

Do clinical findings associate with radiographic osteoarthritis of the knee?

Angela A M C Claessens, Jan S A G Schouten, Frank A van den Ouweland, Hans A Valkenburg

Abstract

From a population survey of 2865 subjects, test characteristics of a number of clinical findings relating to knee osteoarthritis were calculated against the standard of radiographic diagnosis. The clinical findings included from the history were age, gender, current pain in the knee, swollen knee, pain in both hands, morning stiffness, osteoarthritis in any joint, pain or stiffness, or both, in knees or hips when rising from seated position, and pain in knees or hips while climbing stairs; and from the physical examination: Quetelet's index, Heberden's nodes, bony enlargement, palpable effusion, soft tissue swelling, limitation of knee function, pain with knee flexion, bony tenderness and, finally, the latex fixation test.

Of 18 clinical variables, all but Heberden's nodes, palpable knee effusion, pain in both hands, and latex fixation test showed a significant association after adjustment for age. Neither one single variable nor a combination could predict radiographic osteoarthritis of the knee with reasonable accuracy and thus be applicable in clinical practice. The x ray film, therefore, keeps its place in the diagnosis of knee osteoarthritis in general practice as well as in epidemiological research.

In epidemiological research diagnosis of osteoarthritis of the knee is based traditionally on the radiographic appearance of the joint. Osteoarthritis (OA) is judged according to Kellgren's criteria, described in the *Atlas of Standard Radiographs*.¹ In 1986, the subcommittee on classification criteria of osteoarthritis of the American Rheumatism Association (ARA) prepared criteria for the classification and reporting of OA of the knee. The proposed criteria were developed by a Delphi procedure² and subsequently tested in a group of patients with knee pain referred to a rheumatological clinic.³ Variables to construct the criteria were obtained from medical history, physical examination, and laboratory tests. The subcommittee presented several sets of criteria and indicated that one of these, a combination of findings merely from medical history and physical examination, could be applied in epidemiological research. Substitution of the x ray film by a small number of clinical findings might be helpful indeed in future surveys. Apart from the problem of precision or reproducibility of the clinical diagnosis OA,^{4,5} little is known about the association of clinical findings with radiographic OA. To test the value of single clinical and laboratory variables for the prediction of

radiographic diagnosis we studied a population based data set covering the relevant variables from medical history, physical and laboratory examinations, and radiography. By stepwise logistic regression analysis we investigated which combination of variables was most predictive of the radiographic diagnosis of knee OA.

Subjects and methods

The study data are derived from a population survey conducted in the Dutch town of Zoetermeer between 1975 and 1978. The prevalence of rheumatic diseases⁶ and other chronic conditions was investigated. A total of 13 614 inhabitants, aged 5 years and older, dwellers of two town districts, were invited to participate. The overall response was 78.2%. Standard antero-posterior weight bearing knee radiographs were taken from those aged 45 and older, irrespective of complaints. The study group thus comprised 2865 subjects (1320 men (46.1%) and 1545 women (53.9%)). The results are presented for the right knee only. Clinical and radiographic findings therefore refer to the same joint. A preliminary analysis of both joints showed no significant differences between the knees. From the data set all the variables relevant to diagnosis of knee OA were selected. The variables included all those finally present in the ARA subcommittee's study except crepitus and palpable warmth. The former was not investigated and the latter was found in only four knees.

Variables obtained from the history included age, gender, current pain in the knee, swollen knee joint, pain in both hands, morning stiffness of less than 30 minutes in arms or legs, or both, previous medical treatment for OA in any joint, pain in knees or hips, or both, when rising from a seated position, stiffness in knees or hips, or both, when rising from a seated position, and pain in knees or hips, or both, while climbing stairs. Subjects with no pain in any joint, including the spine, were not questioned specifically about the last three variables as we assumed that they would have given negative answers there.

Variables from the physical examination included Quetelet's index (in kg/m²), clinical Heberden's nodes at the right or left distal interphalangeal joints, bony enlargement, palpable effusion, soft tissue swelling, limitation of knee function, pain with knee flexion during examination, and bony tenderness. Function of the knee was scored on a scale ranging from 0 to 4. Grade 1 or more was regarded as limited

Department of
Epidemiology and
Biostatistics,
Erasmus University
Medical School,
Rotterdam,
The Netherlands
A A M C Claessens
J S A G Schouten
F A van den Ouweland
H A Valkenburg

Correspondence to:
Dr Frank A van den Ouweland,
Department of Epidemiology
and Biostatistics,
Erasmus University Medical
School, PO Box 1738,
3000 DR Rotterdam,
The Netherlands.

Accepted for publication
8 December 1989

function (grade 1 is minimal limitation of extension or flexion at physical examination). The clinical observations were made by six doctors; interobserver variability was reduced by combined three month training sessions.

From the laboratory tests available the latex fixation test was selected. A normal test result was defined as one with a titre $<1/20$. The cut off is lower than that applied for clinical use (1/640) in our laboratory.

Radiographs of the knees were studied without knowledge of the clinical findings. Radiographic OA was expressed on a five point scale.¹ In this study the diagnosis radiographic OA refers to grade 2 or more on the x ray film of the right knee (grade 2=definite osteophytes, absent or questionable narrowing of the joint space).

Prevalence, sensitivity, specificity, predictive value of a positive finding, likelihood ratio of a positive finding, and odds ratio are defined as follows. The prevalence figure indicates the percentage of subjects with a positive finding at the time of the examination. Sensitivity is the percentage of subjects with a positive finding from among all those who have radiographic OA; likewise, specificity is expressed by the percentage of subjects with a negative finding conditional on the absence of radiographic OA. The predictive value is the proportion of subjects with radiographic OA from among all those with a positive finding. The likelihood ratio of a positive finding expresses the chance that a positive finding is expected in a subject with radiographic OA, over that in one without radiographic OA.⁷ The odds ratio gives the ratio of the odds for a positive finding in subjects with radiographic OA, over the odds for a positive finding in persons without radiographic OA. Odds ratios are also given after adjustment for age as the prevalence of radiographic OA is strongly associated with age.

Combinations of variables may show a stronger association with radiographic OA than any single variable. To determine the best predictive combination of variables an initial choice was made from all available clinical

findings on the basis of their relevance as indicated by significance and by reported data. Then, for the total population, as well as for those with pain in the right knee at the time of study entry, the optimal combination of clinical findings was selected by a stepwise logistic regression analysis. The combination of variables selected by this procedure included for the population: age, gender, Quetelet's index, pain in knee or hips, or both, while rising from a seated position, bony enlargement, soft tissue swelling, and limitation of function; and for the group with knee pain: age, gender, Quetelet's index, pain in knee or hips, or both, while rising from a seated position, and bony enlargement. To test the value of these two most predictive combinations sensitivity and specificity were determined as follows. Each subject's individual variables were applied to the risk function (see Addendum) to calculate the predicted risk, or predicted probability of having radiographic OA. For each decile cut off point in the two distributions of risk-function outcomes sensitivity and specificity for either the whole population or the group with knee pain were calculated. These points are presented in a so called receiver-operator-characteristic curve (ROC curve)⁷ for both the population and the group with knee pain. Also, as radiographic OA is strongly associated with age, risk functions including only age as continuous variable were developed, again for both the whole population and the group with knee pain.

The BMDP statistical software package⁸ was used.

Results

Right knee radiographic OA was detected on the films of 564/2865 (19.7%) of the population: in 191/1320 (14.5%) men and in 373/1545 (24.1%) women. Three hundred and seventy one subjects (12.9%) in the population had pain in the right knee at the time of study: 100 (7.6%) men and 271 (17.5%) women. One hundred and thirty five subjects (4.7%) had both radiographic OA and current knee pain: 37

Table 1 Prevalence, sensitivity, specificity, predictive value of a positive finding, and likelihood ratio of a positive finding for clinical variables against the standard of radiographic knee osteoarthritis

Variable		Prevalence (%)		Sensitivity (%)		Specificity (%)		Predictive value (%)		Likelihood ratio	
		A†	B†	A	B	A	B	A	B	A	B
Pain during flexion	E‡	2	13	6	22	99	92	56	60	5.3	2.7
Function limitation	E	6	22	17	38	96	88	52	64	4.4	3.1
History of swelling	H	2	16	6	23	99	88	49	52	4.0	1.9
Swelling of soft tissue	E	0.5	3	1	3	100	97	43	36	3.7	1.0
Bony enlargement	E	11	28	26	50	93	85	47	65	3.6	3.2
Bony tenderness	E	3	15	7	22	98	89	44	53	3.2	1.9
Quetelet's index >30 kg/m ²	E	10	17	20	28	93	90	42	62	2.8	2.8
Previous OA	H‡	2	6	4	10	99	96	40	58	2.7	2.5
History of pain	H	13	—	24	—	90	—	36	—	2.3	—
Pain on rising from chair	H	18	65	29	73	85	40	33	41	2.0	1.2
Heberden's nodes	E	12	17	17	24	90	88	29	53	1.7	2.0
Stiff on rising from chair	H	22	68	32	72	81	34	29	38	1.7	1.1
Pain in climbing stairs	H	13	54	19	46	88	42	28	31	1.6	0.8
Morning stiffness	H	13	31	17	30	88	68	26	34	1.5	0.9
Palpable effusion	E	2	6	2	5	99	93	27	30	1.5	0.8
Pain in both hands	H	7	21	9	21	94	79	26	36	1.4	1.0
Female gender	H	54	73	66	73	49	27	24	36	1.3	1.0
Age >50 years	H	79	80	90	92	23	26	22	42	1.2	1.2
Latex test normal	L‡	73	72	73	72	27	28	20	37	1.0	1.0

*Variables are ordered by magnitude of likelihood ratio for the total population.

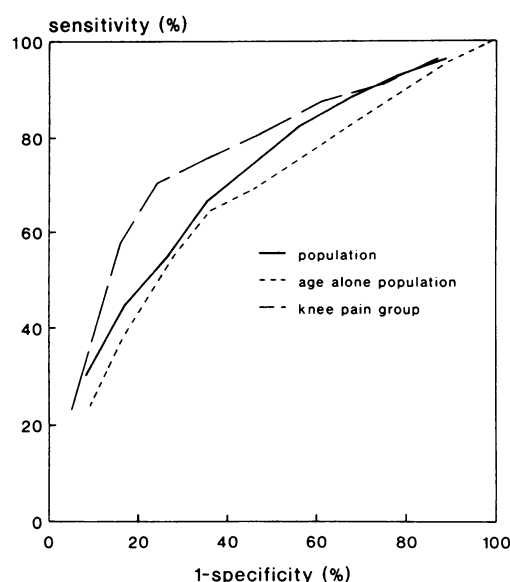
†A=total population; B=the group with pain in the right knee at study entry.

‡E=physical examination; H=history; L=laboratory test result.

(2.8%) men and 98 (6.3%) women. Table 1 lists the sensitivity, specificity, predictive value of a positive finding, and the likelihood ratio of a positive finding for the population and for the group with knee pain separately. Prevalences of age over 50 years and a normal latex test were more than 70% in the population. Some variables had a prevalence of less than 5%: swelling of the right knee, previous medical treatment for OA in any joint, palpable effusion, swelling of soft tissue, pain with movement of the knee at examination, and bony tenderness. The prevalences of other variables, except gender, varied between 6 and 22% of the population. Overall, high specificities were accompanied by low sensitivities. The predictive values of positive findings varied between 22 and 65% and showed no obvious improvement in the group with knee pain. Almost all likelihood ratios in the group with knee pain were lower than the corresponding values in the total population. The best single variables were pain during flexion, limitation of function, history of swelling, swelling of soft tissue, bony enlargement, and bony tenderness.

Table 2 lists the adjusted and unadjusted odds ratios with their 95% confidence intervals. The rank order by magnitude of the odds ratio point estimates is essentially unchanged after adjustment for age. Exceptions for the total population (column A) are the variables swelling of soft tissue (up by five places), pain climbing stairs (up by three places), Heberden's nodes (down by four places), bony tenderness (down by three places) and for the group with knee pain (column B) the variables swelling of soft tissue (up by five places), history of swelling (up by four places), pain on rising from chair (up by three places), history of OA (down by six places), female gender (down by four places), and Heberden's nodes (down by three places). Most variables when adjusted for age showed a small decrease in odds ratio. When the odds ratios were adjusted for age all were significant in the total population except the variables pain in right and left hand, palpable effusion, Heberden's nodes, and a normal latex test.

The ROC curve shows the test characteristics



Receiver-operator-characteristic (ROC) curves showing the characteristics of the combinations of clinical findings that are most predictive for the population, as well as for the group with knee pain at study entry, against the standard of radiographic osteoarthritis of the right knee. The ROC curve constructed by applying age alone in the risk function for the population is included in the plot. The lines connect nine points in the distribution of risk-function outcomes, for which sensitivity and specificity were calculated.

of the best combinations of clinical findings that were most predictive in the population against the standard of radiographic knee OA (figure). The best combination of clinical variables performed somewhat better in the group with knee pain: the ROC curve ascends towards the upper left corner of the plot. In a separate ROC curve the result when age alone was applied as a continuous variable in the risk function is shown in comparison with the combination of variables for the whole population.

Discussion

The answer to the question 'Do clinical findings associate with radiographic osteoarthritis of the knee?' is 'yes'. There is a significant association between radiographic OA and 14 of the 18 clinical findings studied. Also there is consistency in the rank order of variables according to the magnitude of likelihood ratios (table 1) and odds ratios (table 2) for the population as well as the group with knee pain. The strength of the associations of the different variables with radiographic OA is somewhat disappointing as no single clinical finding can accurately predict radiographic OA by means of its sensitivity and specificity, likelihood ratio, or adjusted odds ratio. Interestingly, when the odds ratios are arranged in order of magnitude after adjustment for age the variables Heberden's nodes and bony tenderness in the population column and Heberden's nodes, female gender, and previous OA in the knee pain column decrease in rank order. These indices being apparently related to age lose significance after adjustment for age.

In the analysis of variables, age is used in two different ways: firstly, as in the study of the

Table 2 Odds ratio, unadjusted and adjusted for age with 95% confidence intervals (CI) for the adjusted odds ratios

Variable		Odds ratio		Adjusted odds ratio (95% CI)	
		A*	B*	A	B
Pain during flexion	E*	5.5	3.1	4.3 (2.5 to 7.4)	2.1 (1.1 to 4.2)
Function limitation	E	5.1	4.3	3.1 (2.2 to 4.4)	2.3 (1.3 to 4.1)
History of swelling	H*	4.2	2.1	4.0 (2.4 to 6.8)	2.2 (1.2 to 4.0)
Swelling of soft tissue	E	3.1	1.0	4.2 (1.3 to 13.8)	1.7 (0.5 to 6.5)
Bony enlargement	E	4.5	5.5	2.8 (2.1 to 3.6)	3.2 (1.9 to 5.6)
Bony tenderness	E	3.4	2.2	2.5 (1.6 to 4.0)	1.7 (0.9 to 3.1)
Quetelet's index >30 kg/m ²	E	3.3	3.6	2.6 (1.9 to 3.8)	2.4 (1.3 to 4.4)
Previous OA	H	2.8	2.6	1.9 (1.1 to 3.5)	1.7 (0.7 to 4.5)
History of pain	H	2.8	—	2.6 (2.0 to 3.3)	—
Pain on rising from chair	H	2.4	1.8	2.2 (1.7 to 2.7)	2.0 (1.2 to 3.3)
Heberden's nodes	E	1.8	2.3	1.2 (0.9 to 1.6)	1.6 (0.9 to 3.0)
Stiff on rising from chair	H	2.0	1.3	1.7 (1.4 to 2.1)	1.2 (0.8 to 2.1)
Pain on climbing stairs	H	1.7	0.6	1.9 (1.4 to 2.4)	0.8 (0.5 to 1.3)
Morning stiffness	H	1.6	0.9	1.5 (1.1 to 1.9)	1.0 (0.6 to 1.6)
Palpable effusion	E	1.5	0.8	1.4 (0.7 to 2.8)	0.7 (0.3 to 1.8)
Pain in both hands	H	1.4	1.0	1.3 (0.9 to 1.9)	1.0 (0.6 to 1.8)
Female gender	H	1.9	1.0	1.7 (1.4 to 2.1)	0.7 (0.4 to 1.2)
Age >50 years	H	2.7	4.0	—	—
Latex test normal	L*	1.0	1.0	1.0 (0.8 to 1.3)	1.0 (0.6 to 1.6)

*For abbreviations see table 1.

ARA subcommittee³ by dividing subjects into those above and below 50 years and, secondly, as a continuous variable in the risk-function expression. The variable, age over 50 years, shows a high sensitivity, the highest of all, and a low specificity. This can be explained by the fact that our study is population based and includes only those over 45 years. Age is strongly associated with radiographic OA and, therefore, age alone was applied as continuous variable in the risk function, separately from the calculation of the combination of clinical findings, to show the eventual gain in predictive value by the clinical findings.

The combination of variables, judged by the position of its ROC curve (figure), is a better predictor of radiographic OA than is age alone. The ROC curve of the combination of variables in the group with knee pain also performs better than age alone in that group (not shown in the figure). The difference, however, is marginal in both groups and it implies that there is little gain when a composite of clinical findings is used to predict radiographic OA. Overall, the ROC curves are far from ideal: a clinically useful test characteristic should include at least one point in the extreme upper left corner of the ROC plot.

The most authoritative paper with which to compare our work is one published by the ARA subcommittee on classification criteria of osteoarthritis.³ The set of clinical criteria for knee OA in the ARA classification tree reached a sensitivity of 89% and a specificity of 88% against expert opinion as the standard. Consequently, the subcommittee concluded that clinical examination alone was a useful classification tool in epidemiological studies. If crepitus had been one of the variables in our analysis this index, eventually, might have improved the characteristics of the most optimal combination. It is very unlikely, however, that this would have led to a comparably good result for sensitivity and specificity. On the other hand, Quetelet's index—a strong predictor of knee OA⁹⁻¹¹—was included in our analysis, but not in the ARA subcommittee's classification tree. Our results confirm those of a preliminary report by Spector *et al*, who calculated a sensitivity of 59% and a specificity of 72% for two clinical signs of OA against x ray diagnosis as the standard.⁴ The figures were obtained in a sample of 41 women drawn from the general population. These characteristics match well with the ROC curve for the total population presented here. We differ from Spector, however, as we believe that radiographs are still necessary for ascertaining OA in epidemiological studies.

Moreover, our study shows that even in the group with knee pain at the time of the survey (subjects more likely to be general practitioners' patients) clinical findings are a poor classification tool. In general practice also, an x ray examination will be necessary to diagnose knee OA. We conclude that a number of findings from medical history, physical examination, and laboratory tests are associated with radiographic knee OA; nevertheless, the strength of association is insufficient to predict radiographic OA.

In fact, the best combination of variables proves to be only slightly better than age alone to predict radiographic OA in the population. Clinical findings, either separately or in combination, cannot suffice as a diagnostic tool for knee OA and cannot be an alternative to x ray examination.

Addendum

The most predictive variables for radiographic OA in both the population and the group with knee pain were selected by a stepwise logistic regression analysis. These variables define the risk function.¹² The risk function can be expressed as follows:

$$Y = \frac{\exp(\alpha + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \dots + \beta_n \cdot X_n)}{1 + \exp(\alpha + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \dots + \beta_n \cdot X_n)}$$

where Y = predicted probability of having radiographic OA; α = constant; β_n = coefficient for variable n; X_n = independent variable n (binary for all except Quetelet's index and age, which are continuous variables).

The selected clinical findings represent the independent X variables in the risk function. For every subject each clinical variable, either binary or continuous, has an individual value. The outcome of the risk function or the dependent Y value ranges from 0 to 1 for each respondent. To calculate the sensitivity and the specificity a certain Y value has to be chosen as a cut off point. Above this Y value the diagnosis radiographic OA is assumed to be present and below it absent. The series of 2865 Y values was split into deciles by defining nine cut off points: decile 1 counts the 10% lowest Y values and thus subjects with the smallest chance of showing radiographic OA, and so on for each decile. For each cut off point sensitivity and specificity were calculated.

- 1 Kellgren J H, Jeffrey M R, Ball J. *Atlas of standard radiographs. The epidemiology of chronic rheumatism*. Vol II. Oxford: Blackwell Scientific, 1963.
- 2 Altman R D, Meenan R F, Hochberg M C, *et al*. An approach to developing criteria for the clinical diagnosis and classification of osteoarthritis: a status report of the American Rheumatism Association diagnostic subcommittee on osteoarthritis. *J Rheumatol* 1983; **10**: 180-3.
- 3 Altman R, Asch E, Bloch D, *et al*. Development of criteria for the classification and reporting of osteoarthritis. *Arthritis Rheum* 1986; **29**: 1039-49.
- 4 Spector T D, Brown P, Hart D, *et al*. The clinical diagnosis of osteoarthritis in the general population. *Br J Rheumatol* 1988; **27** (suppl 2): 29.
- 5 McAlindon T, Dieppe P. Osteoarthritis: definitions and criteria. *Ann Rheum Dis* 1989; **48**: 531-2.
- 6 van Saase J L C M, van Romunde L K J, Cats A, Vandenbroucke J P, Valkenburg H A. Epidemiology of osteoarthritis: Zoetermeer survey. Comparison of radiological osteoarthritis in a Dutch population with that in 10 other populations. *Ann Rheum Dis* 1989; **48**: 271-80.
- 7 Sackett D L, Haynes R B, Tugwell P. *Clinical epidemiology: a basic science for clinical medicine*. Boston/Toronto: Little, Brown, 1985.
- 8 Dixon W J, Brown M B, Engelman L, *et al*. eds. *BMDP statistical software manual*. Berkeley: University of California Press, 1988.
- 9 van Saase J L C M, Vandenbroucke J P, van Romunde L K J, *et al*. Osteoarthritis and obesity in the general population: a relationship calling for an explanation. *J Rheumatol* 1988; **15**: 1152-8.
- 10 Felson D T, Anderson J J, Naimark A, *et al*. Obesity and knee osteoarthritis. The Framingham study. *Ann Intern Med* 1988; **109**: 18-24.
- 11 Anderson J J, Felson D T. Factors associated with osteoarthritis of the knee in the first National Health and Nutrition Examination Survey (HANES I). *Am J Epidemiol* 1988; **128**: 179-89.
- 12 Hennekens C H, Buring J E. *Epidemiology in medicine*. Boston/Toronto: Little, Brown, 1987.