J Musculoskelet Neuronal Interact 2011; 11(3):249-256

Original Article

Hylonome

The accuracy of clinical kyphosis examination for detection of thoracic vertebral fractures: comparison of direct and indirect kyphosis measures

K. Siminoski[1], R.S. Warshawski[2], H. Jen[2], K-C Lee[2]

[1]Department of Radiology and Diagnostic Imaging, and Division of Endocrinology and Metabolism, Department of Medicine, University of Alberta, Edmonton, Alberta, Canada; [2]Department of Radiology and Diagnostic Imaging, University of Alberta, Edmonton, Alberta, Canada

Abstract

Objective: To compare the accuracies of two simple physical examination maneuvers for detecting the presence of thoracic vertebral fractures (VF) diagnosed by radiography: direct measurement of kyphosis angle (KA, in degrees) and indirect measurement using wall-occiput distance (WOD, in cm). **Methods:** Subjects were 280 women (average age, 54.5 years; range, 18-92) referred for assessment of osteoporosis. KA was measured from T4 to T12 using a digital inclinometer while WOD was measured with the patient in a standardized position. VF were diagnosed on radiographs using semi-quantitative morphometry. **Results:** KA and WOD were moderately correlated ($r=0.72, p<10^{-11}$). KA increased by 3.7° (95% CI, 2.6-4.8°) for each VF ($p=4x10^{-11}$) and WOD rose 1.3 cm (95% CI, 0.8-1.7 cm) per VF ($p=2x10^{-11}$). The areas under the receiver operating characteristic curves were 0.72 (95% CI, 0.65-0.79) for KA and 0.76 (95% CI, 0.69-0.82) for WOD, which were not significantly different (p=0.13). **Conclusions:** Given similar performances of direct and indirect measures of kyphosis, we propose that WOD should be used in clinical practice, with a clinical threshold of WOD>4.0 cm as an indication to consider spine radiography. At this WOD threshold, sensitivity was 41% (95% CI, 31-52%) and specificity was 92% (95% CI, 87-95%). WOD should be considered for use in the clinical assessment of osteoporosis patients.

Keywords: Hyperkyphosis, Kyphosis, Osteoporosis, Thoracic Spine, Vertebral Fracture

Introduction

Dear Aunt Dode, all bent over like a comma.
John Updike[1]

His manner was brisk, and yet his general appearance gave an undue impression of age, for he had a slight forward stoop and a little bend of the knees as he walked.
Sir Arthur Conan Doyle[2]

Osteoporotic vertebral fractures (VF) can be detected through physical examination[3-6]. Loss of vertebral body height causes loss of total body stature, which can be detected by evaluating height loss[5-8]. Historical height loss (HHL; height loss from the tallest recalled height) can detect prevalent VF[5,7,8] and prospective height loss (height loss observed during a period of monitoring) can detect incident VF[6,9]. Although height loss identifies the presence of VF, it is a global measure and cannot localize them to a particular location within the spine. Specialized physical examination maneuvers may be capable of localizing VF, at least to the extent of determining whether they are likely to be in the lumbar or thoracic spine regions. In the lumbar spine, for example, loss of vertebral body height due to fracture moves the rib cage closer to the pelvic brim, resulting in a measurable narrowing of the rib-pelvis distance (RPD)[4,7,10]. Consequently, decreased RPD on physical examination can assist in pinpointing VF to the lumbar spine. Limited data is available, however, on the clinical application of physical examination maneuvers that can localize fractures to the thoracic spine[11-13].

The most common morphology of VF is the anterior wedge compression deformity[5,6,8]. Within the thoracic region, this wedging leads to increased anterior curvature, or hyperkyphosis[11-18]. The extent of kyphosis can be assessed directly by

The authors have no conflict of interest.

Corresponding author: Dr. Kerry Siminoski, 6628-123 Street, Edmonton, AB, Canada T6H 3T6.
E-mail: kerrygs@telusplanet.net

Edited by: F. Rauch
Accepted 12 July 2011

measuring the angle between two points on the thoracic spine, referred to here as the kyphosis angle (KA)[12,13,15-17,19]. It can also be assessed indirectly by measuring anterior displacement of the head as a result of the augmented kyphosis when the body is positioned to unmask the curvature, referred to here as the wall-occiput distance (WOD)[3,7,14,20-22]. When changes in spinal shape are advanced, as in Aunt Dode, the structural abnormality is apparent even to untrained observers[1]. There is a great deal of data in the literature linking hyperkyphosis to vertebral fractures. There is much more limited information to tell us whether physical examination can specifically detect and localize thoracic VF, with data in the literature limited to DXA-derived vertebral fracture assessment (VFA)[7,15], rather than the gold standard of conventional radiography[23]. In this study we have evaluated the accuracies of KA and WOD for detection of thoracic VF (diagnosed by conventional radiography) in order to compare direct and indirect measures of kyphosis, and to determine whether WOD has sufficient accuracy to be used as a simple and rapid clinical assessment.

Materials and Methods

Study subjects

A series of 280 consecutive Caucasian women referred for assessment of osteoporosis was studied. Average age was 54.5 years (range: 18-92 years). The only reasons for exclusion were morphological abnormalities or skeletal distortions that prohibited either clinical measurements or morphometric assessment of skeletal radiographs. This study was approved by the local Research Ethics Review Board and written informed consent was provided by each participant.

Physical examination maneuvers

Physical measurements were made by a single examiner (KS). Stature was assessed with a Harpenden stadiometer as previously described, using the average of three measurements[4-6]. To measure KA, the patient stood unassisted, with the head facing forward in the Frankfort plane (where the head is positioned so that the inferior left orbital ridge is in the same horizontal plane as left tragion, the notch above the tragus of the ear)[24]. A handheld digital inclinometer (Fabrication Enterprises Inc.) was positioned over the spine at T12 and was zeroed. It was then moved to the second location over T4. The angle between the two positions was recorded to the closest degree. The average of three measurements was used. The precision of this method was 4.1° (root mean square SD). WOD was assessed with the patient standing against a vertical surface with the heels and buttocks touching the surface[3,7]. With the head facing forward in the Frankfort plane, the distance between the occiput and the wall was quantified with a tape measure to the closet 0.5 cm[3]. The average of three measurements was used. The precision was 0.6 cm (root mean square SD).

Vertebral morphometry and bone mineral densitometry

Lateral radiographs of the spine were made by standard methods. Morphometry was done from T4 to L4, with fracture defined as a vertebral height ratio <0.80 for anterior:posterior (defined as a wedge fracture) or middle:posterior (defined as an endplate fracture) height ratios within a vertebra, or posterior:posterior (defined as a crush fracture) height ratios when compared to an adjacent vertebra[4-6,25,26]. This corresponds to a fracture grade of one or worse on a four-point semiquantitative scale (grade 0=no fracture, <20% deformity; grade 1, 20 to <25% deformity; grade 2, 25 to <40% deformity; grade 3, 40% or greater deformity)[25,26]. Bone mineral density (BMD) of the lumbar spine and proximal femur was measured by DXA (QDR4500C).

Statistical analysis

Population descriptions are expressed as means (standard deviation, SD) for continuous variables and as percentages for discrete variables. Means were compared by the Mann-Whitney test and proportions by the Z test. Receiver operating characteristics (ROC) curves were generated and the areas under the curves (AUC) were calculated to summarize the ability of KA and WOD to predict the presence of thoracic VF. AUC comparisons were made using Rockit 0.9B (http://xray.bsd.uchicago.edu/krl/rocsoft.htm). Diagnostic accuracy was determined by evaluating the odds ratio (OR; the ratio of the risk of fracture in a given subgroup compared to the risk in a reference sub-group), positive likelihood ratio (LR+; likelihood that a test result will be positive among those with the condition in comparison to the likelihood among those without the condition), sensitivity (probability that a test result will be positive when the condition is present), and specificity (probability that a test result will be negative when the condition is absent). Positive predictive values (PPV; probability that the condition is present when the test is positive) and negative predictive values (NPV; probability that the condition is not present when the test is negative) were determined for several theoretical fracture prevalences. Some of these analyses were also carried out for different numbers of VF, for different VF grades, and for age sub-groups. Logistic regression was performed to determine change in fracture risk per SD change in physical measurements and linear regression models were developed to determine changes in KA and WOD in relation to thoracic VF number. All reported p values are two-tailed, and those less than 0.05 were considered to be statistically significant. Data was processed using SPSS version 12.0 (SPSS).

Results

One or more vertebral (thoracic or lumbar) fractures were present in 97 subjects (34.6%). Ninety-one subjects had thoracic fractures (70 with only thoracic fractures and 21 with both thoracic and lumbar fractures) while 6 had only lumbar fractures. The intent of the study was to determine whether the physical examination maneuvers under investigation were able to detect the presence of thoracic fractures, even when some of those with thoracic fractures also had lumbar fractures, and even when the comparator group included some people with lumbar fractures but no thoracic fractures. This approach would best reflect how physical examination would perform

Variable	All subjects (N=280)	No thoracic fractures (n=189)	Thoracic fractures (n=91)
Age -years	54.5 (17.5)	50.3 (16.3)	63.2 (16.7)***
Height - cm	160.9 (6.7)	161.7 (6.4)	158.9 (6.9)*
Fracture number	0.8 (1.5)	0	1.9 (1.3)
[a]KA - degrees	29.1 (11.3)	26.1 (8.7)	35.5 (13.6)***
[b]WOD - cm	2.1 (3.7)	0.9 (2.0)	4.4 (5.0)***
Spine [c]BMD - g/cm^2	0.891 (0.156)	0.923 (0.144)	0.818 (0.160)***
Spine T-score	-1.4 (1.4)	-1.1 (1.3)	-2.1 (1.4)***
Total hip [c]BMD - g/cm^2	0.813 (0.151)	0.839 (0.152)	0.751 (0.132)**
Total hip T-score	-1.1 (1.3)	-0.9 (1.3)	-1.6 (1.1)***
Femoral neck [c]BMD - g/cm^2	0.676 (0.139)	0.697 (0.139)	0.625 (0.126)***
Femoral neck T-score	-1.6 (1.3)	-1.4 (1.3)	-2.1 (1.1)***

*Values in parentheses are standard deviations. [a]KA=kyphosis angle; [b]WOD=wall-occiput distance. [c]BMD=bone mineral density. Mann-Whitney test, versus no thoracic fracture: *p<0.02, **p<0.01, ***p<0.001.*

Table 1. Characteristics of study subjects.

Variable	Tier	Cases with fracture	Cases without fracture	[a]LR+	[b]OR	[c]Sens	[d]Spec
				(95% [e]CI)			
[f]KA - degrees	0 to 20	9	42	0.5 (0.2-0.9)	1.0 reference	100 (95-100) (0-3)	0
	21 to 30	19	95	0.4 (0.3-0.7)	0.9 (0.4-2.2)	89 (81-94)	23 (18-31)
	31 to 40	31	32	2.1 (1.4-3.2)	4.5 (1.9-10.8)	67 (57-76)	75 (69-81)
	>40	26	13	4.3 (2.3-7.9)	9.3 (3.5-24.9)	31 (22-41)	93 (88-96)
[g]WOD - cm	0.0	25	130	0.4 (0.3-0.5)	1.0 reference	100 (95-100)	0 (0-3)
	0.1 to 2.0	14	11	2.6 (1.2-5.4)	6.6 (2.7-16.2)	71 (60-79)	76 (69-81)
	2.1 to 4.0	11	17	1.3 (0.6-2.7)	3.4 (1.4-8.0)	54 (44-64)	82 (76-87)
	4.1 to 6.0	11	7	3.2 (1.3-7.9)	8.2 (2.9-23.1)	41 (31-52)	92 (87-95)
	>6.0	24	7	6.9 (3.1–15.5)	17.8 (6.9-45.8)	28 (20-39)	96 (92-98)

[a]LR+=positive likelihood ratio; [b]OR=odds ratio. [c]Sens=sensitivity; [d]Spec=specificity. [e]CI=confidence interval; [f]KA=kyphosis angle. [g]WOD=wall-occiput distance.

Table 2. Accuracy of KA and WOD for detection of prevalent thoracic vertebral fractures.

in clinical practice. Therefore, subjects with thoracic fractures (n=91), including those who also had lumbar fractures, were compared to the group without thoracic fractures (n=189), which included those with no fractures (n=183) and those with only lumbar fractures (n=6). Those with thoracic VF had an average of 1.9 (SD, 1.3) fractures per person, were significantly older, shorter, and had lower spine and hip bone mineral densities than those without VF (Table 1). These patients also had greater KA at 35.5° (13.6°) compared to those without thoracic fractures who had KA of 26.1° (8.7°). Similarly, WOD was greater in those with thoracic fractures at 4.4 cm (5.0 cm) versus 0.9 cm (2.0 cm) in those without.

KA and WOD were moderately correlated (WOD= 0.23×KA-4.6; r=0.72, $p<10^{-11}$). The distribution of kyphosis measurements differed between those with thoracic fractures and those without thoracic fractures (Table 2). Among those with thoracic fractures, 67% had KA >30° compared to 25% among those without thoracic fractures. The ROC AUC for the detection of prevalent thoracic fractures using KA was 0.72 (95% CI, 0.65-0.79) and the odds for thoracic VF increased by 2.5-fold (95% CI, 1.8-3.4) for each SD increase in KA ($p=1.4x10^{-8}$). The study population was divided into four arbitrary groups (KA ≤20°, 21 to 30°, 31 to 40°, >40°). Using KA ≤20° as the reference group, the OR for thoracic VF rose as KA increased, reaching 9.3 (95% CI, 3.5-24.9) when KA >40°. The LR+ similarly increased with increasing KA, reaching 4.3 (95% CI, 2.3-7.9).

Among those with thoracic fractures, 41% had WOD >4.0 cm compared to 8% among those without thoracic fractures (Table 2). The odds for thoracic VF increased by 3.3 (95% CI, 2.2-4.8) for each SD increase in WOD ($p=1.6x10^{-8}$). The AUC under the ROC curve for the detection of prevalent thoracic fractures using WOD was 0.76 (95% CI, 0.69-0.82) and did not statistically differ from the corresponding AUC value for KA (p=0.13). The study population was divided into five arbitrary groups (WOD=0.0 cm, 0.1 to 2.0 cm, 2.1 to 4.0 cm, 4.1 to 6.0 cm, >6.0 cm). The OR for thoracic fracture rose as WOD increased, reaching 17.8 (95% CI, 6.9-45.8) when WOD >6.0 cm. Similarly, the LR+ rose to 6.9 (95% CI, 3.1-15.5) at WOD >6.0 cm. Sensitivity for thoracic fractures was 41% (95% CI, 31-52%) and specificity was 92% (95% CI, 87-95%) at a threshold of WOD >4.0 cm.

Linear regression models were developed to describe the relationships between measurements of kyphosis and the number of thoracic VF. The relationship between KA and thoracic VF was: KA (°) = (3.7×number of thoracic VF)+26.8 (r=0.37, $p=4 \times 10^{-10}$). Incorporating age as a variable, the relationship was: KA (°) = (2.6×number of thoracic VF)+(0.23×age)+14.8; r=0.50, $p<5x10^{-12}$). Each thoracic VF led to an increase in KA of 3.7° (95% CI, 2.6-4.8°) without adjustment for age and by 2.6° (95% CI, 1.5-3.7°) after adjustment for age. The relationship between WOD and thoracic VF was: WOD (cm)=(1.3×number of thoracic VF)+1.5; r=0.41, $p=2\times10^{-11}$). Incorporating age, the relationship was: WOD (cm)=(0.89×number of thoracic VF)+(0.088 x age)-3.3; r=0.57, $p<10^{-12}$). WOD increased by 1.3 cm (95% CI, 0.8-1.7 cm) per thoracic VF without adjustment for age and by 0.9 cm (95% CI, 0.5-1.2 cm)/fracture when age-adjusted. When ROC AUCs were compared for subjects over and under 50 years of age, there were no statistically significant differences. The AUC for KA was 0.60 (95% CI, 0.44-0.76) for the younger age group and 0.71 (95% CI, 0.63-0.80) for the older age group (p=0.41). The comparable values were 0.67 (95% CI, 0.52-0.82) and 0.75 (95% CI, 0.67-0.83) for WOD (p=0.81).

Further analysis was carried out on WOD for detection of different categories of thoracic VF, including fracture number, fracture grade, and fracture morphology. Forty-nine (54%) patients had one thoracic VF with mean WOD of 4.3 cm (SD, 5.2 cm), 20 (22%) had two VF with mean WOD of 4.3 cm (SD, 3.9 cm), 12 (13%) had 3 VF with mean WOD of 4.1 cm (SD, 6.2 cm), and 10 (11%) had four or more thoracic VF with mean WOD of 5.8 cm (SD, 5.3 cm). There was no statistical difference in detection of increasing numbers of fractures. At a threshold of WOD >4.0 cm, sensitivity was 38% (95% CI, 26-53%) for 1 VF, 50% (95% CI, 29-71%) for 2 VF, 33% (95% CI, 14-61%) for three VF, and 50% (95% CI, 22-79%) for 4 or more fractures.

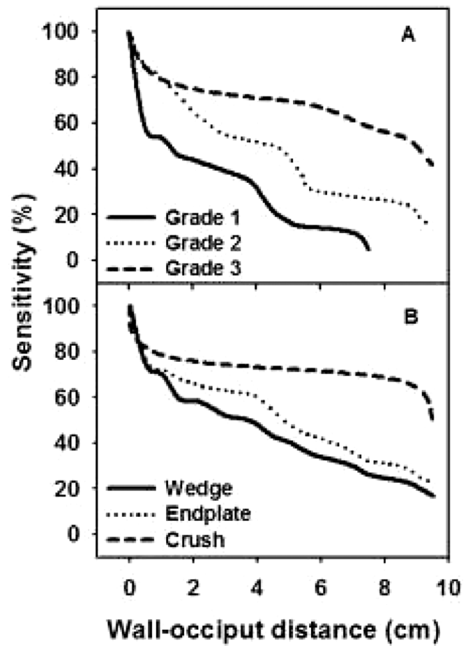

Figure 1. Sensitivity for detection of different fracture categories. **A.** Sensitivity of WOD for thoracic vertebral fractures of different grades. **B.** Sensitivity of WOD for thoracic vertebral fractures of different types.

Forty-four thoracic fracture subjects (48%) had a worst fracture grade of grade 1. This group had a mean WOD of 2.3 cm (SD, 2.9 cm). Grade 2 was the worst grade for 21 subjects (23%) with a mean WOD of 4.3 cm (SD, 3.8 cm), while 26 subjects (29%) had grade 3 fractures and mean WOD of 8.1 cm (SD, 6.6 cm). WOD was significantly different between those with worst fracture grades of 1 and 2 (p=0.025), and between 1 and 3 (p=0.0003). Similarly, sensitivity (at WOD>4.0 cm) increased with worsening fracture grade (Figure 1, panel A). Sensitivity rose from 41% (95% CI, 31-52%) for detection of grade 1 fractures to 59% (95% CI, 44-72%) for grade 2 fractures (p=0.05 compared to grade 1 fractures), and to 67 % (95% CI, 47-82%) for grade 3 fractures (p=0.022 for comparison to grade 1 fractures).

Eighty-two thoracic fracture subjects (90%) had wedge fractures and the subjects had a mean WOD of 4.6 cm (SD, 5.1 cm). Endplate fractures were present in 53 subjects (58%) with a mean WOD of 5.4 cm (SD, 5.0 cm), while 16 subjects (18%) had crush fractures and a WOD of 8.7 cm (SD, 6.4 cm). WOD values for those with wedge fractures and those with crush fractures were significantly different (p=0.009). Similarly, sensitivity (WOD >4.0 cm) was higher for detection of crush fractures (71%; 95% CI, 45-88%) than for wedge frac-

Parameter	Theoretical prevalence[a] (%)					
	1	5	10	15	25	50
Positive predictive value (%)	5	21	36	47	63	84
Negative predictive value (%)	99	97	93	90	82	61
[b]Proportion needing radiographs (%)	8	10	11	13	16	25

[a]Values were derived using sensitivity=41% and specificity=92%. [b]Proportion needing radiographs is the percentage of the population requiring spine radiographs at each fracture prevalence if every person having WOD>4.0 cm is sent for radiography

Table 3. Predictive values (post-test probability) at various fracture prevalences using WOD to detect prevalent thoracic vertebral fractures.

tures (43%; 95% CI, 32-54%) and endplate fractures (54%; 95% CI, 40-67%) (Figure 1, panel B). Only the sensitivities for crush and wedge fractures were significantly different (p=0.032).

PPV and NPV as defined by the sensitivity (41%) and specificity (92%) of WOD>4.0 cm were determined across a range of theoretical thoracic fracture prevalences likely to be encountered in clinical practice (Table 3). PPV remained low and rose to above 80% only at a high prevalence rate of 50%. Conversely, NPV remained high at low prevalences and fell to about 80% at 25% prevalence. If everybody above the WOD threshold was to be sent for X-rays at 25% prevalence, 16% of people would have spine radiography.

Discussion

Hyperkyphosis has been associated with aging since ancient times[27]. One of the earliest identified causes was osteoporotic VF, first linked to hyperkyphosis almost 50 years ago[28]. Since then, multiple possible contributors to hyperkyphosis have been described, including postural change, degenerative disc disease, muscle weakness, loss of ligamentous elasticity, genetic and metabolic disorders, infection, malignancy, and trauma[17,29-31]. Despite the large list of potential etiologies, osteoporotic VF, specifically those in the thoracic spine, remain the greatest single contributor to increased kyphosis, accounting for 42% to 48% of the variation in kyphosis between individuals[32].

VF are the most common type of fracture in people with osteoporosis, but unlike most other types of osteoporotic fractures, they tend to go undiagnosed[33]. One method of detecting prevalent VF is through assessing HHL, which has an AUC of 0.66 (95% CI, 0.59-0.72) for detection of prevalent VF, and at the recommended threshold of >6.0 cm, sensitivity is 30% (95% CI, 22-37%) and specificity is 94% (95% CI, 90-97%)[5]. By virtue of the fact that height loss is an assessment of the total spine, it can detect VF, but cannot determine where in the spine they are present. Assessment of RPD can localize VF to the lumbar spine, giving an AUC of 0.76 (95% CI, 0.71-0.81) for detection of lumbar VF[4]. At the recommended threshold of <2 fingerbreadths (about 3.5 cm), sensitivity for lumbar VF is 87% (95% CI, 81-93%) and specificity 47% (95% CI, 43-51%)[4]. Given the well-documented association of increased kyphosis with thoracic VF, our hypothesis was that measurement of the degree of kyphotic curvature by either direct or indirect assessment could detect thoracic VF with similar clinical accuracy to HHL for total VF and RPD for lumbar VF. Our results confirmed this.

Many devices have been used to directly quantify kyphosis, including mechanical inclinometers, goniometers, a specially designed variant of a goniometer called a kyphometer, a flexible rod to trace spine shape and then subsequently measure angles called a flexicurve or flexicurve ruler, a series of linked inclinometers called a curviscope, and a computer-aided surface measuring device[7,12-17,34]. We directly assessed kyphosis using a digital inclinometer and refer to the measured angle from T4 to T12 as the kyphosis angle, KA. We found that KA was greater in those with thoracic VF compared to those without (35.3º (SD, 13.6 º) vs. 26.1º (SD, 8.7º); p<0.001), that each thoracic VF led to an increase in KA of 3.7º (95% CI, 2.6-4.8º) without adjustment for age and by 2.6º (95% CI, 1.5-3.7º) after age adjustment, and that the ROC AUC for detection of thoracic VF was 0.72 (95% CI, 0.65-0.79). Prince and colleagues used a flexicurve ruler for directly measuring kyphosis to derive a parameter called the kyphosis index (KI), and determined accuracy parameters for detection of thoracic wedge VF at tertile thresholds (with VF diagnosed using DXA VFA rather than the gold standard of conventional radiography)[15]. Although their analysis was done only for wedge VF, not for all thoracic VF as in our study, their results are very close to our results for KA, as the majority of thoracic fractures are wedge morphology. The upper KI tertile had a sensitivity of 63% and specificity of 72%, indistinguishable from results at our KA threshold of >30º (which is approximately at our upper KA tertile) where sensitivity was 67% (95% CI, 57-76%) and specificity was 75% (95% CI, 69-81%). At the second KI tertile threshold, Prince et al. had a sensitivity of 90% and specificity of 29%, again virtually identical to our findings at a KA threshold of >20º, where we found a sensitivity of 89% (95% CI, 81-94%) and specificity of 23% (95% CI, 18-31%). Direct measurement of kyphosis by either flexicurve or digital inclinometer produces similar performance results in detection of thoracic VF.

By the nature of these direct kyphosis measurements, special devices are necessary to perform the quantification and so the techniques are limited to facilities that have such equipment. In our study, we have compared the performance of direct measurement with a simple clinical assessment, which we have referred to as WOD (also called OWD, the wall-occiput gap, and the gap-on-wall in the literature)[22,25], and found them to be correlated ($r=0.72$, $p<10^{-11}$). As with KA, we found that WOD was greater in those with thoracic VF compared to those without (4.4 cm (SD, 5.0 cm) vs. 0.9 cm (SD, 2.0) cm; $p<0.001$) and that each thoracic VF led to an increase in WOD of 1.3 cm (95% CI, 0.8-1.7 cm) without adjustment for age and by 0.9 cm (95% CI, 0.5-1.2 cm) after age adjustment. The WOD ROC AUC for detection of thoracic VF was 0.76 (95% CI, 0.69-0.82), a value not statistically different from the AUC for KA ($p=0.13$). Abe and colleagues assessed the ability of WOD to detect any VF (both thoracic and lumbar fractures as diagnosed using DXA VFA) in a small number of fracture subjects (n=28) and reported that at a threshold of >0.5 cm sensitivity was 70% (95% CI, 50 to 90%) and specificity was 67% (95% CI, 57 to 76%)[7]. These are indistinguishable from our findings using conventional radiography at a threshold of >0 cm, where we found sensitivity of 71% (95% CI, 60-79%) and specificity of 76% (95% CI, 68-81%).

Since WOD performs as well as direct measurement of kyphosis, and since it can be performed anywhere because the only equipment necessary to gauge WOD is a tape measure and a vertical surface, we propose that the most useful maneuver to apply in clinical assessment of kyphosis is measurement of WOD. Further, we suggest that a clinically useful threshold for detection of prevalent thoracic VF is WOD >4.0 cm. Using this threshold, sensitivity for thoracic VF is 41% (95% CI, 31-52%) and specificity is 92% (95% CI, 87-95%). The choice of a specific threshold to apply when using any physical exam maneuver as a clinical test depends on the specific application of the test, with trade-offs between false positives, false negatives, and the fraction of the tested population that will require gold standard testing (which has a bearing on cost and possible detrimental characteristics of the gold-standard test, in this case radiation exposure). Clinical thresholds of WOD >0.0 cm and >0.5 cm have been proposed in the literature[3,7]. Using our current results for the threshold of WOD >0.0 cm, with sensitivity of 71% and specificity of 76% as noted, results would be inferior to our proposed threshold of WOD >4.0 cm in all respects except sensitivity. Using a fracture prevalence of 25% for comparison, the threshold of WOD >0.0 cm would give a lower PPV (50% for WOD >0.0 cm vs. 63% for WOD >4.0 cm), lower NPV (76% vs. 82%), and many more screened subjects would exceed the detection threshold and would undergo radiography (36% for WOD >0.0 cm vs. 16% for WOD >4.0 cm). This supports the recommendation of a clinical threshold for WOD >4.0 cm, unless it is considered appropriate for the given application to achieve a higher sensitivity, with the trade-off of many more false positives and many more screened patients undergoing x-rays.

In assessing WOD, it is important that it be performed properly; in particular, care must be given in positioning the body. In an adult with normal spine shape and posture, the center of gravity passes through the T12/L1 junction and lies over the hips[34,36]. In individuals with hyperkyphosis, the forward displacement of the upper body moves upper body mass anteriorly, shifting the center of gravity anterior to the hips, and would cause the body to fall forward if there were not postural compensation[34,36]. This adjustment is accomplished by dorsiflexing the ankles, flexing the knees, and extending the hips, thereby rotating the thorax backwards so that weight is again over the legs[34-36,37], a posture noted by Dr Watson in describing a visitor to 221-B Baker Street[2]. Due to this compensatory positioning of the head, WOD cannot be measured with the patient standing freely. The hyperkyphosis needs to be unmasked by having the patient stand against a vertical surface, such as a wall or stadiometer[3,7]. By having the heels and buttocks touching the vertical surface, and then positioning the head so that it is facing forward (which is standardized by placing it in the Frankfort plane), the underlying thoracic curve is revealed in a reproducible manner.

There has been concern that grade 1 fractures as defined using the GSM may falsely assign anatomic variants as VF, particularly in the mid-thoracic region where vertebrae are normally more wedge-shaped[38,39]. This concern is clearly negated by the evidence, as grade 1 VF correlate with compromised bone structure, impaired quality of life, the presence of non-bone clinical features, and future fracture risk[8,40-42]. In addition, the ABQ method of VF evaluation, which uses qualitative radiological features rather than vertebral height ratios, finds a mid-thoracic focus of VF, just as seen in the GSM[39,43]. Further, the ABQ method finds a higher rate of false-negative VF than false-positive VF in the mid-thoracic spine[43]. Given the solid body of evidence in support of the GSM, it has become the recommended clinical approach to VF diagnosis by several organizations, and is the standard methodology in automated vertebral fracture assessment (VFA) by DXA[44,45].

A potential limitation of this study is that it was performed in a specialist setting using patients referred for osteoporosis assessment. The fact that results are very similar to those of two epidemiological studies with randomly chosen participants (one in Japan, one in Australia) suggests that our results are generalizable among postmenopausal women[7,15]. Further studies would be needed to verify that the same relationships exist between KA and WOD and vertebral deformities in men, and whether the same thresholds would apply.

Although there are many possible causes and contributors to hyperkyphosis, VF is one of the most common etiologies. Our results show that measurement of kyphosis using WOD can serve as a useful clinical tool for detecting the presence of thoracic VF. It is a quick and reproducible method that does not require specialized equipment and should be incorporated into the clinical assessment of osteoporosis patients and those at risk for osteoporosis, along with evaluation of height loss and measurement of RPD.

References

1. Updike J. In the Beauty of the Lilies. New York: Alfred A Knopf; 1996. p. 8-9.
2. Conan Doyle A. The adventure of the noble bachelor. In: The Complete Sherlock Holmes. Garden City, New York: Doubleday and Company; 1938. p. 291.

3. Green AD, Colón-Emeric CS, Bastian L, Drake MT, Lyles KW. Does this woman have osteoporosis? JAMA 2004;292:2890-900.
4. Siminoski K, Warshawski RS, Jen H, Lee KC. Accuracy of physical examination using the rib-pelvis distance for detection of lumbar vertebral fractures. Am J Med 2003; 115:233-6.
5. Siminoski K, Warshawski RS, Jen H, Lee K. The accuracy of historical height loss for detection of vertebral fractures in postmenopausal women. Osteoporos Int 2005;16:290-6.
6. Siminoski K, Jiang G, Adachi JD, Hanley DA, Cline G, Ioannidis G, et al. Accuracy of height loss during prospective monitoring for detection of incident vertebral fractures. Osteoporos Int 2005;16:403-10.
7. Abe K, Tamaki J, Kadowaki E, Sato Y, Morita A, Komatsu M, et al. Use of anthropometric indicators in screening for undiagnosed vertebral fractures: a cross-sectional analysis of the Fukai Osteoporosis Cohort (FOC) study. BMC Musculoskel Dis 2008;9:157-66.
8. Ismail AA, Cooper C, Felsenberg D, Varlow J, Kanis JA, Silman AJ, et al. Number and type of vertebral deformities: epidemiological characteristics and relation to back pain and height loss. Osteoporos Int 1999;9:206-13.
9. Moayyeri A, Luben RN, Bingham SA, Welch AA, Wareham NJ, Khaw K. Measured height loss predicts fractures in middle-aged and older men and women: the EPIC-Norfolk prospective population study. J Bone Miner Res 2009;23:425-32.
10. Olszynski WP, Ioannidis G, Sebaldt RJ, Hanley DA, Petrie A, Brown JP, et al. The association between iliocostal distance and the number of vertebral and non-vertebral fractures in women and men registered in the Canadian Database for Osteoporosis and Osteopenia (CANDOO). BMC Musculoskeletal Disord 2002;3:22-9.
11. Middleton ET, Gardiner ED, Steel SA. Which women should be selected for vertebral fracture assessment? Comparing different methods of targeting VFA. Calcif Tissue Int 2009;85:203-10.
12. Cortet B, Roches E, Logier R, Houvenagel E, Gaydier-Souquieres G, Puisieux F, et al. Evaluation of spinal curvatures after a recent osteoporotic vertebral fracture. Joint Bone Spine 2002;69:201-8.
13. Ensrud KE, Black DM, Harris F, Ettinger B, Cummings SR. Correlates of kyphosis in older women. J Am Geriatr Soc 1997;45:682-7.
14. Geusens P, Vosse D, vander Heijde D, Vanhoof J, van Tubergen A, Raus J, et al. High prevalence of thoracic vertebral deformities and discal wedging in ankylosing spondylitis patients with hyperkyphosis. J Rheumatol 2001;28:1856-61.
15. Prince RL, Devine A, Dick IM. The clinical utility of measured kyphosis as a predictor of the presence of vertebral deformities. Osteoporos Int 2007;18:621-7.
16. Ettinger B, Black DM, Palermo L, Nevitt MC, Melnikoff S, Cummings SR. Kyphosis in older women and its relation to back pain, disability and osteopenia: the study of osteoporotic fractures. Osteoporos Int 1994;4:55-60.
17. Kado DM, Prenevost K, Crandall C. Narrative review: hyperkyphosis in older persons. Ann Intern Med 2007; 147:330-8.
18. Roux C, Fechtenbaum J, Kolta S, Said-Nahal R, Briot K, Benhamou C. Prospective assessment of thoracic kyphosis in postmenopausal women. J Bone Miner Res 2010; 25:362-8.
19. O'Brien, Culham E, Pickles B. Balance and skeletal alignment in a group of elderly female fallers and non-fallers. J Gerontol A Biol Sci Med Sci 1997;52:221-6.
20. Michel A, Kohlmann T, Raspe H. The association between clinical findings on physical examination and self-reported severity in back pain. Results of a population-based study. Spine 1997;22:296-304.
21. Leidig-Bruckner G, Minne HW, Schlaich C, Wagner G, Scheidt-Nave C, Bruckner T, et al. Clinical grading of spinal osteoporosis: quality of life components and spinal deformity in women with chronic low back pain and women with vertebral osteoporosis. J Bone Miner Res 1997;12:663-75.
22. Vosse D, van der Heijde DMFM, Landewe RBM, Geusens P, Mielants H, Dougados M, et al. Determinants of hyperkyphosis in patients with ankylosing spondylitis. Ann Rheum Dis 2006;65:770-4.
23. Fuerst T, Wu C, Genant HK, von Ingersleben G, Chen Y, Johnston C, et al. Evaluation of vertebral fracture assessment by dual-energy X-ray absorptiometry in a multicenter setting. Osteoporos Int 2009;20:1199-205.
24. Norton K, Whittingham N, Carter L, Kerr D, Gore C, Marfell-Jones M. Measurement techniques in anthropometry. In: Norton K, Olds T editors. Anthropometrica. Sydney, Australia: University of New South Wales Press; 1996. p. 25-75.
25. Genant HK, Wu CY, van Kuikj C, Nevitt MC. Vertebral fracture assessment using a semiquantitative technique. J Bone Miner Res 1993;8:1137-48.
26. Grados F, Roux C, de Vernejoul MC, Utard G, Sebert JL, Fardellone P. Comparison of four morphometric definitions and a semiquantitative consensus reading for assessing prevalent vertebral fractures. Osteoporos Int 2001; 12:716-22.
27. Manolagas SC. Birth and death of bone cells: basic regulatory mechanisms and implications for the pathogenesis and treatment of osteoporosis. Endocrine Rev 2000; 21:115-37.
28. Rowe CR, Sorbie CPD. Fractures of the spine in the aged. Clin Orthop Rel Res 1963;26:34-49
29. Dalbayrak S, Onen MR, Yilmaz M, Naderi S. Clinical and radiographic results of balloon kyphoplasty for treatment of vertebral body metastases and multiple myelomas. J Clin Neurosci 2010;17:219-24.
30. Huang QS, Zheng C, Hu Y, Yin X, Xu H, Zhang G, et al. One-stage surgical management for children with spinal tuberculosis by anterior decompression and posterior in-

strumentation. Int Orthop 2009;33:1385-90.
31. Munting E. Surgical treatment of post-traumatic kyphosis in the thoracolumbar spine: indications and technical aspects. Eur Spine J 2010;19(Suppl.1):S69-73.
32. Milne JS, Williamson J. A longitudinal study of kyphosis in older people. Age Ageing 1983;12:225-33.
33. Majumdar SR, Kim N, Coleman I, Chahal AM, Raymond G, Jen H, et al. Incidental vertebral fractures discovered with chest radiography in the emergency department: prevalence, recognition, and osteoporosis management in a cohort of elderly patients. Arch Intern Med 2005; 165:905-9.
34. Ishikawa Y, Miyakoshi N, Kasukawa Y, Hongo M, Shimada Y. Spinal curvature and postural balance in patients with osteoporosis. Osteoporosis Int 2009;20:2049-53.
35. Davis SR, Kirby C, Weeks A, Lanzafame A, Piterman L. Simplifying screening for osteoporosis in Australian primary care: the Prospective Screening for Osteoporosis: Australian Prmary Care Evaluation of Clinical Tests (PROSPECT) study. Menopause 2010;18:1-7.
36. Chow RK, Harrison JE. Relationship of kyphosis to physical fitness and bone mass on post-menopausal women. Am J Phys Med 1987;66:219-27.
37. Bot SDM, Caspers M, Van Royen BJ, Toussaint HM, Kingma I. Biomechanical analysis of posture in patients with spinal kyphosis due to ankylosing spondylitis: a pilot study. Rheum 1999;38:441-3.
38. Grigoryan M, Guermazi A, Roemer FW, Delmas PD, Genant HK. Recognizing and reporting osteoporotic vertebral fractures. Eur Spine J 2003;12(Suppl):S104-S112.
39. Ferrar L, Jiang G, Clowes JA, Peel NF, Eastell R. Comparison of densitometric and radiographic vertebral fracture assessment using the algorithm-based qualitative (ABQ) method in postmenopausal women at low and high risk of fracture. J Bone Miner Res 2008;23:103-111.
40. Siris ES, Genant HK, Laster AJ, Chen P, Misurski DA, Krege JH. Enhanced prediction of fracture risk combining vertebral fracture status and BMD. Osteoporos Int 2007; 18:761-770.
41. Delmas PD, Genant HK, Crans GG, Stock JL, Wong M, Siris E, et al. Severity of prevalent vertebral fractures and the risk of subsequent vertebral and non-vertebral fractures: results from the MORE trial. Bone 2003;33:522-532.
42. Tanko LB, Christiansen C, Cox DA, Geiger MJ, McNabb MA, Cummings SR. Relationship between osteoporosis and cardiovascular disease in postmenopausal women. J Bone Miner Res 2005;20:1912-1920.
43. Ferrar L, Jiang G, Cawthon PM, San Valentin R, Fullman R, Lambert L, et al. Identification of vertebral fracture and non-osteoporotic short vertebral height in men: the MrOS study. J Bone Miner Res 2007;22:1434-1441.
44. Vokes T, Bachman D, Baim S, Binkley N, Broy S, Ferrat L, et al. Vertebral fracture assessment: the 2005 ISCD official positions. J Clin Densitom 2006;9:37-46.
45. Lentle BC, Brown JP, Khan A, Leslie WD, Levesque J, Lyons DJ, et al. Recognizing and reporting vertebral fractures: reducing the risk of future osteoporotic fractures. Can Assoc Radiol J 2007;58:27-36.