Osteoporosis Int (1991) 2:48–53

Osteoporosis International

Original Article

Screening for Vertebral Osteoporosis Using Individual Risk Factors

The Multicentre Vertebral Fracture Study Group: C. Cooper (Epidemiologist)[1], S. Shah (Co-ordinator)[2], D. J. Hand (Professor of Statistics)[3], J. Adams (Consultant Radiologist)[4], J. Compston (Consultant Physician)[5], M. Davie (Consultant Physician)[6] and A. Woolf (Consultant Rheumatologist)[7]

[1]Rheumatology Unit, Bristol Royal Infirmary, [2]Medical Department, Sandoz Pharmaceuticals, [3]Faculty of Mathematics, The Open University, Milton Keynes, [4]Department of Diagnostic Radiology, University of Manchester Medical School, [5]Department of Medicine, University of Cambridge Medical School, [6]Robert Jones and Agnes Hunt Orthopaedic Hospital, Oswestry, and [7]Royal Cornwall Hospital (City), Truro, UK

Abstract. Osteoporosis is a major cause of ill health in postmenopausal women. Several risk factors for osteoporosis have been identified, and they have been widely recommended as a means of identifying subgroups of postmenopausal women who might benefit from prophylaxis and therapy. Evidence to support this use of risk factors is currently lacking, however. We have constructed and evaluated a profile of putative risk factors as a means of identifying women attending general practitioners who have sustained vertebral fractures. The overall prevalence of vertebral fractures in the 1012 women (mean age 64.4 years) studied was 7.8%. Women who had sustained vertebral fractures in this population were significantly ($p<0.05$) older and shorter than those without fractures. They reported a significantly ($p<0.05$) earlier menopause, lower parity and a greater prevalence of hyperthyroidism. However, the best screening instrument devised was not sufficiently predictive to warrant widespread use.

Keywords: Risk factors; Screening; Vertebral fractures; Osteoporosis

Correspondence and offprint requests to: Dr Cyrus Cooper, Senior Research Fellow, Section of Clinical Epidemiology, Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota 55905, USA.

Introduction

Osteoporosis constitutes a major public health problem through its association with fractures of the hip, vertebrae and distal radius [1]. The most frequent manifestation of osteoporosis in women aged less than 75 years is believed to be vertebral fracture [2]. These fractures are often asymptomatic and a clinical diagnosis is thus frequently delayed until multiple vertebral fractures, height loss and spinal deformity are present. Several forms of treatment may benefit patients with established osteoporosis [3]. These are more likely to be effective if started early in the course of the disease.

At present, vertebral fractures can only reliably be demonstrated by conventional radiography of the dorsal and lumbar spine. A number of risk factors for osteoporosis have been identified [4], raising the possibility that risk factor profiles might provide an alternative and less invasive approach to the identification of women with vertebral fractures. Such a method would be inexpensive and suitable for use in general practice. We have therefore constructed a profile for obtaining information on putative risk factors for osteoporosis and estimated its performance as a predictor of vertebral fracture in a sample of women attending general practitioners in the United Kingdom.

Patients and Methods

The study included 1012 peri- and postmenopausal women aged between 48 and 81 years who attended their general practitioner for any complaint and agreed

to participate in the study. This sample was recruited from 19 general practices located in 10 towns throughout the UK (Belfast, Bristol, Cambridge, Dundee, Edinburgh, Leeds, Liverpool, Macclesfield, Oswestry and Truro). These centres were chosen so as to cover a broad range of the socio-economic conditions found within the UK, and for their proximity to centres with experience in the management of osteoporosis. In each practice, women attending their physician for any complaint were invited to attend for interview. A total of 1209 eligible women were approached, of whom 1136 (94%) agreed to interview and 1012 (89%) consented to further thoracolumbar radiography.

All women who fulfilled entry criteria were interviewed by a trained research nurse who administered a structured questionnaire. This obtained information on 24 variables for which published evidence suggests an association with osteoporosis. The variables are listed in the Appendices. They included age, reproductive variables, previous medical and family history, cigarette smoking, alcohol consumption, physical activity and dairy product intake. For some variables, the questionnaire was constructed so as to elicit a yes/no response (Appendix 1). For the remainder, responses were treated as continuous variables (Appendix 2). Cut-off values for certain variables such as activity and cigarette smoking were decided upon *a priori*, based upon available data about their role as risk factors for osteoporotic fractures [5,6]. The questionnaire was designed by a panel of hospital physicians engaged in the routine management of osteoporosis, in collaboration with an epidemiologist and biostatistician. Measurements of weight and height were made using the scales and stadiometers available in each practice.

On completion of the questionnaire, all the study subjects were invited to attend their local hospital radiology department for lateral thoracolumbar spine radiography. Two coned lateral spinal radiographs, one of the thoracic and one of the lumbar spine, were performed in each subject. All radiographs in each centre were performed according to a standardized technique and included views from T4 to L5. They were assessed for evidence of vertebral fracture by a single consultant radiologist (J.A.) who manually measured the anterior, posterior and mid-point heights of each vertebral body using a ruler. Wedge fractures were defined as a 20% or greater reduction in anterior vertebral height. Crush fractures were defined as a 20% or greater reduction in anterior and posterior height compared with adjacent vertebrae. Biconcavity was defined as a 20% or greater reduction in mid-vertebral height. The reproducibility of radiographic assignment of fracture according to this algorithm was evaluated in a sample of 561 study radiographs, presented in random order to the radiologist on two successive occasions. At the second reading she was unaware of her assessment at the initial reading. In order to maximize comparability of prevalence rates with estimates for other studies, the radiologist also recorded fracture cases according to a 15% cut-off rule for wedge, crush and biconcave deformities.

The prevalence of vertebral fractures was calculated within 5-year age bands for both the 20% and 15% radiographic criteria. Comparison with other prevalence estimates was made by age standardization to the United States white female population in 1990. Reproducibility of the radiographic assessment was quantified using the kappa statistic, around which 95% confidence intervals were calculated [7]. Individual questionnaire items were examined by comparing vertebral fracture cases and non-fracture controls using chi-squared tests for the categorical variables and *t*-tests for the continuous variables. Classical discriminant analysis [8] was used to identify the group of questions which discriminated best between cases and controls, and the sensitivity, specificity and predictive value of this instrument was quantified. Whenever the performance of a classification rule is assessed within the data set which was used to design the rule, bias is introduced as a result of the sampling process. This bias acts towards optimizing the performance of the classification rule. The most efficient statistical means of allowing for this bias is the leaving-one-out method [9]. This builds the rule based on all cases but one, classifies this case, and repeats the process for each case in turn. The overall misclassification rate is then an almost unbiased estimate of that which would be obtained using the complete data-set, since the design and test sets for each classification rule are independent.

Ethical approval for the study was obtained from the Ethics Committee of the Royal College of General Practitioners.

Results

The mean age of the 1012 women in the study was 64.4 years. Seventy-nine (7.8%) of them were found to have one or more vertebral fractures according to criteria incorporating a 20% reduction in anterior, with or without posterior, height. Table 1 shows the prevalence of vertebral fractures with advancing age. The prevalence rate rose from 4.3% at 55–59 years, to 27.8% at 80 years and over. It also shows the higher prevalence obtained in each age category if the criteria for vertebral fracture definition incorporate a 15% reduction in vertebral heights. Of the 110 fractured vertebrae in the 79 women, the majority (82, 74%) were of the wedge variety, 22 (21%) were biconcave fractures and only 6 (5%) were crush fractures. Although the prevalence of each fracture type increased with age, the small number of crush and biconcave fractures limited statistical analysis of the individual trends. The prevalence of reported back pain in the previous year was similar in those with vertebral fractures (35.4%) to that in patients without fractures (36.1%).

Table 2 shows that in the sample of 561 radiographs which were used in the assessment of within-observer variation, there was agreement on fracture classification using the 20% rule in 545 (97%). This gave a kappa statistic of 0.73 (95% confidence interval 0.62–0.84).

Table 1. Prevalence of all vertebral fractures with age among 1005 women attending British general practitioners

Age group (years)	Total number of women	Prevalence of vertebral fracture			
		20% criterion		15% criterion	
		No.	%	No.	%
<50	1	0	0	0	0
50–54	17	1	5.9	4	23.5
55–59	282	12	4.3	49	17.4
60–64	244	17	7.0	42	17.3
65–69	218	23	10.6	53	24.3
70–74	120	9	7.5	30	25.0
75–79	105	11	10.5	22	21.2
>79	18	5	27.8	7	38.9
Total	1005[a]	78	7.8	207	20.6

[a] Seven radiographs not evaluated due to poor technical quality.

Table 2. Within-observer reproducibility of radiographic assessment method in a sample of 561 study radiographs

Second reading	First reading		
	Fracture	Non-fracture	Total
Fracture	22	14	36
Non-fracture	2	523	525
Total	24	537	561

Kappa = 0.73 (0.62–0.84).

Table 3. Discriminant function resulting from backwards stepwise elimination

Variable	Coefficient	Coefficient of standardized function
1. Age (years)	0.053	0.367
2. Height (cm)	−0.076	−0.491
3. Broken bone in back after age 45 (1=yes, 2=no)	−7.015	−0.377
4. Age of last menstrual period (years)	−0.052	−0.269
5. Number of children (1=1 or more, 2=none)	1.212	0.494
6. Ever-use of oral corticosteroids (1=yes, 2=no)	1.043	0.296

The distribution of each of the variables in the risk factor questionnaire was compared between the 79 women with vertebral fractures and the 933 without fractures (Appendixes 1 and 2). Those with fractures were statistically significantly ($p<0.05$) older and shorter. They reported an earlier menopause, lower parity and a greater frequency of hyperthyroidism. There was also a tendency for them to report an earlier menarche and to have a heavier alcohol intake, although these differences did not attain statistical significance.

Various statistical discrimination techniques were used to construct screening rules. There was very little difference in performance between the resulting rules. The best screening rule derived using backwards stepwise elimination in classical discriminant analysis is given in Table 3. This included age, height, a history of a broken bone in the back after age 45, age at menopause, parity and ever-use of oral corticosteroids. A history of medically attended hyperthyroidism was the only variable showing a significantly different distribution between cases and non-cases in the univariate analysis, which was not selected in the classical discriminant model. The table provides the coefficients in the model which best discriminate between the case/non-case groups. The overall score for each subject is constructed using these coefficients and compared with a threshold to give a case/non-case classification. As the variables in the table have different ranges, we also present the standardized coefficients. These permit the relative importance of the variables in the discriminant function to be assessed – a larger value signifying that a change in that variable has a greater impact.

Even using this 'best' statistical model, very poor discrimination was achieved between cases and non-cases. It was not possible to select a threshold such that cases predominated on one side and non-cases on the other. Table 4 shows the classification tables with sensitivity, specificity and positive predictive value at three cut-off points – low, intermediate and high sensitivity. When a cut-off was chosen in the high sensitivity region (0.91), the resulting classification rule was highly non-specific (0.23) and had poor positive predictive value (0.09). Thus the questionnaire identifies 79% of all women screened as having vertebral fractures, when in truth only 9% of this group have fractures on radiography. At the other extreme, a high specificity (0.94) is associated with very low sensitivity (0.15). In this scenario, although the number of false-positives is smaller, the questionnaire only identifies 17% of the

Table 4. Leaving-one-out performance estimates of the discrimination rule

	Classification table		Specificity	Sensitivity	PV+[a]
	Predicted				
True	Case	Non-case			
Case	8	47			
Non-case	41	579	0.94	0.15	0.17
True	Case	Non-case			
Case	28	27			
Non-case	199	433	0.69	0.51	0.12
True	Case	Non-case			
Case	49	6			
Non-case	484	148	0.23	0.91	0.09

[a] PV+, positive predictive value.

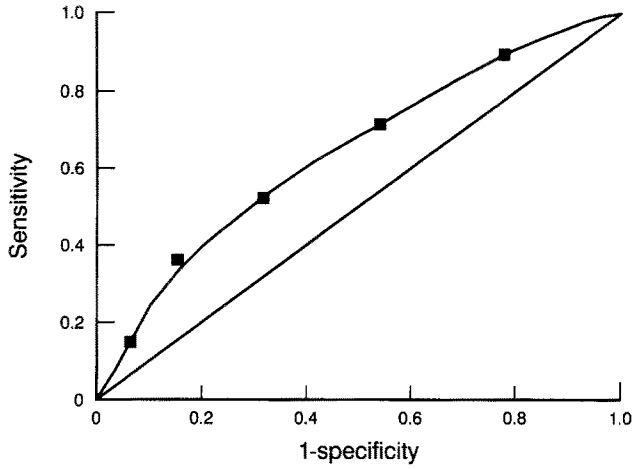

Fig. 1. Receiver–operator characteristic (ROC) curve obtained for the optimal screening instrument using risk factors to distinguish women with vertebral fractures from those without fractures

women with radiographically proven fracture. This generally poor performance of the screening instrument is illustrated graphically by the receiver–operator characteristic (ROC) curve (Fig. 1).

Discussion

We report the design and evaluation of an osteoporotic risk factor profile in a multicentre general practice survey of vertebral fractures in women. Our results suggest that the optimal screening instrument derived from the 24 historical risk factors studied is unlikely to be sufficiently predictive of vertebral fracture among women attending general practitioners as to warrant widespread use. An instrument, for example, with a sensitivity of 91% would only have a specificity of 23%, and succeed in identifying around one-tenth of women with vertebral fractures.

We also report the first prevalence estimates for vertebral fracture in postmenopausal British women attending their general practitioners. The age-specific prevalence rates are comparable with those derived in population-based studies in the United States [2] and Scandinavia [10]. Such comparisons must be made with caution, however, as there is no general consensus on the definition of vertebral fracture which should be used in epidemiological studies. The difficulty in deciding whether a vertebral body is, indeed, fractured stems from variation in the shape of vertebrae both at different levels in the spine, and between individuals. The most recent approach to this problem has been to measure vertebrae from healthy subjects to establish normal ranges for their anterior, posterior and mid-point heights, and then to construct an algorithm according to which moderate or severe deformity can be estimated [11–13]. However, most currently available epidemiological data on vertebral fracture adopt less rigorous criteria. Data on the reproducibility of vertebral radiogrammetry are also scarce. Although the precision of digitized vertebral morphometry is known to be high [12], information is not available on the observer error introduced when morphometry is manually undertaken. The within-observer variation in our study, characterized by a kappa statistic of 0.73, suggests that the technique might be unsuitable for use in individual subjects, especially if they are being sequentially studied. It is, however, sufficiently reproducible for studies in populations, or large subgroups within populations.

In the best data available from the United States [2], the prevalence of vertebral fractures, defined according to an algorithm approximating to a 15% cut-off, rose from 5.9% at age 50–54 years to 46.5% in those aged 85–89 years. Age adjustment to the US white female population in 1990 gives an overall prevalence rate of 18.7% among women 50 years and over in the Rochester study [2], and one of 10.4% using our 20% cut-off. When our criteria were relaxed to 15%, however, the age-adjusted prevalence rate rose to 23.9% in our study. In a Danish study of women aged 70 years, Jensen et al. [10] reported a prevalence of one or more vertebral fractures in 22.8%, as compared with our rates of 7.5% and 25% using the 15% and 20% cut-off criteria, respectively. Our decision to adopt the 20% definition in analysing the risk factor profile was based on the observation that fracture definition according to the 15% rule has a higher frequency of false-positives than definitions using more stringent criteria [12].

The study was based on a large sample of women from 10 general practices in Britain. It is unlikely, therefore, that the populations sampled would have been markedly biased in their socio-economic constitution. Bias may have arisen, however, in the choice of general practice attenders as the study sample. Women would thus have been excluded who could not make the journey to see their general practitioner. Although such under-representation of co-morbidity might have influenced the observed prevalence of vertebral fracture, it would not have altered the performance characteristics of the risk factor questionnaire. Entry criteria into the study were deliberately broad for two reasons. First, it seemed appropriate to assess the questionnaire in the population sample for which it was ultimately intended. Second, we also wished to examine the prevalence of vertebral fractures in this large population of women. Thus, women who reported a previous diagnosis of vertebral fracture were not excluded. The number of such women (2 cases, 3 non-cases) was small, however, and would not have diminished the observed quality of our screening instrument.

The range of risk factors to be included in a profile such as this is large. In designing the questionnaire we attempted to include those variables which have been suggested as either determinants of bone density or risk factors for fracture. In many instances the strongest evidence implicating variables as risk factors for osteoporosis comes from analytical studies of hip fracture. It is difficult to know how applicable these might be to vertebral fracture. The pathogenesis of hip fracture

represents a complex interaction between osteoporosis and propensity to trauma. Vertebral fractures, on the other hand, often occur without preceding trauma and may be more closely associated with low bone density [14]. The few case-control studies of vertebral fracture to have been performed [15,16] have suggested that body weight, early menopause in women, smoking, alcohol consumption and various diseases are important risk factors. These variables are known determinants of bone density [17] and were included in our questionnaire.

Several studies have concluded that risk factor questionnaires for osteoporosis are poor predictors of both axial and appendicular bone density in peri- and postmenopausal women [18, 19]. In addition, recent studies from the United States [20, 21] have reported that historical risk factors demonstrate limited sensitivity and specificity in discriminating between women with vertebral fractures and outpatient controls. Our results are in accord with these findings and suggest that such questionnaires are also likely to have limited value in the general practice setting.

Appendices. Variables Included in the Questionnaire and Their Distribution Among the Subjects

Appendix 1. Categorical variables: distribution among women with and without vertebral fractures

Variable	Number positive		χ^2	p-value
	With fracture (n=79)	Without fracture (n=933)		
1. Persistent back pain in last year	28	337	0.00	1.00
2. Previous broken bone in back	2	3	3.38	0.07
3. Previous wrist fracture	7	91	0.01	0.94
4. Oophorectomy	14	132	0.55	0.46
5. Hysterectomy	15	193	0.06	0.80
6. Use of hormone replacement	7	138	1.69	0.19
7. Temporary amenorrhoea >6 months	2	59	1.20	0.27
8. Parity (1 or more children)	54	763	7.73	0.005[a]
9. Lactation >1 month per child	37	541	0.00	1.00
10. Ever-use of corticosteroids	5	82	0.25	0.62
11. Ever-use of thyroxine	10	87	0.62	0.43
12. Ever-diagnosis of hyperthyroidism	13	68	7.12	0.008[a]
13. Ever partial gastrectomy	5	39	0.39	0.53
14. Family history of fracture	18	223	0.002	0.97
15. Smoker (>10 cigarettes/day, >10 years)	18	243	0.28	0.60
16. Walking outdoors (>½ hour/day)	51	605	0.00	1.00
17. Participation in sports (> monthly)	22	291	0.26	0.61

[a] $p<0.05$.

Appendix 2. Continuous variables: distribution among women with and without vertebral fractures

Variable	Mean value (SD)		t-statistic	p-value
	Fracture (n=79)	Non-fracture (n=933)		
1. Age (years)	67.2 (7.4)	64.4 (7.1)	3.30	0.001[a]
2. Height (cm)	158.1 (6.7)	160.0 (6.3)	−2.52	0.012[a]
3. Weight (kg)	65.6 (12.2)	65.7 (10.9)	−0.06	0.95
4. Age at menarche (years)	13.6 (1.7)	13.3 (1.7)	1.49	0.14
5. Age at menopause (years)	47.3 (5.2)	48.6 (5.1)	−2.00	0.045[a]
6. Alcohol consumption (units/week)	1.7 (2.9)	1.6 (3.7)	—	0.80[b]
7. Consumption of milk (pints/week)	4.0 (1.9)	4.2 (2.3)	−0.75	0.46

[a] $p<0.05$.
[b] Skewed distribution; comparison made using Mann–Whitney test.

Acknowledgements. This study was supported by Sandoz Pharmaceuticals Ltd. We are grateful to the general practitioners, research staff and hospital radiology department staff who assisted in collection of the data.

References

1. Cooper C. Osteoporosis: an epidemiological perspective. A review. J R Soc Med 1989; 82:753–7.
2. Melton LJ, Kan SH, Frye MA, Wahner HW, O'Fallon WM, Riggs BL. Epidemiology of vertebral fractures in women. Am J Epidemiol 1989; 129:1000–11.
3. Consensus Development Conference. Prophylaxis and treatment of osteoporosis. Br Med J 1987; 295:914–15.
4. Cummings SR, Kelsey JL, Nevitt MC, O'Dowd KJ. Epidemiology of osteoporosis and osteoporotic fractures. Epidemiol Rev 1985; 7:178–208.
5. Cooper C, Barker DJP, Wickham C. Physical activity, muscle strength and calcium intake in fracture of the proximal femur. Br Med J 1988; 297:1443–6.
6. Cooper C, Wickham C. Cigarette smoking and the risk of age-related fractures. In: Wald N, Frogatt P, eds. Smoking and hormone related disorders. Oxford University Press, 1990:93–100.
7. Fleiss JL. Statistical methods for rates and proportions. New York: Wiley, 1981:212–36.
8. Hand DJ. Discrimination and classification. Chichester: Wiley, 1981.
9. Hand DJ. Recent advances in error rate estimation. Pattern Recognition Lett 1986; 4:335–6.
10. Jensen GF, Christiansen C, Boesen J, Hegedus V, Transbøl I. Epidemiology of postmenopausal spinal and long bone fractures: a unifying approach to postmenopausal osteoporosis. Clin Orthop 1982; 166:75–81.
11. Eastell R, Cedel SL, Wahner HW, Riggs BL, Melton LJ. Classification of vertebral fractures. J Bone Miner Res 1991; 6:207–15.
12. Smith-Bindman R, Cummings SR, Steiger P, Genant HK. A comparison of morphometric definitions of vertebral fracture. J Bone Miner Res 1991; 6:25–34.
13. Hedlund LR, Gallagher JC. Vertebral morphometry in diagnosis of spinal fractures. Bone Miner 1988; 5:59–67.
14. Hayes WC, Piazza SJ, Zysset PK. Biomechanics of fracture risk prediction of the hip and spine by quantitative computed tomography. Radiol Clin North Am 1991; 29:1–18.
15. Aloia JF, Cohn SH, Vaswani A, Yeh JK, Yuen K, Ellis K. Risk factors for postmenopausal osteoporosis. Am J Med 1985; 78:95–100.
16. Seeman E, Melton LJ, O'Fallon WM, Riggs BL. Risk factors for spinal osteoporosis in men. Am J Med 1983; 75:977–83.
17. Riggs BL, Melton LJ. Involutional osteoporosis. N Engl J Med 1986; 314:1676–86.
18. Elders PJM, Netelenbos JC, Lips P, Khoe E, van Ginkel FC, Hulshof KFAM, van der Stelt PF. Perimenopausal bone mass and risk factors. Bone Miner 1989; 7:289–99.
19. Slemenda CW, Hui SL, Longcope C, Wellman H, Johnston CC. Predictors of bone mass in perimenopausal women. A prospective study of clincial data using photon absorptionmetry. Ann Intern Med 1990; 112:96–101.
20. Wasnich RD, Ross PD, Maclean CJ, Davis JW, Vogel JM. The relative strengths of osteoporotic risk factors in a prospective study of postmenopausal osteoporosis. In: Christiansen C, Johansen JS, Riis BJ, eds. Osteoporosis 1. Proceedings of the International Symposium on Osteoporosis, Denmark, 1987. Copenhagen: Osteopress ApS, 1987:394–7.
21. Kleerekoper M, Peterson E, Nelson D, Tilley B, Phillips E, Schork MA, Kuder J. Identification of women at risk for developing postmenopausal osteoporosis with vertebral fractures: role of history and single photon absorptiometry. Bone Miner 1989; 7:171–86.

Received for publication 11 January 1991
Accepted in revised form 9 July 1991