

· 临床研究 ·

不同骨质疏松风险评估工具的筛检效果评价

李浩月¹ 荣爽^{1,2} 程静^{1,2} 杨晨露¹ 李本超¹ 李婷婷¹ 程光文^{1,2*}

1. 武汉科技大学医学院公共卫生学院营养与慢性病研究所,湖北 武汉 430065

2. 职业危害识别与控制湖北省重点实验室,湖北 武汉 430065

中图分类号: R592 文献标识码: A 文章编号: 1006-7108(2019)09-1307-05

摘要: 目的 通过对7种骨质疏松风险评估工具的对比,评价其社区与临床应用价值,为预防和筛查骨质疏松提供科学依据。**方法** 共纳入299名40周岁以上的中老年人,其中女性均已绝经,应用双能X线骨密度仪测量其骨密度值并诊断是否患有骨质疏松症,计算各个工具得分,比较变量、灵敏度、特异度和曲线下面积(area under the curve, AUC)。**结果** 绝经后妇女的各个工具得分均有统计学意义($P<0.05$),亚洲人骨质疏松自评工具(osteoporosis self-assessment tool for asian, OSTA)的灵敏度为96.3%,特异度为6.3%,AUC为0.710;骨质疏松风险评估工具(osteoporosis risk assessment instrument, ORAI)的灵敏度为90.1%,特异度为12.7%,AUC为0.661;骨质疏松风险简单评估(simple calculated osteoporosis risk estimation, SCORE)的灵敏度为25.9%,特异度为81.9%,AUC为0.686;骨质疏松危险指数(osteoporosis index of risk, OSIRIS)的灵敏度为90%,特异度为30.8%,AUC为0.734;骨质疏松预筛选风险评估(osteoporosis prescreening risk assessment, OPERA)的灵敏度为38.2%,特异度为84%,AUC为0.658;美国骨质疏松基金会快速诊断法(National Osteoporosis Foundation, NOF)的灵敏度为90.4%,特异度为26.6%,AUC为0.652。仅NOF筛检的中老年男性骨质疏松具有统计学意义($P<0.05$),NOF的灵敏度为93.5%,特异度为25.8%,AUC为0.697;而男性骨质疏松症风险评估(male osteoporosis risk estimation score, MORES)筛检的灵敏度为74.1%,特异度为29.0%,AUC为0.575,不具有统计学意义($P=0.190$)。**结论** OSTA的灵敏度最高,且仅有2个变量,使用最为简便,适合筛检大样本人群;OSIRIS灵敏度和特异度均较好,用于临床筛检更为准确;ORAI和NOF灵敏度较高,适合筛检阳性人群;SCORE和OPERA特异度较好,适合筛检阴性人群;NOF可用于筛检中老年男性骨质疏松。

关键词: 骨质疏松;风险评估;筛检;工具

Evaluation of screening value of different osteoporosis risk assessment tools

LI Haoyue¹, RONG Shuang^{1,2}, CHENG Jing^{1,2}, YANG Chenlu,¹, LI Benchao¹, LI Tingting¹, CHENG Guangwen^{1,2*}

1. Institute of Nutrition and Chronic Diseases, department of Nutrition and Food Hygiene, School of Public Health, Medical College, Wuhan University of Science and Technology, Wuhan 430065

2. Hubei Province Key Laboratory of Occupational Hazard Identification and Control, Wuhan 430065, China

* Corresponding author: CHENG Guangwen, Email: chenggw@wust.edu.cn

Abstract: Objective The community and clinical application value of seven kinds of osteoporosis risk assessment tools was evaluated to provide scientific basis for the prevention and screening of osteoporosis. **Methods** A total of 299 middle-aged and elderly people over 40 years old were included, among which all the women were postmenopausal, and their osteoporosis was diagnosed by dual-energy X-ray absorptiometry. The tools scores were calculated, and the variables, sensitivity, specificity and area under the curve (AUC) were compared. **Results** The scores of various tools of postmenopausal women were statistically significant ($P<0.05$), and the sensitivity of Osteoporosis Self-assessment Tool for Asian (OSTA) was 96.3%, the specificity was 6.3%, and the area under the curve (AUC) was 0.710. The sensitivity of Osteoporosis Risk Assessment Instrument (ORAI) was 90.1%, the specificity was 12.7%, and AUC was 0.661. The sensitivity of Simple Calculated Osteoporosis Risk Estimation (SCORE) was 25.9%, the specificity was 81.9%, and AUC was 0.686. The sensitivity of Osteoporosis Index of Risk (OSIRIS) was 90%, the specificity was 30.8%, and AUC was 0.734. The sensitivity of Osteoporosis Prescreening Risk Assessment (OPERA) was 38.2%, the specificity was 84%, and AUC was 0.658. The sensitivity of National Osteoporosis Foundation (NOF)

* 通信作者: 程光文,Email:chenggw@wust.edu.cn

was 90.4%, the specificity was 26.6%, and AUC was 0.652. Only NOF screening of elderly men showed statistical significance ($P < 0.05$). The sensitivity of NOF was 93.5%, the specificity was 25.8%, and AUC was 0.697. The sensitivity of MORES was 74.1%, the specificity was 29.0%, and AUC was 0.575, did not reach statistical significance ($P = 0.190$). **Conclusion** OSTA is the most convenient (only two variables) and sensitive for screening large sample populations. OSIRIS had good sensitivity and specificity, and was the most accurate in clinical screening. ORAI and NOF were highly sensitive and suitable for screening positive population. SCORE and OPERA had good specificity and were appropriate for screening negative population. NOF could be used for screening osteoporosis in elderly men.

Key words: osteoporosis; risk assessment; screen; tool

骨质疏松症是骨的脆性增高的一种全身性骨病,其最严重的后果是导致骨质疏松性骨折的发生^[1-2]。随着老龄化社会的到来,骨质疏松症已成为影响中老年人健康和家庭负担的一种重要疾病^[3-4]。全世界骨质疏松的患病率较高,该病在中国13.2%的患病率也居高不下,并随年龄的增大而升高^[5]。老年骨质疏松患者并发骨折将导致患者生命质量下降及家庭护理负担加重,也因此带来医疗支出的增加,所以对社区老年人进行骨质疏松的筛查,早发现、早干预对降低发病率和死亡率具有重要意义^[6]。

传统的骨质疏松诊断主要靠仪器测量,包括超声骨密度仪和双能X线骨密度仪,但仪器筛查成本高,操作复杂,不适合大规模的人群筛查。因此,很多研究者致力于快速筛查方法的研究,目前比较流行的快速筛查问卷就是其中的一种。目前,同类研究中应用的快速问卷种类较多,究竟哪一种问卷更适合中国居民的筛查却未见报道。本研究拟通过采用目前国际公认的骨质疏松诊断金标准-双能X线吸收法(dual energy x-ray absorptiometry, DEXA)的诊断结果,对应用较为广泛的亚洲人骨质疏松自评工具(osteoporosis self - assessment tool for asian, OSTA)、骨质疏松风险评估工具(osteoporosis risk assessment instrument, ORAI)、骨质疏松风险简单评估(simple calculated osteoporosis risk estimation, SCORE)、骨质疏松危险指数(osteoporosis index of risk, OSIRIS)、骨质疏松预筛选风险评估(osteoporosis prescreening risk assessment, OPERA)、美国骨质疏松基金会快速诊断法(National Osteoporosis Foundation, NOF)、男性骨质疏松症风险评估(male osteoporosis risk estimation score, MORES)7种问卷进行评估,以寻找最适合中国人群的骨质疏松筛查工具,为社区居民骨质健康管理提供科学的筛查方法。

1 材料和方法

1.1 研究对象

以2017年6月至2018年9月在华中科技大学附属协和医院核医学科接受DEXA检查的299例中老年人为受试对象,所有受试者均知情同意,收集其年龄、体重、既往病史、手术史、用药情况等资料。排除标准:①近6个月使用过类固醇类药物、抗抑郁药物以及其他对骨代谢造成影响的药物;②继发性骨质疏松症;③女性有子宫或者卵巢切除术史。

1.2 测量方法

1.2.1 体重和骨密度测量:使用身高体重仪测量体重。用双能X线骨密度仪测定所有受试者腰椎L₁₋₄以及双侧股骨颈骨密度,由专业医师操作以减少人为因素误差。

1.2.2 骨质疏松风险评估工具:由专人面对面指导受试者完成各个风险评估问卷,并计算其得分。OSTA、ORAI、SCORE、OSIRIS、OPERA、NOF以及MORES的变量和判断阈值见表1。其中ORAI、SCORE、OSIRIS、OPERA四种适用于女性,MORES适用于男性,OSTA、NOF则男性、女性皆可。

1.3 诊断标准

根据《原发性骨质疏松诊疗指南》(2017版)^[7]的诊断标准诊断骨质疏松。骨密度较正常人骨量峰值平均值低约1个标准差内,为正常;低1~2.5个标准差为骨量减少;低于2.5个标准差以上为骨质疏松。

1.4 统计学处理

使用SPSS 19.0统计学软件分析数据,计量资料用均数±标准差表示,两样本之间采用t检验。以DEXA的诊断结果为金标准,使用SPSS绘制OSTA、ORAI、SCORE、OSIRIS、OPERA、NOF及MORES在中老年男性和女性中骨质疏松的ROC曲线,并计算曲线下面积以及灵敏度、特异度、似然比、阳性预测值和阴性预测值。以 $\alpha=0.05$ 作为检验水准进行双侧检验, $P<0.05$ 为差异具有统计学意义。

表1 骨质疏松评估工具的变量和判断阈值

Table 1 Variables and thresholds of osteoporosis assessment tools

	OSTA	ORAI	SCORE	OSIRIS	OPERA	NOF	MORES
年龄	+	+	+	+	+	+	+
体重	+	+	+	+	+	+	+
雌激素		+	+	+			
绝经过早						+	
骨折史			+	+	+	+	
家族骨折史						+	
种族			+				
吸烟史						+	
慢性阻塞性肺病							+
类风湿性关节炎			+				
类固醇药物					+		
阈值	3	9	6	1	2	1	6

2 结果

2.1 人口学特征

研究人群中女性为175例,年龄41~88岁,绝经时间为1~37年,体重38~81公斤;男性为124例,年龄40~92岁,体重37~99公斤,详见表2。

表2 纳入人群的年龄和体重特征

Table 2 Age and weight characteristics of subjects

	女性/%	男性/%	
年龄/岁	40~ 50~ 60~ 70~ 80~	19(10.9) 81(46.3) 56(32.0) 15(8.6) 4(2.3)	30(24.2) 29(23.4) 45(36.3) 15(12.1) 5(4.0)
平均年龄/岁		60.0±8.5	
体重/kg	30~ 40~ 50~ 60~ 70~ 80~	1(0.6) 37(21.1) 91(52.0) 31(17.7) 14(8.0) 1(0.6)	1(0.8) 5(4.0) 22(17.7) 41(33.1) 34(27.4) 21(16.9)
平均体重/kg		57.2±8.4	

根据诊断标准将受试者分为非骨质疏松组

(187例)和骨质疏松组(112例)。*t*检验结果表明,女性两组间年龄和体重的差异有统计学意义,*P*<0.05;男性两组间的体重差异具有统计学意义,而年龄差异无统计学意义,*P*>0.05,见表3。

2.2 骨质疏松评估工具得分

关于骨质疏松组与非骨质疏松组OSTA、ORAI、SCORE、OSIRIS、OPERA、NOF、MORES得分的对比,通过独立样本*t*检验可知,OSTA、ORAI、SCORE、OSIRIS、OPERA和NOF在女性骨质疏松组和非骨质疏松组中都具有统计学意义(*P*<0.05);在男性的检验结果中只有MORES无统计学意义(*P*>0.05),而NOF和OSTA具有统计学意义。见表4。

2.3 筛查骨质疏松的灵敏度、特异度以及ROC曲线分析

评估OSTA、ORAI、SCORE、OSIRIS、OPERA、NOF及MORES对绝经后女性和中老年男性骨质疏松的筛查能力,在不同阈值的条件下,不同工具的灵敏度、特异度、似然比、阳性预测值、阴性预测值和曲线下面积如表5中所示。

表3 骨质疏松组与非骨质疏松组的年龄、体重比较

Table 3 Comparison of age and weight between osteoporosis group and non-osteoporosis group

分组	女性			男性		
	例数/n	年龄/岁	体重/kg	例数/n	年龄/岁	体重/kg
骨质疏松组	81	61.87±7.66	55.47±8.75	31	61.36±12.06	65.77±10.67
非骨质疏松组	94	58.33±8.93	58.72±7.73	93	59.36±10.52	71.26±11.58
<i>t</i> 值		2.792	2.613		0.883	2.326
<i>P</i> 值		0.006	0.010		0.379	0.022

各筛查工具对绝经后女性骨质疏松的筛查效果均有统计学意义(*P*<0.05),而三种筛查男性的工具

仅有NOF具有统计学意义,MORES和OSTA则无统计学意义(*P*>0.05)。NOF对男性的筛查灵敏度

为 93.5%, 但特异度是 25.8%, AUC 仅有 0.697。分析 ROC 曲线, 在女性人群中, OSTA、ORAI、OSIRIS 和 NOF 指数都能在推荐阈值下达到较高的灵敏度, 分别是 96.3%、90.1%、90.0% 和 90.4%。而 SCORE 和 OPERA 指数在推荐阈值下灵敏度较低, 分别是

25.9% 和 38.2%, 降低阈值才能达到较好的灵敏度。比较筛查女性骨质疏松不同工具的 AUC 可知, 最低的是 NOF, AUC 为 0.652, 最高的是 OSIRIS, AUC 达到了 0.734, 筛查能力最高。

表 4 骨质疏松组和非骨质疏松组的各评估工具得分对比

Table 4 Comparison of evaluation tools between osteoporosis group and non-osteoporosis group

组别		骨质疏松组	非骨质疏松组	t 值	P 值
女性	OSTA	-1.28±2.25	0.08±2.25	3.987	0.000
	ORAI	15.21±5.32	12.73±5.16	-3.121	0.002
	SCORE	0.22±7.54	-2.41±8.13	-2.204	0.029
	OSIRIS	-1.95±2.63	-0.18±2.29	4.753	0.000
	OPERA	1.33±0.85	0.87±0.77	-3.741	0.000
	NOF	1.48±0.88	0.99±0.78	-3.876	0.000
男性	MORES	7.03±2.55	6.22±2.97	-2.164	0.173
	OSTA	0.88±3.66	2.38±3.22	-3.672	0.032
	NOF	1.77±0.99	1.05±0.79	-1.370	0.001

表 5 骨质疏松筛选工具的 ROC 曲线分析

Table 5 ROC curve analysis of osteoporosis screening tools

ROC 指标	女性										男性		
	OSTA	ORAI	SCORE	OSIRIS	OPERA	NOF	MORES	OSTA	NOF				
阈值	≤3*	≤0	≥9*	≥12	≥6*	≥-4	≤1*	≤0	≥2*	≥1	≥1*	≥2	≥6*
SE	96.3	76.5	90.1	70.3	25.9	64.2	90.0	82.7	38.2	85.1	90.4	43.2	74.1
SP	6.38	57.4	12.7	56.3	81.9	71.2	30.8	51.0	84.0	31.9	26.6	78.7	29.0
+LR	1.03	1.80	1.03	1.61	1.43	2.24	1.30	1.69	2.40	1.25	1.23	2.03	1.05
-LR	0.58	0.41	0.77	0.53	0.90	0.50	0.32	0.34	0.73	0.46	0.37	0.72	0.89
PPV	47.2	60.8	47.1	58.2	55.3	65.8	52.9	59.3	67.4	51.9	51.4	63.6	25.8
NPV	66.7	74.0	60.0	68.8	56.2	69.8	78.4	77.4	61.2	71.4	75.8	61.7	77.1
AUC	0.710		0.661		0.686		0.734		0.658		0.652		0.575
P 值	0.000		0.000		0.000		0.000		0.000		0.000		0.190
													0.072
													0.000

注: * 为工具筛选骨质疏松的推荐得分阈值。

3 讨论

近年来, 骨质疏松及其导致的脆性骨折的发生率越来越高^[8], 其影响因素包括营养状况、遗传因素、内分泌因素以及生活方式等^[9]。由于早期骨量减少并没有明显的临床症状, 临幊上发现的骨质疏松患者大多是已出现脆性骨折。因此早期诊断骨质疏松或者骨量降低、预测发生的可能性对预防骨质疏松和骨折有着重要的意义。

OSTA 是一种有效的骨质疏松症早期筛查方法, 因为只需要年龄和体重两个变量, 具有快速、简单的优势, 在韩国、印度等亚洲国家应用 OSTA 筛查绝经后妇女骨质疏松高危人群有着良好的效果^[10-11]。本研究利用 OSTA 对于武汉地区中老年人的研究所得灵敏度高达 96.3%, 筛检阳性人群取得了良好的效果。而景彩霞等^[12]对延安地区体重偏高绝经期女性

的研究表明 OSTA 的灵敏度不高, 但是特异度却高达 94.6%, AUC 为 0.789, 与本研究结果相反; ORAI 则在 OSTA 的基础上多了一个是否使用雌激素的问题, 该工具在国外运用较多, 国内却很少有研究探讨, 其原因可能是国内极少有绝经期妇女使用雌激素替代疗法。研究^[13]表明行子宫和卵巢切除术的妇女, 其体内激素水平下降明显, 容易引起骨质疏松。ORAI 在推荐阈值下能达到较高的灵敏度, 如果更改阈值到 12, 其灵敏度和特异度都能同时保持较高水平。韩国的一个研究^[14]中 ORAI 的筛查阈值提高到 16, AUC 达到了 0.782, 灵敏度降低为 65.1%; OSIRIS 则有年龄、体重、雌激素和骨折史四个变量, 其灵敏度同 ORAI 一样, 如果降低其阈值, 提高特异度, 各方面指标都明显提高。陈平等^[15]在成都地区绝经期妇女中验证 OSIRIS 工具时, 降低了筛查阈值, 能让其灵敏度和特异度优于其他几种工具, 与本研究的结果一致;

NOF 则是除了年龄、体重和骨折史外,还需要提供家族骨折史和吸烟史,变量数增多,适用于男性和女性,其 AUC 在几种工具中最低;SCORE 和 OPERA 在推荐阈值下,其灵敏度只有 25.9% 和 38.2%,筛检效果并不理想。如果将 SCORE 的阈值降低到 -4 则能提高灵敏度到 64%,将 OPERA 的阈值降低到 1 也能提高其灵敏度到 85%。这两种工具在国内的研究极少,而且 SCORE 有一项是关于类风湿性关节炎,需要经过确诊才能评定,大大降低了其在社区人群中的筛查效率。

而对于可以筛查中老年男性骨质疏松的三种工具来说,只有 NOF 有统计学意义,而且灵敏度也可达到 93.5%,但是特异度仅有 25.8%,AUC 为 0.697,筛检能力不高;OSTA 和 MORES 这两种工具主要根据年龄和体重计算,可能因为男性两组间的年龄没有差异,导致两组得分不具有统计学意义。

当改变工具的得分诊断阈值时,灵敏度和特异度都相应改变,特异度的升高将带来灵敏度的降低。本研究发现,如果改变阈值提高特异度,则大多数骨质疏松筛查工具的灵敏度和特异度能同时保持在一个较高的水平。因而,根据工具在特定人群中的灵敏度和特异度比较,在不同环境下选择适当的诊断阈值,应考虑达到更高的灵敏度以筛选真正的阳性人群,还是达到更高的特异度来筛选真正的阴性人群,与王玉鹏^[16]等给出的建议相似。

综合比较曲线下面积、灵敏度和变量可知,OSIRIS 和 OSTA 用于预测绝经后女性骨质疏松的效果好,ORAI 和 NOF 的筛查效果次之。由此可见,除了体重和年龄,是否使用雌激素和是否具有骨折史是判断女性骨骼健康的重要辅助条件。然而国内绝经期妇女在临幊上多不愿意接受激素类治疗,问卷工具是否需要加入使用雌激素这一变量还需探讨。SCORE 和 OPERA 在不改变阈值的情况下,适用于筛查绝经后骨质疏松女性的阴性人群,但如果改变阈值提高灵敏度,却不如使用前几种工具简便。在男性群体中,NOF 的灵敏度较高,与 OSTA 的变量比较,提示家族骨折史、自身骨折史、吸烟史与男性的骨骼健康可能有关联。

本研究纳入的人群并无风湿性关节炎和使用过类固醇药物,因此 SCORE 和 OPERA 在推荐阈值下无法达到较高的灵敏度,还需要更多的研究去论证。由于样本量偏小和地域局限性,今后需要扩大样本量以提高研究的精确性。

综上,OSTA、OSIRIS、ORAI 和 NOF 用于筛查绝

经后妇女骨质疏松有着较高的灵敏度,SCORE 和 OPERA 有较高的特异度,仅 NOF 对筛查中老年男性骨质疏松有统计学意义,可为临床预防提供指导。对于社区大规模人群则可考虑利用最为简便的 OSTA 来筛查骨质疏松高危患者。

【参考文献】

- [1] Nanes MS, Kallen CB. Osteoporosis [J]. Seminars in Nuclear Medicine, 2014, 44(6): 439-450.
- [2] 何渝煦, 魏庆中, 熊启良, 等. 骨质疏松性骨折与骨密度关系的研究进展[J]. 中国骨质疏松杂志, 2014, 20(2): 219-224.
- [3] Fujiwara S, Zhao X, Teoh C, et al. Disease burden of fractures among patients with osteoporosis in Japan: health-related quality of life, work productivity and activity impairment, healthcare resource utilization, and economic costs[J]. Journal of Bone & Mineral Metabolism, 2019, 37(2): 307-318.
- [4] Rodrigues AM, Eusébio M, Santos MJ, et al. The burden and undertreatment of fragility fractures among senior women [J]. Archives of Osteoporosis, 2018, 13(1): 22.
- [5] 韩亚军, 帖小佳, 伊力哈木·托合提. 中国中老年人骨质疏松症患病率的 Meta 分析[J]. 中国组织工程研究, 2014, 18(7): 1129-1134.
- [6] Mitchell P, Harvey N, Dennison E, et al. IOF Compendium of Osteoporosis[Z]. International Osteoporosis Foundation, 2017: 32-36.
- [7] 中华医学会骨质疏松和骨矿盐疾病分会. 原发性骨质疏松症诊疗指南(2017)[J]. 中华骨质疏松和骨矿盐疾病杂志, 2017, 20(5): 413-443.
- [8] 黄雪珍, 张建东, 宦荣光, 等. 1 768 例中老年妇女绝经后骨质疏松患病率及危险因素调查分析[J]. 检验医学与临床, 2016, 13(13): 1841-1843.
- [9] 陈玉平, 蔡德鸿, 刘雪琴, 等. 社区中老年人骨质疏松症的患病危险因素调查[J]. 中国临床康复, 2005, 9(15): 156-157.
- [10] Oh SM, Nam BH, Rhee Y, et al. Development and validation of osteoporosis risk-assessment model for Korean postmenopausal women[J]. J Bone Miner Metab, 2013, 31(4): 423-432.
- [11] Satyaraddi A, Shetty S, Kapoor N, et al. Performance of risk assessment tools for predicting osteoporosis in south Indian rural elderly men[J]. Arch Osteoporos, 2017, 12(1): 35.
- [12] 景彩霞, 李二乐, 薛亚娟. OSTA 指数与体重指数对于绝经后妇女骨质疏松的预测效果评价[J]. 中国骨质疏松杂志, 2015, 21(9): 1083-1086.
- [13] 范文玲, 刘志奇, 王九焕, 等. 子宫切除术后更年期症状及骨质疏松情况观察[J]. 河北医药, 2012, 34(22): 3441-3442.
- [14] Yoo TK, Kim SK, Kim DW, et al. Osteoporosis risk prediction for bone mineral density assessment of postmenopausal women using machine learning [J]. Yonsei Med J, 2013, 54(6): 1321-1330.
- [15] 陈平, 黄际远, 孙勤, 等. 骨质疏松症危险评价工具在成都地区绝经后妇女中的效果评价[J]. 中国骨质疏松杂志, 2011, 17(2): 113-116.
- [16] 王玉鹏, 李宁, 张秋菊, 等. 三种骨质疏松症筛查方法的效果评价[J]. 中国卫生统计, 2012, 29(2): 193-195.

(收稿日期: 2018-05-29; 修回日期: 2018-10-15)