

如何撰写高质量的流行病学研究论文

第三讲 肿瘤标志物预后研究报告规范——REMARK 介绍

邓秋菊 陈茹 詹思延

【导读】 尽管肿瘤标志物的研究已历经多年且产出颇丰,但真正能够应用到临床者却很少。其中不规范、不全面的报告严重影响了研究质量及结果评估,在一定程度上制约了肿瘤标志物的发展及应用。本文旨在介绍肿瘤标志物预后研究报告规范(Reporting Recommendations for Tumor Marker Prognostic Studies,REMARK)的清单内容,并对一些重要条目进行说明及解读。

【关键词】 肿瘤标志物预后研究报告规范;肿瘤标志物;清单

How to write high-quality epidemiological research paper III. Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK) —introduction and explanation DENG Qiu-ju¹, CHEN Ru², ZHAN Si-yan². 1 Laboratory of Genetics, Peking University Cancer Hospital and Institute, Beijing 100142, China; 2 Department of Epidemiology and Biostatistics, School of Public Health, Peking University Health Science Center

Corresponding author:ZHAN Si-yan, Email:siyan-zhan@bjmu.edu.cn

This work was supported by a grant from the Beijing Natural Science Foundation (No. 7111006).

【Key words】 Reporting Recommendations for Tumor Marker Prognostic Studies; Tumor marker; List

1. REMARK 的产生背景:肿瘤标志物是由肿瘤细胞合成、释放或由宿主对肿瘤进行反应而产生的一类物质。根据其功能可分为诊断标志物、预后标志物、预测标志物及监测标志物。肿瘤标志物检测对肿瘤的早期诊断及判断疗效、转移、复发和预后等具有重要意义。尽管对肿瘤标志物的研究已历经数年且文献报道颇多,但像甲胎蛋白、癌胚抗原等肿瘤标志物一样能真正应用到临床者却很少。通常最初的研究结果总是提示某肿瘤标志物有重大临床意义,但后续研究却得到不一致甚至完全相反的结论^[1]。因此当务之急是探索针对同一肿瘤标志物的不同研究得到不一致结果的深层原因。遗憾的是,大部分肿瘤标志物研究并没有按照严格的规范进行报告,发表的文献也缺乏足够信息,以致于读者不能充分评估其研究质量及结果。为此肿瘤标志物预后研究报告规范(Reporting Recommendations for Tumor Marker Prognostic Studies,REMARK)应运而生,为规范相关研究的报告格式、提高报告质量提供了参考^[1]。

2. REMARK 的制定、目标及使用:REMARK 最

初由美国国立癌症研究所(NCI)/欧洲癌症研究与治疗组织(EORTC)于 2000 年在第一届国际肿瘤诊断会议上提出。鉴于 CONSORT^[2]对随机对照试验及 STARD^[3]对诊断试验的成功应用,REMARK 指南在研究设计、预先假设、患者与标本特征、检测方法及统计分析方法上均提供了相关信息,并对如何展示数据及重要部分的讨论提供了有用的建议。2012 年 REMARK 推出了解释与说明文件^[4],针对清单条目逐一进行了举例和阐释,为进一步理解及执行 REMARK 声明提供了指导。

REMARK 旨在鼓励透明及完整的报告,使其他人能够利用文献中相关信息判断研究数据的可信度及结论的可推广性,并不规定其特有的研究设计或分析策略。目前该指南已在多家肿瘤专业期刊上公开发表并被推荐使用。

3. REMARK 内容:REMARK 清单(表 1)共包含 20 个条目,即前言(条目 1)、材料与方法(条目 2~11)、结果(条目 12~18)、讨论(条目 19~20)。本文对 REMARK 的清单内容进行解读。

4. REMARK 清单解读:

条目 1 研究目的包括评估肿瘤标志物与临床结局的关联,以及确定肿瘤标志物是否比传统临床或病理提供更多的临床结局相关信息。标志物的描述既包括标志物的生物学特性,也包括其在临床过

DOI: 10.3760/cma.j.issn.0254-6450.2013.09.017

基金项目:北京市自然科学基金(7111006)

作者单位:100142 北京大学肿瘤医院北京市肿瘤防治研究所遗传室(邓秋菊);北京大学公共卫生学院流行病与卫生统计学系(陈茹、詹思延)

通信作者:詹思延, Email:siyan-zhan@bjmu.edu.cn

表 1 肿瘤标志物预测研究的推荐报告清单

条目	条目序号	解 释
前言	1	说明检测的标志物、研究目的及任何预先特定的假设
材料与 方法	2	描述研究对象的特征(例如疾病阶段或合并症),包括其来源以及纳入和排除标准
患者	3	描述接受治疗及其如何选择(例如随机或基于某种规则)
样本特征	4	描述使用的生物材料类型(包括对照样本)及其保存储藏方法
检测方法	5	明确所使用的试验方法,并提供(或引用)详细方案,包括使用的特定试剂或试剂盒、质量控制程序、可重复性评估、定量分析方法、得分及报告方案;明确是否以及如何对研究终点盲法检测
研究设计	6	说明研究对象选择的方法,包括是否为前瞻性或回顾性、是否使用了分层或匹配(例如按疾病阶段或年龄);明确病例采集的时间段、随访终点和中位随访时间
	7	精确定义所有检测的临床终点
	8	列出所有初步检测或考虑纳入模型的变量
	9	给出样本量确定的理由;如果研究设计是为了检测一个特定的效应值,还应给出目标效能和效应值
统计分析方法	10	明确所有的统计方法,包括变量选择详细过程和其他的建模问题、如何验证模型假设以及如何处理缺失数据
	11	阐明分析中如何处理标志物数值;如有可能,应描述用于决定界值方法
结果		
数据	12	清晰描述整个研究过程中病例数量及其变化,包括每个分析阶段纳入和排除的病例数量(如设计一个流程图则更好)及其原因,即无论是总研究人群还是各亚组纳入人群均应明确报告事件发生的病例数量
	13	报告各变量分布情况及其缺失值,包括基本人口学特征(至少年龄和性别)、传统的(疾病相关)预后变量及肿瘤标志物
分析和展示	14	报告标志物与传统预后变量间的关联
	15	报告单因素分析结果,包括肿瘤标志物与结局间的关联以及相应估计值[例如风险比(HR值)及生存概率],最好同时报告其他所有变量的类似分析。研究肿瘤标志物对时间-事件结局的效应时,建议用Kaplan-Meier法进行分析
	16	对于重要的多因素分析,或者至少在最终模型中应报告肿瘤标志物及其他所有变量的估计值(如HR值)及其可信区间(CI)
	17	无论是否有统计学意义,在结果中均应报告标志物及传统预后变量的估计效应及其CI
	18	如果有其他结果,如假设检验、敏感性分析及内部验证应予以报告
讨论		
	19	结合先前特定假设及其他相关研究对结果做出解释,包括对研究的局限性进行讨论
	20	讨论本文对未来研究的意义及临床价值

程中的分析时间。生物学特性包括检测的分子或结构类型(如蛋白质、RNA、DNA等)及其相应评估指标(如表达水平、拷贝数等)。应基于以往相关研究或生物学机制,在研究中阐明预先设定的假说。例如,“在潜在的结直肠癌根治性切除术后患者中,胸苷酸合成酶(thymidylate synthase)表达水平和p53表达状态是其整体生存的标志物”^[5]。

条目2 应明确说明研究对象来源(如社区或医院)及纳入排除标准。纳入标准通常基于临床(如肿瘤位置和分期)、病理特征(如肿瘤的组织学类型)或其他因素(如患者的年龄、治疗方式)。排除标准可包括有肿瘤史、肿瘤治疗史、肿瘤样本不足、未取得知情同意或者关键数据严重缺失等。

条目3 应报告患者接受何种治疗及其选择依据,如不确定,必须谨慎解释预后和预测分析结果。例如,“根据淋巴结状态和激素受体的结果,决定是否进行化疗和/或激素等辅助治疗”^[6]。

条目4 应报告用于标志物检测的标本类型(例如肿瘤组织、细胞或DNA)及标本来源详细信息,如标本采集时间、采集方式(活检等)及采集部位(原发灶或转移灶)等。尽可能报告研究中所使用生物材料的收集、处理及储存方式,尤其是检测前其储存的时间和方式,还应提供对照样品的相关信息,如

来源、适用性等。

条目5 清晰完整报告检测方法,确保测量的可重复性。实验室技术应按如下报告:免疫组化应报告特异性抗体、抗原修复步骤、标准和对照、评分标准、得分报告及解释;DNA及RNA检测应说明特异性引物、探针及所使用的评分或定量方法。报告分析所需的最小样本量、样品适用性的评估方法、相关质量控制措施及样品预处理的步骤。还应报告降低测量误差的方法以及对临床结局是否使用盲法。

条目6

(1)研究对象选择:应详细说明患者如何纳入研究,例如是通过临床试验前瞻性招募还是采用数据库回顾性搜索。如患者选择还基于其他考虑应予以报告,如分层或匹配设计等。

(2)时间周期:应报告结局发生时间及随访结束时间,尽可能报告中位随访时间。确定中位随访时间的首选计算方法是reverse Kaplan-Meier法(该法可将删失数据作为感兴趣的终点结局,故使用队列中所有患者数据)。有时还需提供未发生事件患者的中位随访时间,并报告长期随访(>1年)中失访患者数量及完整数据。

条目7 应精确定义终点,若研究死亡,应明确死因如何分类(如全死因死亡,肿瘤相关死亡)并尽

可能描述死因的确定依据(如死亡证明等)。如果研究疾病复发,研究终点及结局(如无病生存期)均应精确定义。如果终点评估需要主观判断,应尽量报告是否使用了盲法。

条目 8 应列出并充分定义最初考虑纳入模型的所有变量,包括标志物测量或其他临床病理变量,以及最终模型未使用的变量。必要时还应解释如何测量这些变量。所有与生存分析相关变量最好在研究开始前或同时进行检测,之后检测的变量也应予以慎重考虑。

条目 9 应阐明样本量确定的理由,必要时需计算。有时样本量大小受限于实际(如可使用的肿瘤样本及其成本),也应报告在既定样本量情况下其检测效应价值。例如,“受经费和实际条件限制,本研究的样本量为 400 例”^[7];“假设对照组生存率为 60%,高 TS 表达或 p53 基因过度表达患者的生存率为 50%,则从 750 名患者的组织样本分析,在 10% 的绝对差下检测到这些标志物的表达与 OS 相关的效能为 80%”^[8]。

条目 10 应报告在分析中使用的所有统计方法,以使读者了解所选择方法的原因及评估结果。

(1)初步的数据准备:在统计分析前对数据的评估(包括修改或删除数值)应予以报告。首先考虑标志物以及任何其他变量的数值分布,以观察其是否存在极值或严重偏态并予以处理。分类变量可将频数小的类别进行合并,其他变量可用图形(例如箱线图)或统计量(例如平均数、中位数、标准差、全距、频率)描述其分布。

(2)标志物与其他变量的关联:在数据分析初期,应用相关系数、 χ^2 检验或 t 检验等描述标志物与标志物或标志物与其他变量(例如既定的临床、病理和人口协变量)的关联,并关注极端或异常关联。

(3)评估标志物与临床结局关联的单因素分析:通常先在未调整其他变量时进行单因素分析以评估标志物或其他某个变量与临床结局的关联。应报告分析方法(如秩和检验或 Cox 回归模型)、检验统计量(如 Wald 检验、似然比检验或得分检验)的选择以及对变量的其他处理(如编码、分组或连续变量转化等)。

(4)多因素分析:可用于评估调整既有预后变量后的标志物效应。为了便于比较调整前后的关联,应尽量报告单因素及多因素使用相同分析的结果。多因素分析也可用于构建涉及若干候选标志物的预测模型,但应避免过度拟合,并严格评估模型的预测能力。也可报告其他多变量统计方法(如分类和回归树以及人工神经网络)。

(5)缺失数据:应报告每个相关变量的缺失值数量。如有可能,还应给出缺失原因,并在描述研究人群的流程图中注明因缺失数据而被排除的人数。还应报告缺失值假设(如随机缺失)及其处理方法(如多重插补)。

(6)变量选择:应描述所使用的变量选择程序及其对应的变量纳入或排除标准,例如“当使用逐步回归法确定生存预后的独立因素时,变量增加采用前进法,纳入标准为 $P < 0.05$,变量删除采用后退消除法,排除标准为 $P > 0.05$ ”^[9]。

(7)检验模型假设:对变量分布或变量间的关系均可通过统计模型做出特定的假设。此时应报告使用经验性方法进行的模型假设检验。例如,“通过二次对数生存曲线图检查比例风险的假设”^[10]。

(8)模型验证:利用不参与原始分析的数据对结果验证,即“外部验证”,应提供验证所用研究人群、设计及分析方法的基本细节。“内部验证”是指先保留数据的一部分(“测试集”),在剩余部分上(“训练集”)建立模型;模型在训练集上完全确定后,再在测试集上进行评估,应描述其所使用的特定验证算法(例如靴襻法、10 倍或留一法交叉验证)。

条目 11 对连续型的标志物值应报告其如何纳入分析,是保持连续还是对变量进行了其他处理(如对数转换,作为线性或非线性建模及其如何建模)。对于分类变量,应该指定分界点以及如何选择。如有 ≥ 3 个有序分类变量,应明确其在分析中是如何处理的,如创建哑变量,亦要明确如何定义及其分析。

条目 12 结构化陈述是保证重要信息得以完整报告的重要途径。该指南提供了关于预后研究概况的两部分结构化报告模式。第一部分为报告肿瘤标志物具体情况及其他变量,也可以报告其他重要信息,如研究人群及其纳入和排除标准、入组患者及结局事件数量;第二部分为报告数据的分析概况。有关肿瘤标志物预后的研究均应包括单因素、多因素及亚组分析,甚至还需报告几种多因素分析的结果。

条目 13 应报告各变量分布及缺失情况,包括实际纳入的研究人群基本资料(如年龄、性别、民族或种族等)、相关预后变量(如病理分期信息,包括淋巴结情况、肿瘤大小、是否转移或一些临床报告如体能状况等)及肿瘤标志物。此外应尽可能将研究人群与未纳入研究人群进行人口学特征比较,以便读者评价该结果的普遍性。

条目 14 应报告标志物与传统预后变量的关联强度及关联本质,以正确评估标志物的效应。如

果均为连续变量,可用散点图及相关系数;如果既有连续变量也有分类变量,建议采用箱式图、散点图或直方图描述在分类变量不同水平下连续变量的检测值,同时应避免对连续变量分类;如果均为分类变量,可按标志物或其他传统变量的类别在交叉列联表中展示。

条目 15 应先报告在未调整其他变量时标志物与结局的简单关联,即单因素分析结果,这表明不考虑其他因素时该肿瘤标志物的预后强度。

如果临床结局是二分类变量,若标志物为分类变量,研究者应报告每组肿瘤标志物观察到的结局发生概率;若标志物为连续型变量,可分别报告出现或未出现临床结局这两组间肿瘤标志物的值。对于时间-事件结局,建议绘制 Kaplan-Meier 生存曲线。当标志物是分类变量时,可以用对数秩检验描述标志物与结局之间的关联,并报告危险比估计或其他估计值,并注明估计的精度及不确定性(如 *CI* 及 *P* 值等);对于连续型肿瘤标志物,可以报告未对肿瘤标志物进行分类时其对结局的影响,若分类后进行分析,还应报告分类方法。

条目 16 多因素分析通常是由全模型经变量筛选及建模,得到一个简约的“最终模型”。应该报告重要的多因素模型(至少对于最终模型)中所有变量的 *CI* 及 *P* 值,以确定估计效应的精确度及显著性。如果根据重要变量分亚组进行多因素分析(例如分男女性别分别做多因素分析),还应报告每个亚组模型中的相关指标。

条目 17 在评估一种新的肿瘤标志物是否与临床结局有关联时,需要考虑一些已被证明具有估计预后价值或已被临床指南推荐使用的经典预后变量,如疾病分期及其相关变量(肿瘤大小及淋巴结状态),或人口学变量如年龄、性别等。通过“标准化模型”,即将所研究的标志物及全部传统变量均纳入模型分析,得到调整所有变量后标志物的效应以便于不同研究间进行比较。应报告标准化模型及模型中标志物和每个传统变量(无论是否有统计学差异)的估计值、精确度及其显著性(*CI* 和 *P* 值)。

条目 18 如果模型的假设(例如比例风险,协变量的线性及缺失数据的处理机制等)均为合理,或其结果对某种假设并不敏感,那么预后分析结果具有较大可信度。如为检验假设或者证明结果的稳定性做了相关分析,也应予以简要陈述。如果采用敏感性分析,或进行内部及外部验证,其结果也应如实报告。

条目 19 在“讨论”部分可先简单重述研究目

的以及研究假设,然后对研究主要发现进行简要概括,客观评价报告结果(如数据存在的偏倚及不一致,检测方法的局限性,研究设计或数据分析方法的局限性等)。研究者也应评估本研究结果是否可外推至其他人,并讨论本研究结果与以往研究报道是否一致,若不一致则尽可能做出适当解释。

条目 20 研究任何一个标志物最根本原因就是为获得疾病生物学的相关信息,寻找新方法以帮助临床决策或开发新的治疗方法。若研究结果表明肿瘤标志物与临床结局之间有一定关联,可能提示其对于临床诊疗有重要价值;若标志物不能提供更多预后信息,但更易于评估、成本更低或重复性更好,也应有一定临床意义。因此讨论中还应明确本研究在多大程度上实现了此目的以及下一步如何去做。

5. 结语:完整清晰的报告便于评估研究的质量及对研究结果进行更好地解释和推广。REMARK 清单可以帮助研究者规范报告研究结果,同时对于开展新项目的研究者以及审稿人、编辑均有一定帮助。当然,REMARK 仍会根据新的文献证据及建议不断发展和完善,相关信息可浏览 EQUATOR 网站(<http://www.equatornetwork.org>)。相信随着 REMARK 的不断改进和广泛应用,肿瘤标志物的预后研究报告质量将不断提高,并推动该项研究的快速发展。

参 考 文 献

- [1] Mcshane LM, Altman DG, Sauerbrei W, et al. Reporting recommendations for tumor marker prognostic studies (REMARK). *J Natl Cancer Inst*, 2005, 97(16): 1180-1184.
- [2] Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA*, 2001, 285(15): 1987-1991.
- [3] Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ*, 2003, 326(7379): 41-44.
- [4] Altman DG, Mcshane LM, Sauerbrei W, et al. Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): explanation and elaboration. *PLoS Med*, 2012, 9(5): e1001216.
- [5] Popat S, Matakidou A, Houlston RS. Thymidylate synthase expression and prognosis in colorectal cancer: a systematic review and meta-analysis. *J Clin Oncol*, 2004, 22(3): 529-536.
- [6] de Mascarel I, Bonichon F, Durand M, et al. Obvious peritumoral emboli: an elusive prognostic factor reappraised. Multivariate analysis of 1320 node-negative breast cancers. *Eur J Cancer*, 1998, 34(1): 58-65.
- [7] Barratt PL, Seymour MT, Stenning SP, et al. DNA markers predicting benefit from adjuvant fluorouracil in patients with colon cancer: a molecular study. *Lancet*, 2002, 360(9343): 1381-1391.
- [8] Popat S, Chen Z, Zhao D, et al. A prospective, blinded analysis of thymidylate synthase and p53 expression as prognostic markers in the adjuvant treatment of colorectal cancer. *Ann Oncol*, 2006, 17(12): 1810-1817.
- [9] Efficace F, Bottomley A, Smit EF, et al. Is a patient's self-reported health-related quality of life a prognostic factor for survival in non-small-cell lung cancer patients? A multivariate analysis of prognostic factors of EORTC study 08975. *Ann Oncol*, 2006, 17(11): 1698-1704.
- [10] Lutterbach J, Sauerbrei W, Guttenberger R. Multivariate analysis of prognostic factors in patients with glioblastoma. *Strahlenther Onkol*, 2003, 179(1): 8-15.

(收稿日期:2013-07-24)
(本文编辑:张林东)