

• 循证医学 •

原发性骨质疏松危险自我测评工具筛检效果系统评价

吴秀云¹

摘要 目的:评价骨质疏松症危险测评工具 OST/OSTA 的筛检效果。**方法:**检索 MEDLINE, EMBASE and CBM 及 PubMed 数据库 2000—2007 年英文相关文献,同时手工检索参考文献中的相关文献。按照事先制定入选与剔除标准筛选文献。采用循证医学 Meta 分析方法,根据异质性检验结果,Meta 分析采用随机效应模型。**结果:**无论用于男性与女性人群,OST/OSTA 具有较高的敏感性,某些研究同时具有较高的特异性。妇女危险评价工具合并敏感度与合并特异度分别是 87.0%、51.0%,合并受试者诊断特征曲线下面积(SAUC)均数为 0.734。**结论:**OST/OSTA 具有一定筛检效果,敏感性较高,在不同种族人群中均取得较好效果,但某些研究特异性较低。评价工具的应用为临床医生筛选可疑患者,确定那些患者最可能从骨密度测量中受益提供了快捷、可行的手段。对 OST/OSTA 的效果验证研究尚有待于深入进行,并需要与其他筛检工具进行比较。

关键词 骨质疏松症; Meta 分析; 骨质疏松危险自我测评工具

中图分类号:R681,R493 **文献标识码:**B **文章编号:**1001-1242(2008)-12-1102-04

骨质疏松症是全球重大公共卫生与健康问题之一。世界上 60—70 岁的妇女约有 1/3 患有骨质疏松症,80 岁以上的妇女 2/3 患有骨质疏松症。骨质疏松症的最主要后果是导致骨折,给社会、医疗卫生服务系统、家庭与个人带来沉重经济负担^[1-3]。鉴于骨质疏松症及其后果骨折对人群健康构成的严重危害,研究应用经济实用、快速准确、安全无痛苦的筛检方法,以识别骨质疏松症可疑患者对于早期发现、早期诊断,早期治疗骨质疏松症非常重要。近年来,国内外骨质疏松学术界致力于骨质疏松症筛检方法与效果的研究,发展并使用了许多简易的筛检工具。其中骨质疏松症危险自我测评工具(the osteoporosis self-assessment tool,OST)与骨质疏松症亚洲人群危险自我测评工具(the osteoporosis self-assessment tool for Asians,OSTA),分别在我国及世界其他国家不同人群中验证其有效性。OST/OSTA 评价工具仅包括年龄和体重两个危险因素,使用简单,便于个人进行自我评价,那么这一筛检工具的总体效果如何,是否可以作为一种有效的筛检工具在人群中推广应用,目前还没有肯定而一致的意见。因此本文采用循证医学系统评价方法,对 OST/OSTA 筛检效果进行评价。

1 资料与方法

1.1 研究内容

研究内容包括筛检工具文献基本信息,如研究文献来源、作者及发表年限;研究设计及样本特征等;筛检效果评价包括不同截断值、不同 t 值下的敏感度(sensitivity)与特异度(specificity),阳性预测值,阴性预测值,受试者诊断特征曲线下面积(area under curve,AUC),合并敏感度、特异度与合并受试者诊断特征曲线下面积(SAUC)等。

1.2 文献检索及数据提取方法

利用 MEDLINE,EMBASE 及 PubMed 数据库检索 2000—2007 年英文相关文献,生物医学文献数据库检索同期发表中文文献,同时用手工检索查找参考文献中的相关文献。MeSH 检索关键词包括骨质疏松症、筛检、诊断实验、OST、OSTA、危险评价、敏感度与特异度,受试者诊断特征

曲线。对入选文献,按照事先制定的标准化数据提取表,提取所需要的数据,录入计算机进行统计分析。

1.3 研究入选与剔除标准

文献入选标准是研究设计包括横断面研究,队列研究,随机对照实验,病例对照研究;研究疾病为原发性骨质疏松症患者,研究对象为 30 岁及以上妇女和男性;筛检工具以双能 X 线吸收法(dualenergy X-ray absorptiometry,DEXA)作为骨密度测量的金标准;研究报告筛检方法的敏感度与特异度或操作特征曲线。文献剔除标准是继发性骨质疏松症;筛检工具以 DEXA 以外的其他设备作为骨密度测量和骨质疏松症诊断的金标准;非 OST/OSTA 筛检工具;研究没有报告筛检方法的敏感度与特异度或受试者诊断特征曲线;文献为评论、信件、报告、指南或综述。

1.4 研究质量评价方法

采用由 Penny Whiting 教授等提出的筛检实验研究质量评价量表“诊断准确性研究质量评价工具(QUADAS tool)”对入选文献进行质量评价^[4]。该量表包括 14 个条目,主要从偏倚、变异、报告质量三方面对文献质量进行评价。质量评价内容包括研究设计是否科学合理;样本含量及代表性;选择试验对象标准是否明确,分组是否随机化;筛检实验是否独立于诊断“金标准”;实验的可重复性;实验结果解释是否独立于诊断“金标准”的结果;是否说明试验对象退出情况;统计分析方法是否正确等。每一条目以“是”、“否”、“不清楚”评价。满足此条标准为“是”,不满足或未提及此条标准为“否”,部分满足或从文献无法得到有关信息记为“不清楚”。为综合评分考虑,我们把评价“是”的条目标记为 5 分,“否”记为 1 分,“不清楚”记为 3 分。请 3 个专家采用盲法对每篇文献进行评价,然后根据各研究的综合得分确定文献质量。所谓盲法即隐去作者姓名和期刊、发表时间等信息。文献质量最终分为低质量文献和高质量文献两个等级,综合评分在 3.5 分以上

1 山东省潍坊医学院卫生管理学院,261042

作者简介:吴秀云,女,教授

收稿日期:2008-04-01

为符合质量标准的文献。

1.5 统计学分析

描述性分析入选研究的基本情况,比较不同研究的筛检效果。对符合 Meta 分析质量要求的研究采用 Meta-DiSc1.2 进行 Meta 分析。其中异质性检验用森林图(forest plot)分析,随机效应模型(random effect model),DerSimonian & Laird 方法分析显示结果。

根据骨质疏松症危险测评工具研究对象的特征及骨密度测量部位的不同,应进行亚组分析。本研究按性别、 t 值、截断值及密度测量部位的不同进行分组分析。

2 结果

2.1 OST/OSTA 研究基本情况

骨质疏松症危险自我测评工具 OST/OSTA 是由 Koh,L 教授等在亚洲 8 国更年期后妇女人群中发展使用的^[5]。随后有数项研究分别在亚洲人群及欧美白人人群中验证此工具的筛检效果。它是基于年龄与体重两个危险因素,危险分数计算简单,便于个人自我评价。筛检工具危险分数计算方法是: $0.2 \times (\text{体重} - \text{年龄})$ 。

通过文献检索得到相关文献 28 篇,其中有 14 篇文章属于其他筛检工具被排除。13 篇为危险自我测评工具 OST/OSTA,其中一篇因质量不符合纳入标准(样本为社区教堂志愿者)被剔除。因此有 13 篇符合纳入标准入选本研究。纳入研究中有 4 篇为男性骨质疏松症危险评价工具。入选研究设计全部为横断面研究,各研究基本情况见表 1。

2.2 OST/OSTA 用于妇女骨质疏松症筛检效果

表 1 危险评价工具 OST/OSTA 入选研究基本特征

编号	发表时间,作者及国家	样本量及分组样本量(人)	样本来源	样本种族或国家	平均年龄(岁)(最小与最大值)
13	2007,Huang QR(中国)	N=405	医院骨质疏松门诊	中国人	63(47-90)
12	2005,Kung AWC(中国)	N=776(男性)1 组:发展队列 420; 2 组:验证队列 356	三级医院病例记录	中国人	1 组: 65 (50-93) 2 组:64 (50-90)
11	2005, Lynn HS(中国)	N=1970(男性)1 组:发展队列 1182; 2 组:验证队列 788	人群样本	中国人	1 组: (65-NR) 2 组: NR (65-NR)
10	2003,Adler RA(美国)	N=181(男性)	医院风湿病与肺部疾病门诊	美国白人	64 (32-87)
9	2005,Gourlay ML(比利时)	N=40351 组:45-64 岁 2539; 2 组:≥65 岁 1496	医院专科门诊	白人	1 组:56(45-64) 2 组:70(65-96)
8	2005,Li-Yu JT(菲律宾)	N=15971 组:妇女 1465; 2 组:男性 132	三级医院病例记录	菲律宾人	1 组:60(40-89) 2 组:63(40-89)
7	2004,Geater S(泰国)	N=388	妇科门诊	泰国人	61(43-91)
6	2004,Richy F(比利时)	N=4035	医院专科门诊	白人	62(45-96)
5	2004,Cadarette SM(加拿大)	N=644	非专科门诊	未报告	62(45-90)
4	2003, Park HM(韩国)	N=1101	妇科门诊	韩国人	59(48-87)
3	2003, Kung AWC(中国)	N=722	社区人群样本	中国人	62(43-81)
2	2002,Geusens P(荷兰)	N=1102	人群样本	白人	61(45-87)
1	2001,Koh L(亚洲)	N=19831 组:亚洲人群 860; 2 组:日本人群 1123	非专科门诊	亚洲 7 国	1 组:62(45-88) 2 组:NR

注:NR 表示未报告

入选研究筛检结果见表 2。Koh 等在 860 名亚洲 8 国更年期后妇女人群中发展了筛检工具 OSTA,并在 1123 名日本更年期后妇女人群中验证 OSTA 的有效性,在截断值取 $OSTA \leq -1$, T 值 ≤ -2.5 ,骨密度测量部位为股骨颈时,发展队列敏感度为 91.0%,特异度为 45.0%,验证队列敏感度为 98.0%,特异度为 29.0%^[5]。之后 Geusens P (荷兰)、Kung(中国)、Cadarette SM (加拿大)、Geater S(泰国)等相继在其研究人群中验证此工具的效果,取相同截断值 $OSTA \leq -1$, T 值 ≤ -2.5 ,敏感度范围为 79.0%—97.5%,特异度为 45.0%—69.5%^[6-14]。中国黄琪仁等^[14]以上海 405 名绝经后妇女为发展队列,敏感度、特异度及 AUC 均数分别为 93.4%、52.6%、

0.818,与 Koh 等研究中发展队列结果十分近似。OSTA 在欧美国家妇女人群中验证时称为 OST,也取得近似效果。

为了与 Koh 的研究保持可比性,本文对妇女骨质疏松症危险评价工具与 Koh 的发展队列采用 t 值、截断值一致的研究,即 t 值 ≤ -2.5 、 $OSTA \leq -1$ 的研究进行了 Meta 分析。森林图合并敏感度、合并特异度及 SROC 见图 1—3。

纳入分析的 10 组数据,异质性检验敏感度与特异度差异较大, $P < 0.01$,资料缺乏齐同性,故采用随机效应模型中 DerSimonian & Laird 方法进行敏感度与特异度的综合分析。Meta 分析后同类研究合并敏感度为 87.0%,合并特异度为 51.0%,合并受试者诊断特征曲线下面积均数为 0.734,大于

图 1 Meta 分析结果——合并敏感度

图2 Meta 分析结果——合并特异度

OST/OSTA 及简易设备的筛检效果评价。本研究有 4 项研究是针对男性骨质疏松症危险评价的。其中中国香港 Kung 教授等以 420 名 65 岁及以上男性作为发展队列,以 356 名 64 岁及以上男性组成验证队列,取截断值 $OSTA \leq -1$, T 值 ≤ -2.5 , 骨密度测量部位为股骨颈时,其敏感度、特异度、AUC 分别是发展队列 81.0%、66.0%、0.830, 验证队列 83.0%、67.0%、0.850。当其他指标不变,骨密度测量部位为任一部位时,其敏感度、特异度分别是发展队列 73.0%、68.0%, 验证队列 71.0%、68.0%^[16]。Li-Yu JT(菲律宾)等在其验证研究中,取截断值 $OSTA \leq -1$, T 值 ≤ -2.5 , 骨密度测量部位为股骨颈,以 132 名男性为对象获得敏感度 90.9%、特异度 66.1%、AUC 0.847^[12]。Lynn HS(中国)等在另外一项研究中,取截断值 $OSTA \leq 2$, T 值 ≤ -2.5 , 骨密度测量为任一部位时,其发展队列敏感度、特异度、AUC 分别是 99.4%、27.1%、0.823^[17]。与 Lynn HS 的研究取相同截断值和 T 值,骨密度测量为任一部位时,Adler RA(美国)的验证研究得敏感度、特异度、AUC 分别为 82.0%、74.0%、0.836。可见 OST/OSTA 用于男性骨质疏松症筛检同样取得较好效果^[18]。见表 2。

图3 Meta 分析结果——合并受试者诊断特征曲线

0.5,因此综合评价认为 OST/OSTA 对妇女人群筛检效果是肯定的。

2.3 OST/OSTA 用于男性骨质疏松症筛检效果

大部分骨质疏松症危险评价工具针对妇女人群进行设计与使用。尽管骨质疏松症男性人群发病率低于女性,但骨质疏松症引起的骨折,男性死亡率要高于女性^[15]。因此近几年学术界加强了男性骨质疏松症筛检工具研究,主要集中在

表 2 男性骨质疏松症危险评价工具 OST/OSTA 筛检效果

编号	发表时间,作者及国家	截断值	BMD 测量部位	T 值	组数	分组描述	敏感度	特异度	操作者特征曲线下面积(均数)
12	2005,Kung AWC(中国)	≤ -1	股骨颈	< -2.5	2	发展队列	81.0	66.0	0.830
12	2005,Kung AWC(中国)	≤ -1	股骨颈	< -2.5	2	验证队列	83.0	67.0	0.850
12	2005,Kung AWC(中国)	≤ -1	任一部位	< -2.5	2	发展队列	73.0	68.0	0.790
12	2005,Kung AWC(中国)	≤ -1	任一部位	< -2.5	2	验证队列	71.0	68.0	0.780
11	2005,Li-Yu JT(菲律宾)	≤ -1	股骨颈	< -2.5	2	验证队列	90.9	66.1	0.847
11	2005,Li-Yu JT(菲律宾)	≤ -1	股骨颈	< -2.0	2	验证队列	72.0	69.1	0.754
10	2003,Adler RA(美国)	1	任一部位	< -2.0	1	验证队列	62.0	89.0	0.815
10	2003,Adler RA(美国)	1	任一部位	< -2.5	1	验证队列	75.0	80.0	0.836
11	2005, Lynn HS(中国)	2	任一部位	< -2.5	2	发展队列	99.4	27.1	0.823
11	2005, Lynn HS(中国)	2	任一部位	< -2.5	2	验证队列	NR	NR	0.839
10	2003,Adler RA(美国)	2	任一部位	< -2.5	1	验证队列	82.0	74.0	0.836
10	2003,Adler RA(美国)	2	任一部位	< -2.0	1	验证队列	69.0	82.0	0.815
10	2003,Adler RA(美国)	3	任一部位	< -2.0	1	验证队列	74.0	72.0	0.815
11	2005, Lynn HS(中国)	3	任一部位	< -2.5	2	发展队列	94.3	45.7	NR
10	2003,Adler RA(美国)	3	任一部位	< -2.5	1	验证队列	93.0	66.0	0.836
11	2005, Lynn HS(中国)	7	任一部位	< -2.5	2	发展队列	30.0	96.4	NR

说明:(1)NR=未报告;(2)任一部位是指股骨颈、腰椎或髋部三者之一

3 讨论

一种好的筛检工具应具有高的敏感性,以便能够尽可能多的检测出人群中真正患病者,同时也应有高的特异性,即识别非患者的能力强,以减少假阳性。Meta 分析表明,危险评价工具 OST/OSTA 具有较高的敏感性(87%),能够识别大部

分骨质疏松症患者,但其特异性(51%)尚不高。这意味着筛检结果将有相当比例的假阳性,近一半的正常人(49%)将被误判为骨质疏松症而建议进行不必要的 DEXA 骨密度检验。而综合受试者诊断特征曲线下面积分析显示,AUC 均在 0.7 以上,说明其筛检效果还是肯定的。是否可以将 OST/OSTA 作为

常规筛检手段尚需要进行深入研究,尤其需要与其他筛检工具进行比较研究。

20世纪60年代以来,国内外对骨质疏松症筛检工具及其效果进行了大量研究,提出了许多不同的筛检手段。研究的重点集中于各种工具筛检的可靠性与准确性评价。除了OST/OSTA工具,还有其他评价工具。如Lydick教授等首先提出了包括六个危险因素、用于绝经期后妇女筛检的简单计算骨质疏松危险评价工具(the simple calculated osteoporosis risk estimation,SCORE)^[19]。Cadarette等提出并验证了包括三个危险因素的骨质疏松危险评价工具的效果(the osteoporosis risk assessment instrument,ORAI)^[20]。许多学者分别在不同国家、不同种族、不同性别人群验证这些筛检工具的有效性和可靠性。那么这些筛检工具筛检效果怎样,哪些(种)筛检方法有效,目前在骨质疏松学术界还没有达成一致意见。采用循证医学系统评价方法对各类筛检工具效果进行综合评价是解决以上问题的关键。

骨质疏松症筛检工具的评价比较复杂,因为敏感度、特异度等指标既与截断值有关,又与T值及骨密度测量部位有关。采用何种截断值与T值,需要权衡疾病的严重性、筛检对象年龄、种族等人口特征、筛检效果、费用与收益等因素综合而定^[21-23]。

发表偏性分析一般采用盒式图(funnel plot)分析。但在纳入研究数量较少时,盒式图分析检出偏倚的效能会受到限制。由于本文纳入Meta分析的研究较少,因此没有进行发表偏性分析。期望今后通过增加纳入研究的数量进行发表偏性分析。

综上所述,危险评价工具OST/OSTA具有一定的筛检效果,能够识别大部分骨质疏松症患者,但有些研究其识别非患者的能力尚不高。为了进一步验证其信度与效度,需要在更多不同种族人群中进行验证。同时需要与其他种类评价工具的效果进行对比研究,如SCORE、ORAI,以确定那种评价工具效果更有效。今后随着相关研究的增多,期望通过增加纳入研究的数量进一步做Meta分析,从循证医学角度提出更有力的证据说明危险评价工具OST/OSTA的筛检效果,为骨质疏松症健康危险评价工具的研究提供参考依据。

参考文献

- [1] World Health Organization. Assessment of Fracture Risk and Application to Screening for Postmenopausal Osteoporosis [M]. Geneva: WHO Technical Report Series, 1994.
- [2] 王志群,施名,周重宛.南京市中老年妇女骨质疏松症患病情况的调查分析[J].中国老年学杂志,1997,17:88—89.
- [3] 郭立燕,周立锋,杨婕,等.中老年骨质疏松相关知识及患病状况调查[J].中国公共卫生,2005,4:470—472.
- [4] Whiting P, Rutjes AWS, Reitsma JB, et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews [J]. BMC Medical Research Methodology, 2003, 3:1—13.
- [5] Koh LK, Sedrine WB, Torralba TP, et al. Osteoporosis Self-Assessment Tool for Asians (OSTA) Research Group. A simple tool to identify Asian women at increased risk of osteoporosis [J]. Osteoporos Int, 2001,12(8):699—705.
- [6] Geusens P, Hochberg MC, van der Voort DJ, et al. Performance of risk indices for identifying low bone density in postmenopausal women[J]. Mayo Clin Proc, 2002,77(7):629—637.
- [7] Park HM, Sedrine WB, Reginster JY, et al. OSTA, Korean experience with the OSTA risk index for osteoporosis: a validation study[J]. J Clin Densitom, 2003,6(3):247—250.
- [8] Kung AW, Ho AY, Sedrine WB, et al. Comparison of a simple clinical risk index and quantitative bone ultrasound for identifying women at increased risk of osteoporosis [J]. Osteoporos Int, 2003,14(9):716—721.
- [9] Cadarette SM, McIsaac WJ, Hawker GA, et al. The validity of decision rules for selecting women with primary osteoporosis for bone mineral density testing[J]. Osteoporos Int, 2004, 15(5):361—366.
- [10] Richy F, Gourlay M, Ross PD, et al. Validation and comparative evaluation of the osteoporosis self-assessment tool (OST) in a Caucasian population from Belgium [J]. QJM, 2004,97(1):39—46.
- [11] Geater S, Leelawattana R, Geater A. Validation of the OSTA index for discriminating between high and low probability of femoral neck and lumbar spine osteoporosis among Thai postmenopausal women [J]. J Med Assoc Thai, 2004,87 (11): 1286—1292.
- [12] Li-Yu JT, Llamado LJ, Torralba TP. Validation of OSTA among Filipinos[J]. Osteoporos Int, 2005,16: 1789—1793.
- [13] Gourlay ML, Miller WC, Richy F, et al. Performance of osteoporosis risk assessment tools in postmenopausal women aged 45—64 years[J]. Osteoporos Int, 2005,16: 921—927.
- [14] 黄琪仁,章振林,周琦,等.上海市绝经后妇女低骨量的简易筛选方法的建立和验证[J].中华医学杂志,2007,87(12):808—811.
- [15] Center JR, Nguyen TV, Scheider D, et al. Mortality after all major types of osteoporotic fracture in men and women: an observational study[J]. Lancet, 1999,353:878—882.
- [16] Kung AW, Ho AY, Ross PD, et al. Development of a clinical assessment tool in identifying Asian men with low bone mineral density and comparison of its usefulness to quantitative bone ultrasound [J]. Osteoporos Int, 2005,16 (7): 849—855.
- [17] Lynn HS, Lau, EMC, Wong, SYS. An osteoporosis screening tool for Chinese men[J]. Osteoporos Int, 2005,16(7):829—834.
- [18] Adler RA, Tran MT, Petkov VI. Performance of the osteoporosis self-assessment screening tool for osteoporosis in American men[J]. Mayo Clin Proc, 2003,78(6):723—727.
- [19] Lydick E, Cook K, Turpin J, et al. Development and validation of a simple questionnaire to facilitate identification of women likely to have low bone density [J]. Am J Manag Care, 1998,4:37—48.
- [20] Cadarette SM, Jaglal SB, Kreiger N, et al. Development and validation of the Osteoporosis Risk Assessment Instrument to facilitate selection of women for bone densitometry [J]. CMAJ, 2000, 162(9):1289—1294.
- [21] 杨霖,杨永红,何成奇.骨质疏松症的康复评定[J].中国康复医学杂志,2006,21(12):1140—1142.
- [22] 叶超群,陈佑学,纪树荣.骨质疏松性骨折的危险因素及预测[J].中国康复医学杂志,2007,22(7):660—664.
- [23] 王凯夫,贾继峰,刘明辉.老年男性骨质疏松及雄激素受体的研究进展[J].中国康复医学杂志,2007,22(3):276—278.