诊断试验准确性研究报告规范——STARD 介绍

诊断试验准确性研究是指在用参考试验或“金标准”确诊的患有某病和未患该病的小 样本中实施的一种评价研究,其目的是评价某种或某些诊断技术区分患和不患该病人群的真实性或效力(validity),有时又叫准确性(accuracy),可以用许多指标来表达,包括灵敏度、特异度、似然比、诊断比值比和 ROC 曲线下面积等。在一个诊断试验引入临床实践之前,对其研究的潜在的偏倚和适用性进行严格的评价,不仅能降低由于错误估计试验准确性而导致的有害临床后果的数量,而且能够通过防止不必要的试验来限制卫生保健费用。然而由于许多诊断试验准确性报告缺乏关于诊断性研究设计、实施和分析的重要信息,使得评价难以进行。诊断试验准确性研究报告标准(STARD)正是为了改进诊断准确性研究报告质量而发起的,通过建立一个科学、规范、循证的报告标准,使得读者能够通过完整、准确的报告评价研究结果的内部有效性(潜在偏倚)和外部有效性(适用性)。STARD 筹备委员会通过文献检索,找出正确实施和报告的诊断性试验,并提取可能有用的 75 项条目形成一份泛表。2000 年 9 月在荷兰阿姆斯特丹举行的共识会议上,由相关领域专家对这份泛表进行缩减,最终形成了一个由 25 项条目组成的清单——STARD,具体见表 1。并制定出典型流程图。这有利于提高诊断准确性研究报告的准确性和完整性,以便读者评估研究的偏倚可能性及推广性,更好地服务于临床医生、研究人员、评价者、期刊及公众。

表 1　STARD 清单

项目与主题		条目号	描述
文题和摘要		1	确定该文章是诊断试验准确性研究(推荐使用 MeSH 主题词灵敏度和特异度)
前言		2	陈述研究的问题或目的,例如评估诊断试验的准确性或比较不同试验或不同研究对象之间的准确性
方法	研究对象	3	描述研究对象的纳入和排除标准,收集资料的来源
		4	描述研究对象情况,是否是基于临床症状,或是以前试验的结果,是否接受过目标试验或参考试验
		5	描述研究对象抽样,是否是符合上述第 3、4 条目标准,如果不是,则描述是如何选择的
		6	描述资料收集,是前瞻性(做新试验和参考试验前计划好),还是回顾性资料
	试验方法	7	描述参考试验(金标准)及其原理、合理性
		8	描述所涉及的材料和方法的技术特点,包括如何和何时测定,列出目标试验和参考试验的文献依据
		9	描述目标试验和参考试验结果的单位、临界点、分类的合理性
		10	描述目标试验和参考试验操作和阅读结果人员的人数和培训情况
		11	对目标试验和参考试验进行结果测定者是否设盲,有无提供其他临床信息给他们
	统计学方法	12	描述计算及比较诊断准确性的测定和统计方法,包括 95% 可信区间
		13	描述测定试验可重复性的方法
结果	研究对象	14	报告其进入和完成试验的日期
		15	报告其临床和一般状况,如年龄、性别、症状谱,同时伴随的疾病,目前治疗和征集中心
		16	描述符合纳入标准进入或未进入目标试验和参考试验的数量,以及其么不能参加试验的原因(推荐用流程图表示)
	试验结果	17	报告从做目标试验和参考试验之间间隔的时间,以及在此期间有无任何治疗措施
		18	报告疾病严重程度的分配情况及对照组的具体情况
		19	报告目标试验与参考试验比较的所有结果(包括不确定结果和遗漏的结果),对于连续性结果报告两种试验的分布情况
		20	报告目标试验与参考试验发生的不良事件
	评估	21	报告试验的准确性估计结果和统计学不确定性测量结果(如 95% CI)
		22	报告如何处理目标试验出现的不确定结果、遗漏的结果和异常结果
		23	报告对试验可变性的估计,包括不同操作者、不同中心或亚组的测定结果的准确性及有效性
		24	报告对试验可重复性的评估
讨论		25	讨论试验结果及研究发现的临床适用性