

Original article

A comparison of 6 osteoporosis risk assessment tools among postmenopausal women in Kuala Lumpur, Malaysia

Li Shean Toh ^{a,*}, Pauline Siew Mei Lai ^b, David Bin-Chia Wu ^c, Brian G. Bell ^d,
Cuu Phuong Linh Dang ^e, Bee Yean Low ^f, Kok Thong Wong ^f, Giuseppe Guglielmi ^g,
Claire Anderson ^a

^a Division of Pharmacy Practice and Policy, School of Pharmacy, University of Nottingham, University Park, Nottingham, UK

^b Department of Primary Care Medicine, University of Malaya Primary Care Research Group (UMPCRG), Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia

^c School of Pharmacy, Monash University Malaysia, Jalan Lagoon Selatan, Selangor, Malaysia

^d School of Medicine, University of Nottingham, University Park, Nottingham, UK

^e Division of Pharmacy, School of Medicine, Faculty of Health, University of Tasmania, Tasmania, Australia

^f School of Pharmacy, Faculty of Science and Engineering, University of Nottingham, Selangor, Malaysia

^g Department of Radiology, University of Foggia, Foggia, Italy

ARTICLE INFO

Article history:

Received 18 March 2019

Received in revised form

20 August 2019

Accepted 3 September 2019

Available online 11 September 2019

Keywords:

Osteoporosis

Screening

Osteoporosis self-assessment tool for asians

Risk assessment

Performance comparison

ABSTRACT

Objectives: To compare and assess the performance of 6 osteoporosis risk assessment tools for screening osteoporosis in Malaysian postmenopausal women.

Methods: Six osteoporosis risk assessments tools (the Simple Calculated Osteoporosis Risk Estimation [SCORE], the Osteoporosis Risk Assessment Instrument, the Age Bulk One or Never Estrogen, the body weight, the Malaysian Osteoporosis Screening Tool, and the Osteoporosis Self-Assessment Tool for Asians) were used to screen postmenopausal women who had not been previously diagnosed with osteoporosis/osteopenia. These women also underwent a dual-energy X-ray absorptiometry (DXA) scan to confirm the absence or presence of osteoporosis.

Results: A total of 164/224 participants were recruited (response rate, 73.2%), of which only 150/164 (91.5%) completed their DXA scan. Sixteen participants (10.7%) were found to have osteoporosis, whilst 65/150 (43.3%) were found to have osteopenia. Using precision-recall curves, the recall of the tools ranged from 0.50 to 1.00, whilst precision ranged from 0.04 to 0.14. The area under the curve (AUC) ranged from 0.027 to 0.161. The SCORE had the best balance between recall (1.00), precision (0.04–0.12), and AUC (0.072–0.161).

Conclusions: We found that the SCORE had the best balance between recall, precision, and AUC among the 6 screening tools that were compared among Malaysian postmenopausal women.

© 2019 The Korean Society of Osteoporosis. Publishing services by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Approximately, one-tenth of women aged 60 years, one-fifth of women aged 70, two-fifths of women aged 80 and two-thirds of women aged 90 will be affected by osteoporosis [1]. In 2005, the prevalence of osteoporosis was 24.1% in Malaysia [2]. Osteoporotic fracture is associated with increased mortality, morbidity and is a

huge economic burden [1,3]. Despite these negative outcomes, only approximately 20% of women who had an osteoporosis-related fracture received either a dual-energy X-ray absorptiometry (DXA) scan, or a prescription to treat osteoporosis [4,5].

Screening for osteoporosis among the general population using the DXA scan alone may not be feasible, as this test is expensive, and is only available in resource-rich hospitals. This has resulted in the development of several osteoporosis risk assessment tools such as the Simple Calculated Osteoporosis Risk Estimation (SCORE) [6], the Osteoporosis Risk Assessment Instrument (ORAI) [7], the Age Bulk One or Never Estrogen (ABONE) [8], the body weight (WEIGHT) [9], the Malaysian Osteoporosis Screening Tool (MOST)

* Corresponding author. Division of Pharmacy Practice and Policy, School of Pharmacy, University Park, NG7 2RD, Nottingham, United Kingdom.

E-mail address: lishean.toh@nottingham.ac.uk (L.S. Toh).

Peer review under responsibility of The Korean Society of Osteoporosis.

[10], and the Osteoporosis Self-Assessment Tool for Asians (OSTA) [11]. These tools were developed to identify potential patients who may be at high risk for osteoporosis, who would subsequently undergo a DXA scan. They do not replace the need for a DXA scan. These risk assessment tools have been validated in the United States [6,12], Canada [7,13], Belgium [14], the Netherlands [12], Japan [15], Korea [16], Thailand [17], Hong Kong [18], and Singapore [19]. In Malaysia, only the MOST and the OSTA have been validated [10,11,20]. The MOST was validated among a small sample of 72 women, aged 45 years and above [10], whilst the OSTA was validated among 860 women aged 45 years and above [20]. Furthermore, a Singaporean study validated the SCORE, ORAI, ABONE, WEIGHT, and OSTA. They concluded that the OSTA was the most effective tool for the Chinese elderly in Singapore [19]. Although the ethnicity of the population in Singapore and Malaysia are similar, a tool needs to be validated in the country that it is to be administered in. To date, there is no consensus on which osteoporosis risk assessment tool is the most suitable to screen for osteoporosis in Malaysian postmenopausal women. Therefore, this study aimed to compare and assess the performance of 6 risk assessment tools for use among postmenopausal women in Malaysia.

2. Methods

2.1. Study design

This performance comparison study was conducted from October 2013 to January 2014, at a hospital-based primary care clinic in Kuala Lumpur, Malaysia.

2.2. Participants

Postmenopausal women, aged 50 years and above, who have not been diagnosed with osteoporosis or osteopenia were included. Participants with a history of metabolic disease, presence of bone metastasis, significant renal impairment, previous bilateral oophorectomy, previous hip fracture or prior use of any bisphosphonates were excluded from the study.

The sample size required for this study was 92 participants (confidence interval = 95%). Calculations were based on the prevalence of osteoporosis as 24.1% [2], among postmenopausal women in Malaysia [21,22]. Adjusting for a 20% loss to follow up, the minimum number of participants required was 110 participants.

2.3. Outcomes measured

The primary outcome was to compare the recall and precision of six risk assessment tools. The secondary outcome was the area under the curve (AUC).

2.4. Procedure

Participants were recruited while they were waiting to see their doctor using a 1:2 systematic random sampling procedure. The medical folders of eligible participants were labelled from 1 to 40, and a number was randomly drawn from a bag to determine the starting number at the start of each day of recruitment. Subsequently, every 2nd medical folder was selected for recruitment. This study was approved by the University Malaya Medical Centre Ethics Committee (approval number: 920.27). At baseline, participants' demographics were obtained and questions from the 6 risk assessment tools were administered. Participants' weight, height and hip circumference were measured using a digital weighing machine, a mechanical height scale (Seca, 799 electronic column

scales class III and 220 telescopic measuring rod, Hamburg, Germany) and measuring tape, respectively. The characteristics and published cutoff points of the 6 osteoporosis risk assessment tools are summarised in Table 1.

Subsequently, a DXA scan appointment was arranged to confirm the presence or absence of osteoporosis. The dual energy DXA machine was used to measure the left femoral neck and lumbar spine (L1–4) bone mineral density (BMD) (IDXA, GE Lunar, Milwaukee, WI, USA). The mean and precision error for the femoral neck and the lumbar spine (L1–4) was 0.936 (1.4% coefficient of variation [CV]) and 1.184 (1.1% CV), respectively. All BMD measurements were performed by a radiographer and confirmed by a radiologist. The femoral neck T-score BMD results were used as a comparison to the risk assessment tools in this study. The tools were also analysed using the combination of either a T-score ≤ -2.5 at the femoral neck or lumbar spine.

Although no single level of BMD should be used as the sole basis for treatment, the femoral neck T-score is the most reliable measure for predicting hip fracture risk [23,24]. In addition, participants were also diagnosed as having osteoporosis if the T-score at any given site was ≤ -2.5 .

Participants were then informed of their BMD results via telephone and were sent a copy of their results via mail. Participants with a T-score ≤ -2.5 were advised to visit their doctor.

2.5. Data analysis

All data were entered into IBM SPSS Statistics ver. 20.0 (IBM Co., Armonk, NY, USA). Normality was assessed using the Kolmogorov-Smirnov test.

Since data was not normally distributed, medians and interquartile ranges were reported. The recall-precision results were generated using Stata ver. 15 (StataCorp LP, College Station, TX, USA).

2.6. The recall and precision of a tool

A tool with better performance demonstrates a higher value of recall and precision. The closer the precision and recall are to 1 the better the performance of the tool. The formulas used to calculate the recall and precision are listed below [25]:

Recall = [Number of women with osteoporosis (T-score ≤ -2.5 at the femoral neck)] / [Number of osteoporotic women incorrectly identified as non-osteoporosis using the risk assessment + Number of women with osteoporosis (T-score ≤ -2.5 at the femoral neck)]

= true positives / (false negatives + true positives)

Precision = [Number of women with osteoporosis (T-score ≤ -2.5 at the femoral neck)] / [Number of nonosteoporotic women incorrectly identified as osteoporosis using the risk assessment + Number of women with osteoporosis (T-score ≤ -2.5 at the femoral neck)]

= true positives / (false positives + true positives)

2.7. Area under the curve

The AUC generated from the precision-recall curve was used to compare the diagnostic performance of the osteoporosis risk assessment tools and the DXA scan. A realistic risk assessment tool should not have an AUC < 0.500 [26].

Table 1

Characteristics of the 6 osteoporosis risk assessment tools, its published cutoff points and its scoring system.

Tool	Developed in	Validated in	No. of questions	Published cutoff point indicating an increased risk for osteoporosis	Scoring system
Simple Calculated Osteoporosis Risk Estimation	USA [6]	USA [6,12], Belgium [14], Netherlands [12], Japan [15], Singapore [19]	6	Score of ≥ 6	Simple additive scoring system: <u>Race</u> : 5 if not black <u>Rheumatoid arthritis</u> : 4 if yes <u>History of minimal trauma fracture after age 45 years</u> : 4 for each fracture of the wrist, hip, or ribs (12 points maximum) <u>Age</u> : 3 x first digit of age in years by 3 <u>Estrogen therapy</u> : 1 if No <u>Weight</u> : 1 multiplied with the weight in pounds (lb) divided by 10 and truncated to an integer
Osteoporosis Risk Assessment Instrument	Canada [7]	USA [12], Canada [7], Belgium [14], Netherlands [12], Japan [15], Singapore [19]	3	Score ≥ 9	Simple additive scoring system: <u>Age</u> : 15 if ≥ 75 years 9 if 65–74 years 5 if 55–64 years <u>Weight</u> : 9 if < 60 kg 3 if < 60.0 –69.9 kg <u>Estrogen use</u> : 2 if not currently taking
Age Bulk One or Never Estrogen	USA [8]	Singapore [19]	3	Score ≥ 2	Simple additive scoring system: <u>Age</u> : 1 if > 65 years <u>Weight</u> : 1 if < 63.5 kg <u>Oestrogen use</u> : 1 never used oral contraceptive or oestrogen therapy for at least 6 months
Body weight Malaysian Osteoporosis Screening Tool	Sweden [9] Malaysia [10]	Singapore [19] Malaysia [10,28]	1 4	Weight of ≤ 70 kg Score ≥ 4	Simple additive scoring system: <u>Age</u> : 20 if > 61 years 6 if 56–60 years 2 if 51–55 years 0 if < 50 years <u>Years post menopause</u> : 22 if > 10 years 6 if 6–10 years 4 if 1–5 years 0 if 0 years <u>BMI</u> : 4 if < 19 kg/m ² 2 if 19–24 kg/m ² 0 if > 24 kg/m ² <u>Hip circumference</u> : 2 if < 90 cm 0 if > 90 cm
Malaysian Osteoporosis Screening Tool	Malaysia [10]	Malaysia [10,28]	4	Score ≥ 4	Simple additive scoring system: <u>Age</u> : 20 if > 61 years 6 if 56–60 years 2 if 51–55 years 0 if < 50 years <u>Years post menopause</u> : 22 if > 10 years 6 if 6–10 years 4 if 1–5 years 0 if 0 years <u>BMI</u> : 4 if < 19 kg/m ² 2 if 19–24 kg/m ² 0 if > 24 kg/m ² <u>Hip circumference</u> : 2 if < 90 cm 0 if > 90 cm
Osteoporosis Self-Assessment Tool for Asians	Eight Asian countries [11]	USA [12], Canada [13], Belgium [14], Netherlands [12], Japan [15], Korea [16], Thailand [17], Hong Kong [18], Singapore [19]	2	Score ≤ -1	<u>Simple calculation</u> : $0.2 \times (\text{body weight [kg]} - \text{age [yr]})$

BMI, body mass index.

Original published cutoff points were defined as cutoff points used by the developers of the risk assessment tools.

3. Results

A total of 164/224 participants were recruited (response rate, 73.2%) (Fig. 1). A hundred fifty participants (91.5%) completed their DXA scan as 14 were lost to follow up. Sixteen participants (10.7%) were found to be osteoporotic based on either the femoral neck or lumbar spine (L1–4) BMD, whilst only 6 (4.0%) were found to be osteoporotic based on femoral neck only. Our study was of mixed ethnicity (Malay [8.0%], Chinese [72.0%], Indian [18.7%], and Eurasian [1.3%]) (Table 2).

The recall and the precision of the six risk assessment tools are presented in Table 3. Based on the femoral neck T-score, the SCORE, ORAI, WEIGHT, and MOST had a maximum recall of 1.00, but had low precision (range, 0.04–0.05). Overall, the recall for all tools were high (range, 0.83–1.00) except the OSTA with the lowest recall value of 0.50. However, the precision for all the tools were low ranging from 0.04 to 0.05. The SCORE had the highest AUC (0.072), whilst the WEIGHT had the lowest AUC (0.027).

Based on either the femoral neck or the lumbar spine (L1–4) T-score, the SCORE, WEIGHT, and MOST achieved a maximum recall of 1.00, but had low precision (range, 0.11–0.12). The OSTA had the lowest recall values (0.69). Overall, the precision for all tools was very low (range, 0.11–0.14). The SCORE had the highest AUC (0.161) whilst the OSTA had the lowest AUC (0.078) (Table 3).

4. Discussion

Our study found that all 6 risk assessment tools performed equally with a high recall ranging from 0.83 to 1.00 except for the OSTA which had lower recalls of 0.50–0.69. However, all tools had poor precision ranging from 0.04 to 0.12. The SCORE had the best balance between recall (1.00), precision (0.04–0.12), and AUC (0.072–0.161).

Most studies compared the risk assessment tools by using the receiver operating curve to generate sensitivity and specificity values. As the event of rate for osteoporosis in our study was low, the precision-recall curve was a more appropriate method of analysis [25]. Our findings were compared to the other studies by comparing the sensitivity (true positives/[true positives + false negatives]) obtained from these studies with the recalls from our study. Based on the published cutoff points and T-score ≤ -2.5 , our results were similar to the original studies [6,7,9–11]. In our study, the recall was high for all the tools (range, 0.83–1.00) except for OSTA when based on the T-score of the femoral neck. The SCORE, ORAI, WEIGHT, and MOST had a sensitivity of 91%, 94%, 94%, and 80% respectively in the original study [6,7,9,10]. The OSTA had the lowest recall of 0.50, while the original study obtained a sensitivity of 91% [11]. The difference in findings could be due to the type of analysis used, and the study design utilised to compare these tools. However, when the performance comparison of the six risk assessment tools was based on the T-score of the femoral neck or

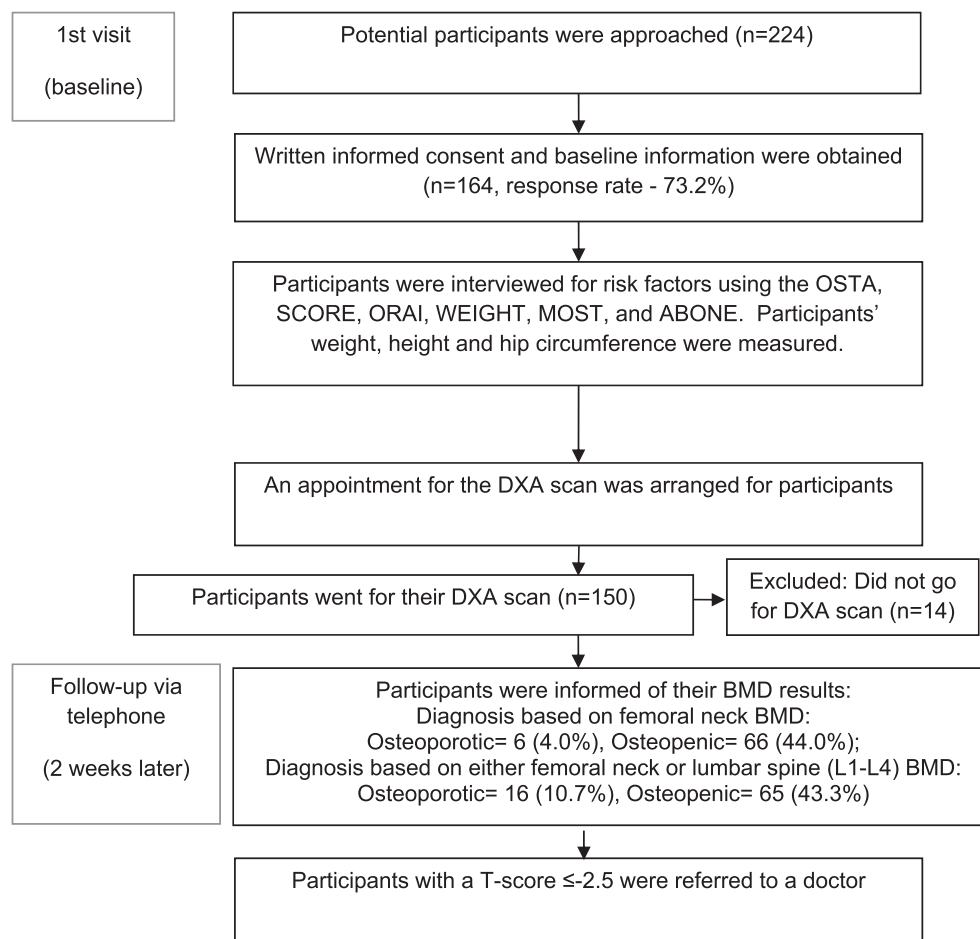


Fig. 1. Flowchart of the performance comparison of the 6 osteoporosis risk assessment tools. OSTA, Osteoporosis Self-Assessment Tool for Asians; SCORE, Simple Calculated Osteoporosis Risk Estimation; ORAI, Osteoporosis Risk Assessment Instrument; WEIGHT, body weight; ABONE, Age Bulk One or Never Estrogen; DXA, dual-energy X-ray absorptiometry; BMD, bone mineral density; MOST, Malaysian Osteoporosis Screening Tool.

Table 2
Demographics characteristics (n = 150).

Characteristic	Value
Age, yr	
Mean $\pm$ SD	62.0 $\pm$ 7.0
Median (IQR)	62.0 (11.0)
Ethnicity, n (%)	
Malay	12 (8.0)
Chinese	108 (72.0)
Indian	28 (18.7)
Eurasian	2 (1.3)
Weight, kg	
Mean $\pm$ SD	57.9 $\pm$ 10.0
Median (IQR)	55.8 (14.5)
Hip circumference, cm	
Mean $\pm$ SD	93.4 $\pm$ 6.8
Median (IQR)	94.0 (7.6)
BMI, kg/m ²	
Mean $\pm$ SD	23.8 $\pm$ 3.8
Median (IQR)	23.0 (5.4)
BMI, kg/m ² , n (%)	
<18.5 (underweight)	5 (3.3)
18.5–24.9 (normal)	97 (64.7)
25.0–29.9 (overweight)	37 (24.7)
$\geq$ 30.0 (obese)	11 (7.3)
Household income per month, n (%)	
<RM1000 (<\$310.7)	29 (19.3)
RM1000–1999 (\$310.7–621.0)	12 (8.0)
RM2000–2999 (\$621.3–931.7)	15 (10.0)
RM3000–3999 (\$932.0–1242.3)	17 (11.3)
RM4000–4999 (\$1242.6–1553)	12 (8.0)
>RM5000 (>\$1553.3)	65 (43.3)
Level of education, n (%)	
Never been to school	1 (0.7)
Primary (6 years of education)	12 (8.0)
Secondary (11–13 years of education)	47 (31.3)
Diploma/Technical school training (12–14 years of education)	31 (20.7)
Tertiary/Postgraduate (>15 years of education)	59 (39.3)
Bone mineral density, g/cm ²	
Femoral neck	
Mean $\pm$ SD	0.78 $\pm$ 1.78
Median (IQR)	0.80 (0.146)
Lumbar spine L1–4	
Mean $\pm$ SD	1.05 $\pm$ 0.15
Median (IQR)	1.05 (0.171)
T-score	
Femoral neck	
Mean $\pm$ SD	–0.87 $\pm$ 0.93
Median (IQR)	–0.95 (1.10)
Lumbar spine L1–4	
Mean $\pm$ SD	–0.53 $\pm$ 1.28
Median (IQR)	–0.50 (1.40)

SD, standard deviation; IQR, interquartile range; BMI, body mass index; RM, Ringgit Malaysia; BMD, bone mineral density.

the lumbar spine, all the tools had a high recall including the OSTA ranging from 0.69 to 1.00.

In our study, all the tools performed poorly in terms of precision. We identified the SCORE as the most appropriate tool as it had the best balance of recall, precision and the largest AUC when compared to the T-score of the femoral only and of the femoral neck and the lumbar spine. A high recall (sensitivity) is crucial since the primary purpose of a risk assessment tool is to identify patients who are at a higher risk of osteoporosis, who are then referred for a DXA scan. Whereas, low precision means that the tool is likely to over identify women at high risk for osteoporosis. However, this may be acceptable as the tool causes no harm [11]. The AUC yielded by the SCORE was the highest (0.072–0.161).

One of the methods to optimize the tools is to identify the tools' empirical optimum cutoff points. The precision-recall curves and the precision-threshold curves were generated to determine the tools' empirical optimal cutoff points as well as the precision and

recall of each of the six risk assessment tools. However, the precision of all the tools was very low (0.04–0.14), generating empirical cutoff points would only render a slight improvement of the precision and a substantial reduction of the recall. This means that we were unable to improve the performance of the tools with the generated curves and empirical cutoff points.

We noticed that the performance of the SCORE and the ORAI as risk assessment tools in our study was very poor compared to the original studies. This may be due to the fact that the cohort in our study was of mixed ethnicity (Malay [8.0%], Chinese [72.0%], Indian [18.7%], and Eurasian [1.3%]), whereas these tools were mainly developed for the Caucasian population [6,7]. A study of 135 Chinese, postmenopausal women noted higher cutoff points for the SCORE (≥ 8) and the ORAI (≥ 20) [19]. However, in our study, we were unable to provide useful empirical cutoff points. The difference in performance in our study for the OSTA and the MOST may also be explained by demographic differences in the population

Table 3
Results of the 6 risk assessment tools using published cutoff points based on femoral neck bone mineral density T-score ≤ -2.5 , as well as either the femoral neck or the lumbar spine (L1–4) T-score ≤ -2.5

Risk assessment tools	Femoral neck					Femoral neck or lumbar spine (L1–4)						
	T-score		Recall	Precision	AUC	Total	T-score		Recall	Precision	AUC	Total
	≤-2.5	>-2.5					≤-2.5	>-2.5				
SCORE			1.00	0.04	0.072				1.00	0.12	0.161	
High risk ≥ 6	6	133				139	16	123				139
Low risk <6	0	11				11	0	11				11
ORAI			1.00	0.05	0.047				0.94	0.12	0.129	
High risk ≥ 9	6	116				122	15	107				122
Low risk <9	0	28				28	1	27				28
ABONE			0.83	0.04	0.034				0.88	0.13	0.088	
High risk ≥ 2	5	105				110	14	96				110
Low risk <2	1	39				40	2	38				40
WEIGHT			1.00	0.05	0.027				1.00	0.12	0.082	
High risk ≤70 kg	6	127				133	16	117				133
Low risk >70 kg	0	17				17	0	17				17
MOST			1.00	0.04	0.034				1.00	0.11	0.105	
High risk ≥ 4	6	141				147	16	131				147
Low risk < 4	0	3				3	0	3				3
OSTA			0.50	0.04	0.030				0.69	0.14	0.078	
High risk ≤ −1	3	73				76	11	65				76
Low risk > -1	3	71				74	5	69				74

SCORE, Simple Calculated Osteoporosis Risk Estimation; ORAI, Osteoporosis Risk Assessment Instrument; ABONE, Age Bulk One or Never Oestrogen; WEIGHT, body weight; MOST, Malaysian Osteoporosis Screening Tool; OSTA, Osteoporosis Self-Assessment Tool for Asians; BMD, bone mineral density; AUC, area under the curve.

studied. The cohort in our study was of women ≥ 50 years of age, whereas the original OSTA and the MOST study included younger women who were ≥ 45 years. In Singapore and Hong Kong, which validated various risk assessment tools, suggested that the OSTA would be the most practical and accurate tool for daily use when assessing osteoporosis risk based on femoral neck T-score [18,19]. As for the WEIGHT, Malaysian women have a lower mean body weight of 58.4 kg and the published cutoff of 70 kg may be unsuitable [27]. This demonstrates that the tools' optimal cutoff points may vary with different age and ethnic groups.

The proportion of participants diagnosed with osteoporosis in our study was low, 10.7% based on either the femoral neck or lumbar spine (L1–4) BMD, and 4.0% based on the femoral neck only. A previous study conducted in Malaysia noted a higher prevalence of 24.1% [2]. This may be due to the fact that our participants were recruited whilst waiting to see their doctor at a primary care clinic, and thus may be more “health seeking” than participants from a previous study who were recruited from the community via flyers [2].

A limitation of our study was that data was collected based on patient self-report, and thus may be subject to recall bias. Secondly, a majority of the participants recruited were Chinese (72.0%), followed by Indians (18.7%) and Malays (8.0%). Although this ethnic group distribution does not represent the Malaysian population, it is representative of patients who sought treatment for osteoporosis in primary care clinics, where the majority are Chinese and middle class in terms of socioeconomic group, which may reflect the difference in the prevalence of osteoporosis across the three ethnic groups [2]. Nonetheless, this study still adds to the knowledge of a suitable risk assessment tool for Malaysian Chinese. Further studies should involve other age groups and ethnicity such as the Malays and Indians.

5. Conclusions

This study compared the performance of the various osteoporosis risk assessment tools, namely the SCORE, ABONE, ORAI, WEIGHT, MOST, and OSTA were found to be unsuitable to identify Malaysian postmenopausal women with osteoporosis. Our study found that the SCORE, despite the low precision and AUC, had the best balance between recall, precision and AUC of the 6 osteoporosis risk assessment tools compared among Malaysian postmenopausal women.

Conflicts of interest

No potential conflict of interest relevant to this article was reported.

Acknowledgments

Ministry of Science, Technology and Innovation (MOSTI) fund (06-02-12-SF0183). **ORCID.** Li Shean Toh: 0000-0002-9228-6266. Pauline Siew Mei Lai: 0000-0002-9940-9644. Brian G. Bell: 0000-0001-8577-7430. Bee Yean Low: 0000-0002-1841-0169. Giuseppe Guglielmi: 0000-0002-4325-8330. Claire Anderson: 0000-0002-5406-2296.

References

- [1] Kanis JA. Assessment of osteoporosis at the primary health care level. WHO Scientific Group Technical Report. Geneva (Switzerland): World Health Organization; 2007. p. 66.
- [2] Lim PS, Ong FB, Adeeb N, Seri SS, Noor-Aini MY, Shamsuddin K, et al. Bone health in urban midlife Malaysian women: risk factors and prevention. *Osteoporos Int* 2005;16:2069–79.

- [3] Johnell O, Kanis JA. An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. *Osteoporos Int* 2006;17:1726–33.
- [4] Hajcsar EE, Hawker G, Bogoch ER. Investigation and treatment of osteoporosis in patients with fragility fractures. *CMAJ (Can Med Assoc J)* 2000;163:819–22.
- [5] Cosman F, de Beur SJ, LeBoff MS, Lewiecki EM, Tanner B, Randall S, et al. Clinician's guide to prevention and treatment of osteoporosis. *Osteoporos Int* 2014;25:2359–81.
- [6] Lydick E, Cook K, Turpin J, Melton M, Stine R, Byrnes C. Development and validation of a simple questionnaire to facilitate identification of women likely to have low bone density. *Am J Manag Care* 1998;4:37–48.
- [7] Cadarette SM, Jaglal SB, Kreiger N, McIsaac WJ, Darlington GA, Tu JV. Development and validation of the Osteoporosis Risk Assessment Instrument to facilitate selection of women for bone densitometry. *CMAJ (Can Med Assoc J)* 2000;162:1289–94.
- [8] Weinstein L, Ullery B. Identification of at-risk women for osteoporosis screening. *Am J Obstet Gynecol* 2000;183:547–9.
- [9] Michaëlsson K, Bergström R, Mallmin H, Holmberg L, Wolk A, Ljunghall S. Screening for osteopenia and osteoporosis: selection by body composition. *Osteoporos Int* 1996;6:120–6.
- [10] Shan LP, Bee OF, Suniza SS, Adeeb N. Developing a Malaysian Osteoporosis Screening Tool (MOST) for early osteoporosis detection in Malaysian women. *Sex Reprod Healthc* 2011;2:77–82.
- [11] Koh LK, Sedrine WB, Torralba TP, Kung A, Fujiwara S, Chan SP, et al. A simple tool to identify asian women at increased risk of osteoporosis. *Osteoporos Int* 2001;12:699–705.
- [12] Geusens P, Hochberg MC, van der Voort DJ, Pols H, van der Klift M, Siris E, et al. Performance of risk indices for identifying low bone density in postmenopausal women. *Mayo Clin Proc* 2002;77:629–37.
- [13] Cadarette SM, McIsaac WJ, Hawker GA, Jaakkimainen L, Culbert A, Zarifa G, et al. The validity of decision rules for selecting women with primary osteoporosis for bone mineral density testing. *Osteoporos Int* 2004;15:361–6.
- [14] Gourlay ML, Miller WC, Richy F, Garrett JM, Hanson LC, Reginster JY. Performance of osteoporosis risk assessment tools in postmenopausal women aged 45–64 years. *Osteoporos Int* 2005;16:921–7.
- [15] Fujiwara S, Masunari N, Suzuki G, Ross PD. Performance of osteoporosis risk indices in a Japanese population. *Curr Ther Res* 2001;62:586–94.
- [16] Park HJ, Park KH, Park GM, Paek YJ, Cho JJ. Evaluation of simple tool as a screening test for osteoporosis and osteopenia in Korean postmenopausal women. *J Korean Acad Fam Med* 2003;24:702–8.
- [17] Saetung S, Ongphiphadhanakul B, Rajatanavin R. The relationship of an Asian-specific screening tool for osteoporosis to vertebral deformity and osteoporosis. *J Bone Miner Metab* 2008;26:47–52.
- [18] Kung AW, Ho AY, Sedrine WB, Reginster JY, Ross PD. Comparison of a simple clinical risk index and quantitative bone ultrasound for identifying women at increased risk of osteoporosis. *Osteoporos Int* 2003;14:716–21.
- [19] Chan SP, Teo CC, Ng SA, Goh N, Tan C, Deurenberg-Yap M. Validation of various osteoporosis risk indices in elderly Chinese females in Singapore. *Osteoporos Int* 2006;17:1182–8.
- [20] Chan S, Yeap S, Hew F, Chee S, Zaitun Y. The osteoporosis self-assessment tool for Asians (OSTA): validation in Malaysia. New York: Elsevier Science Inc.; 2003. p. S176.
- [21] Kish L. Survey sampling. New York: John Wiley and Sons; 1965.
- [22] Kish L. Sampling organizations and groups of unequal sizes. *Am Sociol Rev* 1965;30:564–72.
- [23] Johnell O, Kanis JA, Oden A, Johansson H, De Laet C, Delmas P, et al. Predictive value of BMD for hip and other fractures. *J Bone Miner Res* 2005;20:1185–94.
- [24] Marshall D, Johnell O, Wedel H. Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. *BMJ* 1996;312:1254–9.
- [25] Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS One* 2015;10:e0118432.
- [26] Fawcett T. An introduction to ROC analysis. *Pattern Recognit Lett* 2006;27:861–74.
- [27] Azmi Jr MY, Junidah R, Siti Mariam A, Safiah MY, Fatimah S, Norimah AK, et al. Body mass index (BMI) of adults: findings of the Malaysian Adult Nutrition Survey (MANS). *Malays J Nutr* 2009;15:97–119.
- [28] Muslim D, Mohd E, Sallehudin A, Tengku Muzaffar T, Ezane A. Performance of osteoporosis self-assessment tool for asian (OSTA) for primary osteoporosis in post-menopausal Malay women. *Malays Orthop J* 2012;6:35–9.