

替代指标在临床试验中的应用与局限性

徐小国, 蒋 萌

南京中医药大学附属医院临床药理科, 南京 210029, 江苏

摘要 本文依据国内外临床试验中结局指标选择的现状, 论述了替代指标与终点指标的概念及其关系, 以及在临床试验中结局指标选择基本方法及其当前存在的问题。强调只有对患者有直接影响的终点指标才能够证实干预措施的真正疗效, 而以实验室为主的生物学指标和患者的临床症状或体征等替代指标, 只有被证实与真正临床结局具有因果关系且能客观反映临床措施对临床结局净效应的前提下, 才能够作为代替终点指标的评价指标使用。

关键词 替代指标; 终点指标; 临床试验; 应用; 局限性

中图分类号: R969

文献标识码: A

文章编号: 1009-2501(2006)10-1194-04

由于传统临床医学模式的影响, 对于疾病的疗效评价标准过去多着重于症状学、形态学、病理学及生化等指标的改变^[1]。随着循证医学的兴起, 临床医学模式发生了由以经验为基础到以证据为基础的转变, 现代医学逐渐重视对人体功能活动、生存质量和影响健康等重大事件的评价。

在临床干预的疗效评价上, 如何从整体水平选择终点指标(包括重要临床事件、功能状态、生存质量和受试者满意度等)已成为目前国际医学界一个新的研究热点。世界卫生组织(WHO)在1996年的研讨会上提出当前临床试验中应引起高度重视的6个关键环节, 即(1)纳入标准; (2)合适的对照; (3)随机分组; (4)盲法; (5)报告所有入选病例的结果; (6)

选择明确的、有临床意义的以病人为基础的终点指标^[2]。其中将终点指标的选择提升到和随机、对照、盲法几大原则相同的高度上。在临床试验中, 选择何种指标作为观察终点, 对于干预措施与疗效之间的因果关联推断具有十分重要的作用, Herson指出采用不同的评价指标判断疗效甚至可以得出两种截然相反的结论^[3,4]。

1 临床试验的终点指标与替代指标

过去, 大多只选择与疾病相关的临床症状体征、病理学检查、实验室指标等来证明临床措施的有效性, 而这些指标难以对生命质量做出全面的评价。现在的观点提出要分清终点指标(或称主要结局指标)与替代指标(或称次要结局指标)的临床意义, 强调从人体对临床措施的整体反应来选择相关的结局指标^[5]。

替代指标(surrogate outcome): Prentice将其定义为一种能够代替重要的临床结局的实验室测量指标、临床症状或体征, 经过干预可使其变化。这种变化能直接反映病人的感觉、功能状况或生存情况^[6,7]。替代指标只有在被证实与重要临床结局具有相关性, 并确定是由于治疗所带来的结果时, 作为疗效指标才有意义。

终点指标(end point): Fleming给其定义为那些对病人影响最大、最直接、病人最关心、最想避免的临床事件, 包括疾病终点(如死亡、残疾、功能丧失)和某些重要的临床事件(如骨折)^[8]。应该指出: 随着医学模式的转变, 综合评价病人主观感受、功能状态、生命质量等指标也得到了越来越多的应用, 并被认为是病人最直接相关、病人最关心的主要结局指标^[9]。

终点指标是真正的疾病结局, 能反映干预的真正效果, 偏倚较小, 但出现时间晚, 试验所需时间长, 样本需要量大, 花费大。若结局的出现需较长时间

2006-06-12收稿 2006-09-10修回

徐小国, 男, 硕士研究生, 研究方向: 临床药理。

Tel: 13913870831 E-mail: xuxiaoguo1981@sohu.com

蒋萌, 通讯作者, 主任医师, 博士生导师, 研究方向: 临床药理。

Tel: 025-86555083 E-mail: jmjszy@sina.com

干预,那么结局易受其他非干预因素干扰。替代指标不是真正的临床结局,虽出现时间较早,节省试验时间和花费,较客观、变异小、易测量,往往能大大减少样本量,但若不经过严格的效度评价,往往产生偏倚。值得指出的是现在较多常用的替代指标极少满足严格的效度评价条件。

2 临床试验中结局指标的选择与应用条件

由于疾病的复杂性(多因果通道)和治疗的复杂性(多治疗效应),干预和疗效的因果关联推断显得非常复杂,以致多年来在临床措施疗效评价指标的选择上似乎总处于难以抉择的境地。临床试验的最终目的是为临床治疗决策提供依据,因此选择恰当的判效指标是临床试验的重要环节之一。美国替代医学办公室(OAM)顾问委员会在1995年的“替代医学研究方法论”中就明确指出:传统/替代疗法的“有效性评价是一个关键和核心的问题”、“其疗效必须用人们认可的终点指标来加以证实”^[10]。

有专家提出:如果疾病的后果非常严重(在可预见的较短时间内危及生命)且目前无有效的治疗措施,或疾病病程缓慢,从发病到真正临床结局出现需较长时间,伴随临床试验需大量的人力物力、病人依从性差和某些主观指标变异大等情况,那么宁愿先用那些通过替代指标表明可能对病人主要终点有益的药物对病人进行治疗,因为病人得不到可能有效治疗的后果会大于替代终点不当所造成的损害。如果疾病后果不那么严重,或目前已有一定效果的治疗药物,那么评价新的干预措施有效性时应采用对病人更直接、更相关的指标(如生存、死亡、功能残障等)作为临床试验的主要结局指标^[11]。如在心衰治疗的试验中运动能力的提高作为替代指标的使用有时是不恰当的,应该用生存作为其疗效评价的结局指标^[12]。

在疾病的发展过程中,替代指标的变化出现于真正临床结局之前,发生的频率高,容易测量,因而观察其变化可缩短研究时间,节约经费。为此,近年来许多临床研究采用与疾病进程相关的替代指标来评价干预措施对疾病临床结局的总体疗效^[13,14]。Prentice强调一个合格的替代指标必须满足两个条件^[7]:一是该指标必须与真正的临床结局有因果关系,即替代指标与真正的临床结局有高相关性,并且可以预测疾病结局,二是它可以完全解释由治疗引起的临床结局变化的净效应。其中第二条更为重要,但可惜的是现在应用的很多替代指标极少满足

第二个条件。例如在评价冠心病干预措施的临床疗效时,既往多以心绞痛的缓解程度及心电图的改善情况作为评价指标,这些指标往往难以评价急性冠心事件(如心肌梗死、猝死等)的发生情况^[15]。

3 替代指标应用的局限性及其与临床结局因果关系的不确定性

替代指标虽然用起来可以效率更高,但在评价治疗效果上常常存在错误的危险,尤其是使用未经充分验证的不恰当的替代指标,其结论的临床推广性是非常局限的。替代指标之所以会导致如此严重的错误,Fleming罗列了以下最常见的几个原因^[8]:

(1)替代指标虽然与临床终点有关,但不在疾病进展的因果通道上;(2)疾病的因果通道有几个,干预措施只影响了替代指标所在的那个因果通道;(3)干预措施通过替代指标以外的因果通道影响疾病的临床结局;(4)干预措施不仅通过替代指标所在的因果通道影响临床结局,而且还通过其它我们未曾认识的因果通道影响临床结局;(5)虽然干预措施能影响替代指标所在的因果通道,但替代指标与临床终点没有相关性。Ellenberg补充了两点^[16]:(1)替代指标与终点指标存在不一致性,如瘤体大小(或有无)与生存期的相关性。理论上认为,通过手术、放疗和/或化疗可以消灭或缩小肿块,病人的生存期会相应延长,但事实证明并非完全如此。临床实践中经常遇到患者病灶虽然消失了,但人也死了,即瘤消人死亡^[17]。(2)干预措施虽然能够通过替代指标所在的因果通道影响临床结局,但是它所引起的未曾预料的安全性远大于其正性作用,如激素替代疗法(HRT)虽然能预防绝经后骨质疏松性骨折的发生,但同时也增加患浸润性乳腺癌的相对危险性^[18]。

另一个值得关注的问题是,虽然以前的研究资料 and 已知的生物学知识已强有力的证实某些替代指标与临床结局有高度相关性,但用这些指标去评价新的治疗措施(和以前的治疗属同一类)的临床疗效时,仍可能存在很大的风险。

例如血压与心血管事件的发生有非常密切的关系,Collins等研究证实血压每降低1 mmHg高血压人群心血管病的死亡率就下降5%^[19]。另一项早期的研究说明使用小剂量利尿剂在降低高血压病人血压的同时,能有效降低患者心血管事件的发生,使心血管病的死亡率下降17%^[20]。据此人们推测只要降低血压,就能有效降低心血管病的发病率与死亡率,许多国家在批准新型抗高血压药物使用时,都以其

降压作用作为标准,而无须证明是否能够预防心血管事件的发生。80年代初两种新型抗高血压药物 Ca^{2+} -拮抗剂与肾素-血管紧张素转换酶抑制剂(ACEI)问世,现已广泛用于临床,然而愈来愈多的资料表明,尽管这两类药物有良好的降血压作用,且副作用小,但并不能有效降低心血管的发病率与死亡率。Brown等在一项关于高血压药物治疗与心梗发生率的研究中显示,服用 Ca^{2+} -拮抗剂的高血压患者发生心肌梗塞的危险性更大^[21]。由此可知,尽管在SHEP试验中血压值作为利尿剂疗效评价的替代指标是合格的,但在评价其它抗高血压药物临床疗效,而该治疗措施与过去评价和确认有效性的治疗措施作用机制并不相同时,使用血压值作为抗高血压药物结局评价指标就可能存在风险。

再如绝经后妇女由于骨量减少、骨微观结构退化等导致脆性增加,骨折发生率提高。作为一种替代指标,骨矿物质密度(BMD)曾广泛用于骨质疏松症(预防骨质疏松性骨折)干预措施疗效评价。在一项使用阿兰屈酯预防绝经后骨质疏松症的试验中^[22],BMD增高与椎骨新的骨折发生率降低平行,但另一项使用氟化钠对绝经后妇女骨质疏松症预防试验中却得出了相反的结论^[23,24]:虽然氟化钠在四年内使绝经后妇女腰椎BMD增高35%,但与安慰剂相比治疗组发生了更多椎骨新的骨折(氟化钠组101例妇女发生了163次骨折,安慰剂组101例妇女发生了136次骨折)和非椎骨新的骨折(氟化钠组101例妇女发生了72次骨折,安慰剂组101例妇女发生了24次骨折)。虽然BMD与骨质疏松症有高度相关性,但以BMD作为防治骨质疏松症性骨折的替代指标在评价治疗骨质疏松药物的疗效方面存在着很大的不确定性。骨质疏松患者是否发生骨折,除了与BMD外,还与骨骼的几何形状、年龄、曾发生骨折的病史、体重、身体的灵活性、骨生物力学强度和受力强度等有关。这些因素可能在骨质疏松患者是否发生骨折的干预疗效评价中起更主要的地位。

4 评价与展望

目前国内许多临床试验仍较多地依赖于疾病相关实验室检查和临床症状体征的改善来证明临床措施的效能,尽管这些指标的作用不可忽视,但用这些指标评价治疗的效果和临床意义有时是不恰当的。严格的临床试验应尽可能采用功能状态及生存质量等疾病转归的结局指标对临床措施进行疗效评价,使得出的结论更具有临床推广的价值。

参考文献

- 赖世隆. 中医药临床疗效的评价[J]. 中国中医药信息杂志, 2000; 7: 88—9
- World Health Organization Report. Consultation meeting on traditional and modern medicine; harmonizing the two approaches [C]. Beijing, China, 1996; 16
- Jay H. The use of surrogate endpoints in clinical trials [J]. Statistics in Medicine, 1989; 8: 403
- 胡大一, 项志敏. 临床医学模式转变—从经验医学到循证医学[J]. 中国实用内科学杂志, 1998; 18: 668
- 魏华凤, 郑培永, 季光. 中医临床疗效评价的思路与方法[J]. 中西医结合学报, 2005; 3: 184—90
- Temple RJ. A regular authority's opinion about surrogate endpoints [A]. Nimmo WS, Tucker GT, eds Clinical measurement in drug evaluation [M]. New York: John Wiley & Sons Ltd, 1995; 3—22
- Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria [J]. Stat Med, 1989; 8: 431—40
- Fleming TR, DeMets DL. Surrogate endpoints in clinical trials: are we being misled [J]? Ann Intern Med, 1996; 125: 605—13
- 郭新峰, 赖世隆, 梁伟雄. 中医药疗效评价中结局指标的选择与应用[J]. 广州中医药大学学报, 2002; 19: 251—5
- Levin JS, Glass TA, Kushi LH, Schuck JR, Steele L, Jonas WB. Quantitation methods in research on complementary and alternative medicine. A methodological manifesto [J]. Med Care, 1997; 35: 1079
- 郭新峰, 朱泉, 赖世隆. 替代指标和中间指标及其在中医药疗效评价研究中应用价值的思考[J]. 中国中西医结合杂志, 2005; 25: 585—90
- Bucher HC, Guyatt GH, Cook DJ, Holbrook A, McAlister FA. Users' guides to the medical literature: XIX. Applying clinical trial results. A. How to use an article measuring the effect of an intervention on surrogate endpoints [J]. JAMA, 2000; 19: 192—7
- Stuart G. Baker. Surrogate Endpoints: Wishful Thinking or Reality [J]. J Natl Cancer Inst, 2006; 98: 502—3
- 方向华. 替代终点在心血管病临床试验中的运用及其局限性[J]. 循证医学, 2005; 5: 364—6
- 宋军, 陈可冀. 中医药临床疗效评价若干问题思考[J]. 中国中西医结合杂志, 2003; 23: 564—5
- Susan S. Ellenberg surrogate endpoints: the debate goes on [J]. Pharmacoepidemiol Drug Safety, 2001; 10: 493—6
- 王树堂, 周岱翰. 对于目前恶性肿瘤临床疗效评价标准的评价[J]. 天津中医药, 2004; 21: 227—80
- Tseng LN, Sheu WH, Ho ES, Lan HH, Hu CC, Kao CH. Effects of alendronate combined with hormone replacement therapy on osteoporotic postmenopausal Chinese women [J]. Metab Clin Exp, 2006; 55: 741—7
- Collins R, Peto R, MacMahon S, Hebert P, Fiebach NH, Eberlein KA, et al. Blood pressure, stroke, and coronary heart disease. Part 2, Short-term reductions in blood pressure: overview of randomized drug trials in their epidemiological context [J]. Lancet, 1990; 335: 827—38

- 20 SHEP Cooperative Research Group. Prevention of stroke by antihypertensive drug treatment in older persons with isolated systolic hypertension: final results of the systolic hypertension in the elderly program (SHEP) [J] . JAMA, 1991; 265: 3255—64
- 21 Psaty BM, Heckbert SR, Koepsell TD, Siscovick DS, Lemaitre R, Smith NL, *et al.* The risk of incident myocardial infarction associated with anti-hypertensive drug therapies [J] . Circulation, 1996; 91: 925
- 22 Ettinger b, San Martin J, Crans G, Pavo I. Differential effects of teriparatide on BMD after treatment with raloxifene or alendronate[J] . J Bone Mine Res, 2004; 19: 745—51
- 23 Riggs BL, Hodgson SF, O' Fallon WM, Chao EY, Wahner HW, Muhs JM, *et al.* Effect of fluoride treatment on the fracture rate in postmenopausal women with osteoporosis[J] . N Engl J Med, 1990; 322: 802—9
- 24 Riggs BL, Seeman E, Hodgson SF, Taves DR, O' Fallon WM. Effect of the fluoride/calcium regimen on vertebral fracture occurrence in postmenopausal osteoporosis. Comparison with conventional therapy[J] . N Engl J Med, 1982; 306: 446—50

Use and limitation of surrogate outcomes in clinical trials

XU Xiao-guo, JIANG Meng

Clinical Pharmacological Section of Hospital Attached to Nanjiang University of Traditional Chinese Medicine, Nanjing 210029, Jiangsu, China

ABSTRACT According to the status quo of choosing outcome indexes in Chinese and other countries' clinical trials, propose the concepts of surrogate outcomes and end points and correlation between them, discuss the general methods and present problems of choosing true endpoints. It is emphasized that only the primary outcomes which affect the patients directly can prove the real effectiveness of the interventions. Only when the surrogate ones, such as

biological indexes from laboratory and the patients' symptoms or signs, which are proved to have a correlate of the true clinical outcome and fully capture the net effect of treatment on the clinical outcome, can the latter be the substitute for the former.

KEY WORDS surrogate outcomes; end points; clinical trials; use and limitation