顾性,基于索赔的分析。只有 PET (正电子发射断层扫描) 扫描支付的医疗保险可以在分析中检测到。为最小化错误的索赔比例,所有分析仅限于两者的医疗保险受益人医疗保险 A 部分和 B 部分覆盖,不纳入诊断前后 12 个月的注册管理医疗或医疗保险 C。第四,SEER 登记的患者更可能是非白人,生活在相对不贫困的地区,住在城市地区,这可能限制了研究发现的外推性。第五,在研究期间,疾病阶段是基于 SEER 数据在 4 个月内或直到第一次手术获得。2004 年,数据 SEER 的收集更改为协作分期系统。我们不清楚结果与更新后的方法有何差异”^[54]。② “尽管 SEER-Medicare 数据有一些优势,包括一个比较大的样本量,对美国人口的普及性,及详细的处方信息。我们的研究受限于缺乏关于胆固醇,甘油三酯和葡萄糖水平等提示人体的代谢紊乱程度的实验室数据……因此基于实验室的数据可以通过代谢疾病严重程度减少残留混杂。我们还缺乏更多关于癌症进展的粒度数据,这可能会影响他汀类药物使用和死亡之间的关联,因为他汀类药物治疗在预期生存时间较短的患者停止使用”^[60]。

解释 常规卫生数据通常不是针对特定的事先设计的研究问题而收集的,其数据收集的原因可能不同。许多潜在的偏倚,不但包括观察性研究有关的常见的偏倚来源,也包括很多使用常规收集数据的观察性研究特有的偏倚,这些都会影响研究结论。作者应将以下内容作为潜在的偏倚来源进行讨论: ① 识别研究人群、结局、混杂因素或效应修饰因子的编码或算法(错分偏倚); ② 缺失变量(未测量的混杂); ③ 缺失数据; ④ 随时间变化的适用性。

常规数据收集的原则可能影响研究问题所使用数据的质量和适用性。例如,用于回顾性分析的注册管理机构可能比收集其他类型的常规数据的组织实施更好的质量控制。同样,一些行政数据受到严格的质量控制,而其他数据则不然。管理数据尤其会出现不实或机会性编码的错误。例如,当医院报销基于病例组成的复杂性时,医院可能会通过自由地向患者记录中添加复杂疾病编码从而最大限度地报销偿付^[61]。此外,编码策略的变化可能影响数据的有效性或一致性。例如,提供商计费激励

码的引入可以改变随时间使用代码的可能性^[62, 63]。由于患者的污名或提供者的惩罚,可避免其他代码^[64]。此外,代码分类系统(如,从国际疾病分类 ICD-9 到 ICD-10)的版本的变化可能改变使用编码数据确认的有效性^[65, 66]。医院和人群间临床实践的差异可能会导致实验室检查在特定的地点进行,这可能会影响诊断算法。如果存在任何这些潜在偏倚来源,则应将其作为研究局限性进行讨论。

未测量的混杂被定义为在研究中未包含的与混杂相关的变量,导致残余混杂偏倚^[67]。虽然它是所有观察性研究中偏倚的潜在来源,但在使用常规收集的数据的研究中尤为突出。分析可能需要在设计数据库或收集数据时未考虑的变量。现已提出多种方法来解决这种潜在的偏倚来源^[68-71],包括倾向得分。然而,倾向评分分析,类似标准回归分析和匹配,只能保证研究参与者对数据中获得的变量上平衡。特定类型的未测量混杂为指示混杂,这通常是在使用常规收集数据进行(药物)治疗有效性和安全性评价时遇到。那些接受这种(药物)治疗的患者相比不使用治疗的患者,预后更好或更差,但有关预后或疾病严重程度的信息无法从数据中获得^[72]。作者应讨论这类问题,如果存在,应报告处理这类问题所采用的方法。

缺失数据是所有观察性研究都会遇到的问题,并已在 STROBE 解释性文章^[10]的框 6 中提到。缺失数据是常规收集数据的一个特殊问题,因为研究人员无法控制数据收集^[73]。如果在定义研究队列需要用到的变量上有缺失数据或用于链接识别变量有缺失从而使记录无法链接,特别是非随机缺失。作者应描述可能会造成未测量混杂的缺失变量,变量缺失的原因,对研究结果的影响,以及调整缺失变量所用方法。如,吸烟状态对克罗恩病的严重程度具有强烈影响,并且与该疾病的结局相关。然而,卫生管理数据中很少包含吸烟状况的数据。在评估社会经济地位与克罗恩病结局之间关系的研究中,吸烟状态被讨论为潜在的未测量的混杂因素^[74]。通常,仅在使用常规卫生数据开始研究之后才发现缺失的数据/缺失的变量,使得研究者有必要偏离其原始研究方案。不管何种原因造成偏离,都应报告偏离方案的细节。应讨论偏离的原因及其对研究和结论的影响。

另一个重要的潜在局限性是由于数据库人群、研究人群或两者的组成随时间的变化而导致的编码实践或合格标准的变化。数据库人群的定义可在多种情况下改变,例如,如果注册实践停止与数

据库的协作,改变计算机软件或在数据库变化时改变注册的标准,如注册表。如果由于就业,住所或医疗服务提供者的变化,个体的适用性会随时间变化,行政数据来源(例如保险数据库)的研究人群可能会改变。记录编码的改变(不实或如前所述的编码系统改变)可能会改变研究人群^[63, 75, 76]。当讨论局限性时,作者应该解释如何在分析中处理变更的适用性,以便读者可评估偏倚的可能性。正如 STROBE 声明提到的,讨论需要报告所有潜在偏倚的方向和大小以及处理方法。

2.6 其他信息

RECORD 22.1: 作者需提供如何获得所有补充信息,如研究方案、原始数据或程序编码的方法。

实例 ① Taljaard 及其同事^[77]的文章代表了使用加拿大社区健康调查研究完整的研究方案。② 在他们的文章中, Guttman 和同事根据研究方案的要求:“数据共享:技术附录,数据集创建计划/方案和统计编码可从相应的作者(电子邮件地址)获取”^[78]。

解释 我们强烈推荐提供详细的研究方法和结果的信息。如果可能,我们建议提前或同时发表研究方案,原始数据结果和程序代码(如允许)。此信息对同行评审者和读者评估研究结果的有效性十分有用。研究者有很多可用的途径公开发表这些数据。这些途径包括在线期刊补充材料、个人网站、机构网站、基于科学的社交媒体网站(如, ResearchGate.net 和 Academia.edu)、数据存储库(如, Dryad 或 Figshare)或政府开放数据网站^[79]。我们发现,一些研究组织、公司、教育机构或法律可能禁止或限制此类信息的免费提供。虽然对此知识产权的所有权和使用的讨论超出了 RECORD 指南的范围,这些数据的发布通常要符合研究者研究机构法律和伦理指南以及期刊编辑的指南。该信息对于可能希望访问这些数据以复制,再现或扩展正文描述的研究的其他研究人员也是有用的。无论这些可用的补充信息的格式或程度如何,我们建议在正文中明确说明这些信息的位置。

3 讨论

RECORD 规范针对于使用常规收集的卫生数据进行观察性研究,是 STROBE 声明的补充而非替代。RECORD 规范为作者、期刊编辑、同行评议者和其他利益相关者制定,旨在加强使用常规收集的卫生数据进行研究的报告的透明度和完整性。该清单可供使用此类数据的任何研究人员使用,我

们鼓励向所有相关方广泛传播。我们期望期刊对 RECORD 规范的认可和应用,以改善使用常规收集卫生数据实施的研究报告的透明性。

3.1 局限性

STROBE 声明和 RECORD 规范仅用于观察性研究。然而,常规收集的卫生数据有时也用于其他研究设计的研究,例如用于卫生系统评价的群组随机试验。另外,随机试验来源的数据链接与管理数据可用于结局的长期随访,以及关联研究并不属于观察性研究。随着领域的发展,我们期望使用类似的方法将 RECORD 规范扩展到其他研究设计。

尽管 RECORD 规范代表反映利益相关者的兴趣和优先事项,但使用常规收集的卫生数据进行研究的方法正在迅速变化,并且此类研究的数据类型的可用性正在扩大。例如,移动健康应用程序(mHealth 应用程序)正广泛用于智能手机和可穿戴技术。虽然目前仅有有限的研究使用这类数据资源,但我们预计在不久的将来这些数据的使用将快速增长,并将创建管理这种资源的新方法。此外,工作委员会关注于卫生数据,而非健康相关研究所使用到的所有数据资源(如,环境数据和经济数据等)。因此,RECORD 规范可能不能很好反映将来可能会变为重要的内容,并且在某些方面可能需要修订。

在这些规范制定的过程中已认真考虑利益相关者的广泛代表性。我们通过开放电话和使用各种渠道有针对性地邀请招募利益相关者^[16]。然而,利益相关方代表主要来自使用常规收集的卫生数据进行研究的地区,只有几名代表来自发展中国家和非英语国家。我们认为利益相关方代表了目前该领域的研究者和使用者的普遍认识。虽然通过调查和利益相关者团体获得大量反馈,但鉴于可行性考虑,该声明是由 19 名成员组成的较小的工作委员会制定的,正如已有研究^[17]中建议的那样。在未来,技术和社交媒体可能允许更大的团体更积极地参与工作委员会会议。

3.2 未来方向和社区参与

随着常规收集的健康数据的可用性扩大,我们希望来自当前无法获得这些数据的地区的研究人员能更多地参与其中。通过 recordstatement.org 网站和留言板,我们期望持续对 RECORD 规范相关方面进行评论和讨论,这可能促成未来的正式修订。通过这个在线社区,RECORD 将成为一个灵活的文档,以适应不同领域的变化。

公开发表和通过期刊推荐对改善研究报告质

量来说还不足够^[80]。研究者、期刊和同行评审员实施 RECORD 规范的方式是其重要的可测量影响^[81]。因此,在线留言板包含对其应用的讨论论坛。我们还鼓励评价 RECORD 对该领域报告质量影响的评价,从而可使该规范提供可衡量的益处。

4 结论

RECORD 规范是 STROBE 声明在使用常规收集数据实施的观察性研究中的扩展。根据研究和发表的意见,我们以清单的形式编写了规范以及配套的解释性文件。已有研究证明报告指南改进了研究报告质量,从而使读者能够了解结论的优势、局限性和准确性^[12, 82-84]。同时,我们期望 RECORD 规范将随着该领域研究方法的发展而改变,该指南将在未来几年内有助于促进对研究进行充分报告。随着作者、期刊编辑和同行评议人员的使用,我们期望 RECORD 规范可切实改善使用常规收集卫生数据实施的研究报告的透明性、可重复性和完整性。

支持信息 参与本调查的利益相关者名单(见原文表 S1)。

RECORD 规范的制定得到加拿大卫生研究所(基金编号: 130512)、瑞典国家科学基金会(基金编号: IZ32Z0_147388/1)和奥胡斯大学临床流行病学系的资助。资助者并未参与研究设计、数据收集和分析、撰写和发表过程。该工作由东安大略儿童医院研究伦理委员会批准。RECORD 规范的全体作者感谢在有关所需纳入清单内容调查过程中所有参与的利益相关者(见原文表 S1)。同时也十分感谢 STROBE 声明发起组织成员在 RECORD 制定过程中给予的帮助。感谢 Pauline Quach 和 Danielle Birman 所作的协调工作,以及 record-statement.org 网站设计和管理者 Andrew Perlmutter。感谢所有参与推动 RECORD 规范发展的利益相关者。

RECORD 工作委员会成员名单 Douglas Altman (Centre for Statistics in Medicine, Oxford University), Nicholas de Klerk (University of Western Australia), Lars G. Hemkens (University Hospital Basel), David Henry (University of Toronto and Institute for Clinical Evaluative Sciences, Toronto), Jean-Marie Januel (University of Lausanne), Marie-Annick Le Pogam (Institute of Social and Preventive Medicine, University Hospital of Lausanne), Douglas Manuel (Ottawa Hospital Research Institute, University of Ottawa), Kirsten

Patrick (editor, Canadian Medical Association Journal[CMAJ]), Pablo Perel (London School of Hygiene and Tropical Medicine), Patrick S. Romano (University of California, Davis, co-editor-in-chief, Health Services Research), Peter Tugwell (University of Ottawa, editor-in-chief, Journal of Clinical Epidemiology), Joan Warren (National Institutes of Health/National Cancer Institute), Wim Weber (European editor, BMJ), Margaret Winker (formerly senior research editor, PLOS Medicine; current secretary, World Association of Medical Editors)

参考文献

- Spasoff RA. Epidemiologic Methods for Health Policy. New York: Oxford University Press, 1999.
- Morrato EH, Elias M, Gericke CA. Using population-based routine data for evidence-based health policy decisions: lessons from three examples of setting and evaluating national health policy in Australia, the UK and the USA. *J Public Health*, 2007, 29(4): 463-471.
- De Coster C, Quan H, Finlayson A, *et al.* Identifying priorities in methodological research using ICD-9-CM and ICD-10 administrative data: report from an international consortium. *BMC Health Serv Res*, 2006, 6: 77.
- Hemkens LG, Benchimol EI, Langan SM, *et al.* Reporting of studies using routinely collected health data: systematic literature analysis (oral abstract presentation). Edinburgh, UK: Reward/equator conference, 2015.
- Benchimol EI, Manuel DG, To T, *et al.* Development and use of reporting guidelines for assessing the quality of validation studies of health administrative data. *J Clin Epidemiol*, 2011, 64(8): 821-829.
- Herrett E, Thomas SL, Schoonen WM, *et al.* Validation and validity of diagnoses in the General Practice Research Database: a systematic review. *Br J Clin Pharmacol*, 2010, 69(1): 4-14.
- Rothman KJ, Greenland S, Lash TL. Modern Epidemiology, 3rd edition. Philadelphia: Lippincott Williams & Wilkins, 2008.
- 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(5): 263-267.
- Enhancing the quality and transparency of health research (EQUATOR) network library 2015. Available at: www.equator-network.org/library/.
- Vandenbroucke JP, von Elm E, Altman DG, *et al.* Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLOS Med*, 2007, 4(10): e297.
- von Elm E, Altman DG, Egger M, *et al.* The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLOS Med*, 2007, 4(10): e296.
- Sorensen AA, Wojahn RD, Manske MC, *et al.* Using the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement to assess reporting of observational trials in hand surgery. *J Hand Surg Am*, 2013, 38(8): 1584-1589.
- Cobo E, Cortes J, Ribera JM, *et al.* Effect of using reporting guidelines during peer review on quality of final manuscripts submitted to a biomedical journal: masked randomised trial. *BMJ*, 2011, 343: d6783.
- Benchimol EI, Langan S, Guttman A. Call to RECORD: the need for complete reporting of research using routinely collected health data. *J Clin Epidemiol*, 2013, 66(7): 703-705.
- Langan SM, Benchimol EI, Guttman A, *et al.* Setting the RECORD straight: developing a guideline for the REporting of studies Conducted using Observational Routinely collected Data. *Clin Epidemiol*, 2013, 5: 29-31.
- Nicholls SG, Quach P, von Elm E, *et al.* The REporting of Studies Conducted Using Observational Routinely-Collected Health Data (RECORD) Statement: Methods for Arriving at Consensus and Developing Reporting Guidelines. *PLOS One*, 2015, 10(5): e0125620.
- Moher D, Schulz KF, Simera I, *et al.* Guidance for developers of health research reporting guidelines. *PLOS Med*, 2010, 7(2): e1000217.
- Glasziou P, Altman DG, Bossuyt P, *et al.* Reducing waste from incomplete or unusable reports of biomedical research. *Lancet*, 2014, 383(9913): 267-276.
- Blotiere PO, Weill A, Ricordeau P, *et al.* Perforations and haemorrhages after colonoscopy in 2010: a study based on comprehensive French health insurance data (SNIIRAM). *Clin Res Hepatol Gastroenterol*, 2014, 38(1): 112-117.
- Siregar S, Pouw ME, Moons KG, *et al.* The Dutch hospital standardised mortality ratio (HSMR) method and cardiac surgery: benchmarking in a national cohort using hospital administration data versus a clinical database. *Heart*, 2014, 100(9): 702-710.
- Price SD, Holman CD, Sanfilippo FM, *et al.* Use of case-time-control design in pharmacovigilance applications: exploration with high-risk medications and unplanned hospital admissions in the Western Australian elderly. *Pharmacoepidemiol Drug Saf*, 2013, 22(11): 1159-1170.
- Gross CP, Andersen MS, Krumholz HM, *et al.* Relation between Medicare screening reimbursement and stage at diagnosis for older patients with colon cancer. *JAMA*, 2006, 296(23): 2815-2822.
- Vandenbroucke JP. Observational research, randomised trials, and two views of medical science. *PLOS Med*, 2008, 5(3): e67.
- Smith GD, Ebrahim S. Data dredging, bias, or confounding. *BMJ*, 2002, 325(7378): 1437-1438.
- Prokosch HU, Ganslandt T. Perspectives for medical informatics. Reusing the electronic medical record for clinical research. *Methods Inf Med*, 2009, 48(1): 38-44.
- Benchimol EI, Manuel DG, Guttman A, *et al.* Changing age demographics of inflammatory bowel disease in ontario, canada: a population-based cohort study of epidemiology trends. *Inflamm Bowel Dis*, 2014, 20(10): 1761-1769.
- Ducharme R, Benchimol EI, Deeks SL, *et al.* Validation of diagnostic codes for intussusception and quantification of childhood intussusception incidence in ontario, Canada: a population-based study. *J Pediatr*, 2013, 163(4): 1073-1079.

- 28 Herrett E, Shah AD, Boggon R, *et al.* Completeness and diagnostic validity of recording acute myocardial infarction events in primary care, hospital care, disease registry, and national mortality records: cohort study. *BMJ*, 2013, 346: f2350.
- 29 van Herk-Sukel MP, van de Poll-Franse LV, Lemmens VE, *et al.* New opportunities for drug outcomes research in cancer patients: the linkage of the Eindhoven Cancer Registry and the PHARMO Record Linkage System. *Eur J Cancer*, 2010, 46(2): 395-404.
- 30 Fosbol EL, Granger CB, Peterson ED, *et al.* Prehospital system delay in STsegment elevation myocardial infarction care: a novel linkage of emergency medicine services and in hospital registry data. *Am Heart J*, 2013, 165(3): 363-370.
- 31 Manuel DG, Rosella LC, Stukel TA. Importance of accurately identifying disease in studies using electronic health records. *BMJ*, 2010, 341: c4226.
- 32 Narcolepsy in association with pandemic influenza vaccination (a multi-country European epidemiological investigation). Stockholm: ECDC, 2012.
- 33 Lewis JD, Schinnar R, Bilker WB, *et al.* Validation studies of the health improvement network (THIN) database for pharmacoepidemiology research. *Pharmacoepidemiol Drug Saf*, 2007, 16(4): 393-401.
- 34 Sorensen HT, Sabroe S, Olsen J. A framework for evaluation of secondary data sources for epidemiological research. *Int J Epidemiol*, 1996, 25(2): 435-442.
- 35 Baron JA, Lu-Yao G, Barrett J, *et al.* Internal validation of Medicare claims data. *Epidemiology*, 1994, 5(5): 541-544.
- 36 Marston L, Carpenter JR, Walters KR, *et al.* Smoker, ex-smoker or non-smoker? The validity of routinely recorded smoking status in UK primary care: a cross-sectional study. *BMJ open*, 2014, 4(4): e004958.
- 37 Hardelid P, Dattani N, Gilbert R. Estimating the prevalence of chronic conditions in children who die in England, Scotland and Wales: a data linkage cohort study. *BMJ Open*, 2014, 4(8): e005331.
- 38 Murray J, Bottle A, Sharland M, *et al.* Risk factors for hospital admission with RSV bronchiolitis in England: a population-based birth cohort study. *PLOS One*, 2014, 9(2): e89186.
- 39 Berry JG, Hall M, Hall DE, *et al.* Inpatient growth and resource use in 28 children's hospitals: a longitudinal, multi-institutional study. *JAMA pediatrics*, 2013, 167(2): 170-177.
- 40 Shahian DM, Wolf RE, Iezzoni LI, *et al.* Variability in the measurement of hospital-wide mortality rates. *N Engl J Med*, 2010, 363(26): 2530-2539.
- 41 Springate DA, Kontopantelis E, Ashcroft DM, *et al.* ClinicalCodes: an online clinical codes repository to improve the validity and reproducibility of research using electronicmedical records. *PLOS One*, 2014, 9(6): e99825.
- 42 Dommett RM, Redaniel MT, Stevens MC, *et al.* Features of childhood cancer in primary care: a population-based nested case-control study. *Br J Cancer*, 2012, 106(5): 982-987.
- 43 Tsang C, Bottle A, Majeed A, *et al.* Adverse events recorded in English primary care: observational study using the General Practice Research Database. *Br J Gen Pract*, 2013, 63(613): e534-42.
- 44 Harron K, Goldstein H, Wade A, *et al.* Linkage, evaluation and analysis of national electronic healthcare data: application to providing enhanced blood-stream infection surveillance in paediatric intensive care. *PLOS One*, 2013, 8(12): e85278.
- 45 Adams MM, Wilson HG, Casto DL, *et al.* Constructing reproductive histories by linking vital records. *Am J Epidemiol*, 1997, 145(4): 339-348.
- 46 Ford JB, Roberts CL, Taylor LK. Characteristics of unmatched maternal and baby records in linked birth records and hospital discharge data. *Paediatr Perinat Epidemiol*, 2006, 20(4): 329-337.
- 47 Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. *J Am Med Inform Assoc*, 2013, 20(1): 144-151.
- 48 Sandall J, Murrells T, Dodwell M, *et al.* The efficient use of the maternity workforce and the implications for safety and quality in maternity care: a population-based, cross-sectional study. *Health Serv Deliv Res*, 2014, 2: 38. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK260224/>.
- 49 Welch C, Petersen I, Walters K, *et al.* Two-stage method to remove population- and individual-level outliers from longitudinal data in a primary care database. *Pharmacoepidemiol Drug Saf*, 2012, 21(7): 725-732.
- 50 Van den Broeck J, Cunningham SA, Eeckels R, *et al.* Data cleaning: detecting, diagnosing, and editing data abnormalities. *PLOS Med*, 2005, 2(10): e267.
- 51 Bohensky MA, Jolley D, Sundararajan V, *et al.* Data linkage: a powerful research tool with potential problems. *BMC Health Serv Res*, 2010, 10: 346.
- 52 Harron K, Wade A, Muller-Pebody B, *et al.* Opening the black box of record linkage. *J Epidemiol Community Health*, 2012, 66(12): 1198.
- 53 Lariscy JT. Differential record linkage by Hispanic ethnicity and age in linked mortality studies: implications for the epidemiologic paradox. *J Aging Health*, 2011, 23(8): 1263-1284.
- 54 Dinan MA, Curtis LH, Carpenter WR, *et al.* Variations in use of PET among Medicare beneficiaries with non-small cell lung cancer, 1998-2007. *Radiology*, 2013, 267(3): 807-817.
- 55 Horsfall L, Walters K, Petersen I. Identifying periods of acceptable computer usage in primary care research databases. *Pharmacoepidemiol Drug Saf*, 2013, 22(1): 64-69.
- 56 Gerber DE, Laccetti AL, Xuan L, *et al.* Impact of prior cancer on eligibility for lung cancer clinical trials. *J Natl Cancer Inst*, 2014, 106(11): dju302.
- 57 Carrara G, Scire CA, Zambon A, *et al.* A validation study of a new classification algorithm to identify rheumatoid arthritis using administrative health databases: case-control and cohort diagnostic accuracy studies. Results from the record linkage On Rheumatic Diseases study of the Italian Society for Rheumatology. *BMJ Open*, 2015, 5(1): e006029.
- 58 Rait G, Walters K, Griffin M, *et al.* Recent trends in the incidence of recorded depression in primary care. *Br J Psychiatry*, 2009, 195(6): 520-524.
- 59 Wijlaars LP, Nazareth I, Petersen I. Trends in depression and antidepressant prescribing in children and adolescents: a cohort study in The Health Improvement Network (THIN). *PLOS One*, 2012, 7(3): e33181.
- 60 Jeon CY, Pandolfi SJ, Wu B, *et al.* The Association of Statin Use after Cancer Diagnosis with Survival in Pancreatic Cancer Patients: A SEER-Medicare Analysis. *PLOS One*, 2015, 10(4): e0121783.
- 61 Pruitt Z, Pracht E. Upcoding emergency admissions for non-life-threatening injuries to children. *Am J Manag Care*, 2013, 19(11):

- 917-924.
- 62 McIntock K, Russell AM, Alderson SL, *et al.* The effects of financial incentives for case finding for depression in patients with diabetes and coronary heart disease: interrupted time series analysis. *BMJ Open*, 2014, 4(8): e005178.
- 63 Brunt CS. CPT fee differentials and visit upcoding under Medicare Part B. *Health Economics*, 2011, 20(7): 831-841.
- 64 Walters K, Rait G, Griffin M, *et al.* Recent trends in the incidence of anxiety diagnoses and symptoms in primary care. *PLOS One*, 2012, 7(8): e41670.
- 65 Nilson F, Bonander C, Andersson R. The effect of the transition from the ninth to the tenth revision of the International Classification of Diseases on external cause registration of injury morbidity in Sweden. *Inj Prev*, 2015, 21(3): 189-194.
- 66 Jagai JS, Smith GS, Schmid JE, *et al.* Trends in gastroenteritis-associated mortality in the United States, 1985 inverted question mark 2005: variations by ICD-9 and ICD-10 codes. *BMC Gastroenterol*, 2014, 14(1): 211.
- 67 European Network of Centres for Pharmacoepidemiology and Pharmacovigilance Guide on Methodological Standards in Pharmacoepidemiology, 4.2.2.5. London: European Medicines Agency, 2015. Available at: www.encepp.eu/standards_and_guidances/methodologicalGuide4_2_2_5.shtml.
- 68 Toh S, Garcia RLA, Hernan MA. Confounding adjustment via a semi-automated high-dimensional propensity score algorithm: an application to electronic medical records. *Pharmacoepidemiol Drug Saf*, 2011, 20(8): 849-857.
- 69 Stukel TA, Fisher ES, Wennberg DE, *et al.* Analysis of observational studies in the presence of treatment selection bias: effects of invasive cardiac management on AMI survival using propensity score and instrumental variable methods. *JAMA*, 2007, 297(3): 278-285.
- 70 Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behav Res*, 2011, 46(3): 399-424.
- 71 Sterne JA, White IR, Carlin JB, *et al.* Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*, 2009, 338: b2393.
- 72 Freemantle N, Marston L, Walters K, *et al.* Making inferences on treatment effects from real world data: propensity scores, confounding by indication, and other perils for the unwary in observational research. *BMJ*, 2013, 347: f6409.
- 73 Marston L, Carpenter JR, Walters KR, *et al.* Issues in multiple imputation of missing data for large general practice clinical databases. *Pharmacoepidemiol Drug Saf*, 2010, 19(6): 618-626.
- 74 Benchimol EI, To T, Griffiths AM, *et al.* Outcomes of pediatric inflammatory bowel disease: socioeconomic status disparity in a universal-access healthcare system. *J Pediatr*, 2011, 158(6): 960-967.
- 75 Nassar N, Dixon G, Bourke J, *et al.* Autism spectrum disorders in young children: effect of changes in diagnostic practices. *Int J Epidemiol*, 2009, 38(5): 1245-1254.
- 76 Tan GH, Bhoo-Pathy N, Taib NA, *et al.* The Will Rogers phenomenon in the staging of breast cancer-does it matter? *Cancer Epidemiol*, 2015, 39(1): 115-117.
- 77 Taljaard M, Tuna M, Bennett C, *et al.* Cardiovascular Disease Population Risk Tool (CVDPeRT): predictive algorithm for assessing CVD risk in the community setting. A study protocol. *BMJ Open*, 2014, 4(10): e006701.
- 78 Guttmann A, Schull MJ, Vermeulen MJ, *et al.* Association between waiting times and short term mortality and hospital admission after departure from emergency department: population based cohort study from Ontario, Canada. *BMJ*, 2011, 342: d2983.
- 79 Nicol A, Caruso J, Archambault E. Open Data Access Policies and Strategies in the European Research Area and Beyond. Montreal: Science-Metrix Inc, 2013.
- 80 Fuller T, Pearson M, Peters J, *et al.* What affects authors' and editors' use of reporting guidelines? Findings from an online survey and qualitative interviews. *PLOS One*, 2015, 10(4): e0121585.
- 81 Turner L, Shamseer L, Altman DG, *et al.* Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. *Cochrane Database Syst Rev*, 2012, 11: MR000030.
- 82 Armstrong R, Waters E, Moore L, *et al.* Improving the reporting of public health intervention research: advancing TREND and CONSORT. *J Public Health (Oxf)*, 2008, 30(1): 103-109.
- 83 Moher D, Cook DJ, Eastwood S, *et al.* Improving the quality of reports of metaanalyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet*, 1999, 354(9193): 1896-1900.
- 84 Prady SL, Richmond SJ, Morton VM, *et al.* A systematic evaluation of the impact of STRICTA and CONSORT recommendations on quality of reporting for acupuncture trials. *PLOS One*, 2008, 3(2): e1577.

收稿日期: 2017-02-06 修回日期: 2017-02-28

本文编辑: 熊鹰、张永刚