Assessing the prevalence of hand osteoarthritis in epidemiological studies. The reliability of a radiological hand scale

S Kessler, P Dieppe, J Fuchs, T Stürmer, K P Günther

Abstract

Objective—The hands are often involved in the osteoarthritic disease process. A radiological grading scale is presented, derived from a published atlas, to assess the prevalence of hand osteoarthritis (OA) involvement in clinical and epidemiological studies and its reproducibility is studied.

Methods—This hand scale is based on the radiological feature “joint space narrowing”, which represents the macromorphological process of cartilage loss. Osteophytes and sclerosis are less important unless seen in conjunction with joint space narrowing. Nine individual joints per hand (four proximal interphalangeal joints (PIP), four distal interphalangeal joints (DIP), first carpometacarpal joint (CMC-1)) are scored dichotomously for the presence of OA. To save time and to increase reliability a severity grading of radiological features is not performed. To determine inter-rater and intra-rater reliability of the individual joints and the presence of OA in two separate joint groups (≥ 2 PIP or DIP and at least one CMC-1, used to define “generalised OA” in the ongoing Ulm Osteoarthritis Study) 50 pairs of anteroposterior hand radiographs were read by two investigators twice within one month. The κ coefficient was calculated to quantify the strength of associations.

Results—On average five minutes were needed to score one hand radiograph. Both raters were able to reproduce their own readings in all individual joints and for the presence of OA in two separate joint groups after one month. Reliability was highest for the PIP joints (κ : 0.56–1.00) it was slightly lower for the DIP joints (0.38–0.87), for the CMC-1 joints (0.58–0.69) and for OA in two separate joint groups (0.54). The values for inter-rater agreement were good as well, κ coefficients ranged from 0.52 to 0.92.

Conclusion—This grading scale was shown to be reliable within and between readers for all the individual joints as well as for the presence of OA in two separate joint groups. Scoring a limited number of joints dichotomously makes this scale efficient and therefore useful for clinical and epidemiological trials, when dealing with large patient samples.

(Ann Rheum Dis 2000;59:289–292)

The hand is a frequent site of disease involvement in osteoarthritis (OA).¹ It has been estimated that more than 70% of people aged 65 years or older² are affected. While many previous studies on the prevalence of hand OA were based on the presence of Heberden's knodes^{3–5} more recently performed investigations use radiological changes to grade the disease.^{6–9}

The instrument that was mainly applied in the past to score OA in hands was the Kellgren and Lawrence score.^{6,7} This overall score distinguishes five degrees of severity of OA according to the presence of the radiological features: osteophytes, joint space narrowing (JSN), subchondral sclerosis, cysts and flattening of condyles. Subsequent investigations identified certain deficiencies of that score and its appropriateness in clinical and epidemiological studies was questioned.^{8,9} One important point of criticism is the significance of osteophytes as the main radiological feature in scoring the severity of disease.⁸ Identification of OA can be performed only if definite osteophytes are present. Another point of criticism is the moderate reliability of this score.^{8,9}

These deficiencies have led to the development of other radiological scores that grade individual features of OA separately and independently from the presence of osteophytes. New radiological scores on hand OA were created and have been proved to be reliable.^{8,9} However, these scores assess these features in great detail and scoring is therefore time consuming. This might be a disadvantage for clinical and epidemiological studies dealing with large patient samples.

We present a hand scale based on a simple dichotomy whether or not OA is present by grouping several radiographic features, to assess the prevalence of hand OA and to measure the number of joints being included in the disease process. As JSN resulting from cartilage loss is the main radiological feature of OA, our scale is based on the presence of that feature. Other radiological parameters are only used in conjunction with JSN.

Methods

SUBJECTS AND RADIOGRAPHS

Anteroposterior radiographs of both hands of 50 patients from the Ulm Osteoarthritis Study¹⁰ were chosen for this investigation. Subjects ranged in age between 51 and 79 years. The Ulm Osteoarthritis Study is a cross sectional and longitudinal investigation evaluating radiographic and clinical patterns of

Orthopaedic
Department (RKU),
University of Ulm,
Oberer Eselsberg 45,
D-89081 Ulm,
Germany
S Kessler
J Fuchs
K P Günther

Health Services
Research
Collaboration,
University of Bristol,
Bristol
P Dieppe

Department of
Epidemiology,
University of Ulm
T Stürmer

Correspondence to:
Dr Kessler

Accepted for publication
10 December 1999

Table 1 Radiographic osteoarthritis according to the new hand scale for the joints of the right hand (second rating in parentheses) (n=50)

	CMC-1	PIP-2	PIP-3	PIP-4	PIP-5	DIP-2	DIP-3	DIP-4	DIP-5
Rater 1	17 (21)	9 (9)	11 (9)	8 (8)	8 (9)	35 (32)	27 (27)	13 (14)	26 (30)
Rater 2	18 (25)	7 (8)	8 (7)	6 (5)	7 (7)	25 (22)	18 (16)	9 (14)	20 (28)

CMC-1 = first carpometacarpal joint, PIP = proximal interphalangeal joints, DIP = distal interphalangeal joints.

Table 2 Radiographic osteoarthritis according to the new hand scale for the joints of the left hand (second rating in parentheses) (n=50)

	CMC-1	PIP-2	PIP-3	PIP-4	PIP-5	DIP-2	DIP-3	DIP-4	DIP-5
Rater 1	18 (26)	10 (10)	10 (10)	8 (9)	12 (13)	32 (33)	27 (28)	19 (22)	28 (34)
Rater 2	20 (31)	8 (7)	7 (6)	8 (6)	7 (9)	25 (23)	23 (20)	14 (20)	22 (30)

Abbreviations as in table 1.

hand, hip and knee OA in patients with advanced OA of a large weight bearing joint in south west Germany.

THE HAND SCALE

Hand radiographs were scored for the presence of OA in nine finger joints of each hand: the four distal interphalangeal joints (DIP), the four proximal interphalangeal joints (PIP) and the first carpometacarpal joint (CMC-1). These joints were chosen because they have been shown to be most frequently affected in the OA disease process of the hand¹¹ and have been previously used to define generalised OA.¹⁰ Using the atlas published by Altman *et al*¹² as guideline whether or not significant JSN, osteophytes or sclerosis were present, the scoring was performed in the following manner: if JSN grade two or more or JSN grade one and either of sclerosis or osteophytes grade two or more were present, the joint was regarded as being affected. OA was not diagnosed if osteophytes or sclerosis without JSN were observed. The number of affected joints of both hands was counted separately for interphalangeal joints (0–16) and for the CMC-1 joints (0–2).

TRAINING OF RATERS

After two orthopaedic surgeons (SK, JF) had familiarised themselves with the OA hand scale and the atlas of Altman *et al*¹² they subsequently scored 50 hand radiographs (25 patients) in three two-hour training sessions. For that purpose, radiographs were randomly selected from the Ulm osteoarthritis study population.

Both investigators compared their results and discussed them until consent was achieved.

RATING OF RADIOGRAPHS

Both anteroposterior hand radiographs of another 50 patients of the Ulm Osteoarthritis Study (100 hands) were randomly selected. After blinding for individual data (date of birth, name and sex), all radiographs were independently read by both raters. To determine intra-rater reliability, the rating was repeated by both investigators one month later.

DATA ANALYSIS

κ Statistics were used to quantify the inter-rater and intra-rater agreement of both raters for each individual joint as well as for the presence of OA in two or more finger joints (PIP or DIP) and at least one CMC-1 joint (used to define generalised OA in the Ulm Osteoarthritis Study in patients with hip or knee OA). All analyses were performed with SAS (Statistical Analysis System, Version 6.12, SAS Inc, Cary, NC).

Results

EFFICIENCY

The average time required to read one anteroposterior hand radiograph with the new hand scale was five minutes. The average time that was necessary using Kallman's⁸ or Lane's⁹ scale ranged between 10 and 15 minutes per hand.

PREVALENCE OF OA ACCORDING TO THE HAND SCALE

Table 1 and table 2 show the prevalence of radiographic OA according to the hand scale for the nine joints of each hand. Results are presented separately for the right and the left hand as well as for both readings of both raters.

The most frequently affected joint in the right and the left hand in all readings was the second DIP joint, followed by the third and the fifth DIP joint.

INTER-RATER AND INTRA-RATER RELIABILITY OF INTERPHALANGEAL JOINTS, CMC JOINTS AND PRESENCE OF OA IN TWO SEPARATE JOINT GROUPS Figures 1 and 2 show the intra-individual agreement of both readers. Both raters were able to reproduce their own readings in all the joints after one month. Reliability was generally higher for the PIP joints than for the DIP joints and slightly lower for the CMC-1 joints. Results of the inter-rater reliability are shown in figure 3. Raters were able to reproduce each

Figure 1 Intra-rater reliability of rater 1 (κ statistics) for the right (first column) and the left (second column) hand. (CMC-1 = first carpometacarpal joint, PIP = proximal interphalangeal joints, DIP = distal interphalangeal joints).

Figure 2 Intra-rater reliability of rater 2 (κ statistics) for the right (first column) and the left (second column) hand. Abbreviations as in figure 1.

Figure 3 Inter-rater reliability of rater 1 and 2 (κ statistics) for the right (first column) and the left (second column) hand. Abbreviations as in figure 1.

others readings in an acceptable way. Agreement between raters ranged from 0.52 (DIP-2, right hand) to 0.92 (CMC-1, left hand).

The reliability of presence of OA in two separate joint groups was acceptable as well. Table 3 lists the results

Discussion

In this study we present a radiological scale based on the atlas of Altman *et al*¹² to measure the prevalence of hand OA in clinical and epidemiological study samples, which is short, time efficient and as reliable as already existing scores. Our scale is a simple dichotomy whether or not OA is present to assess prevalence and to measure the number of joints being included in the disease process. Reasons for creating such a short scoring system are limitations of already existing radiological grading systems like the Kellgren and Lawrence score^{6,7} and newer parameters, for example, published by Kallman *et al*⁸ or Lane *et al*.⁹ These limitations include the number of joints being graded and so the amount of time needed for the procedure, especially in large patient samples.

The proposed hand scale is based on the presence of JSN, as the most important individual radiological feature in OA. Other

radiological features are regarded to be less important in this index, unless in conjunction with JSN. The background of this index, focusing on the presence of JSN is in contrast with the idea of the Kellgren and Lawrence score,⁶ which has accepted osteophytes to be the dominant radiological feature of OA. In hand OA, investigators argue that finger joints often appear narrowed and sclerotic on radiographs without showing significant osteophytes.⁸ We had the same impression, when reading the anteroposterior hand radiographs within the Ulm Osteoarthritis Study.¹⁰

Scoring features dichotomously makes our scale easy to handle and less time consuming (on average five minutes for one radiograph). An additional effect of the dichotomous reading seems to be an increase in the reliability of the scale. Within this scale only nine individual joints per hand are scored (eight IP joints and the CMC-1 joint). We have selected these joints, because they have shown to be the most often affected ones¹¹ in the OA disease process and those joints have been used in conjunction with knee or hip OA to decide whether or not generalised OA is present.^{10,13,14,15}

Other investigators like Kallman *et al*⁸ or Lane *et al*⁹ have included the scaphotrapezoid joint into their radiographic grading scores, a joint that we have tested to be difficult and not reliable enough to score, at least within our scale that is based on JSN, because of the oblique projection of this feature on anteroposterior hand radiographs. Including further hand or finger joints certainly would bring additional information but judging the benefit by doing this, the disadvantage losing the time

Table 3 Inter-rater and intra-rater agreement for the presence of osteoarthritis (OA) in two separate joint groups (≥ 2 finger joints (PIP or DIP) and at least 1 CMC-1 joint)

	Inter-rater reliability (κ)	Intra-rater reliability (rater 1) (κ)	Intra-rater reliability (rater 2) (κ)
Presence of OA in two separate joint groups	0.73	0.54	0.35

Abbreviations as in table 1.

efficiency of our scale in handling large patient samples is seen more severe.

Within the training period both readers tested in a limited number of hand radiographs (n=10) the amount of time, necessary to read films with other scales, having included more joints. On average more than double the time (10–15 minutes) was necessary when compared with our hand scale.

Our instrument was reliable between and within raters for all the individual joints (figs 1, 2 and 3). A comparison with the results of the Kellgren and Lawrence score^{6,7} or with those of Kallman *et al* and Lane *et al*^{8,9} is not possible because of the nature of our assessment. We have measured the reliability of scoring an individual joint whether or not OA is present, whereas other authors have measured the reliability of individual radiological features of OA. In our scale, reliability was highest for PIP joints and slightly lower for DIP joints and for CMC-1 joints.

Regarding the frequencies of joint affection, we registered an increasing prevalence of OA in CMC joints but not in interphalangeal joints within all second readings of both raters. We think that this phenomenon is a result of the ongoing learning process in scoring hand joints, especially CMC joints, which are, possibly because of a slightly oblique projection of the joint space on anteroposterior radiographs, more difficult to score when compared with interphalangeal joints.

In summary, our hand scale has been shown to be reliable within and between readers, for all individual joints as well as the presence of OA in two separate joint groups. Scoring a limited number of joints dichotomously makes

this scale very efficient and therefore interesting for investigations on hand OA in large patient samples.

- 1 Hochberg MC. Epidemiology of osteoarthritis: current concepts and new insights. *J Rheumatol* 1991;18 (suppl 27):4–6.
- 2 Van Saase JL, Van Romunde LKJ, Cats A, Vandenbroucke JP, Valkenburg HA. Epidemiology of osteoarthritis: Zoetermeer survey. *Ann Rheum Dis* 1989;48:271–80.
- 3 Adams R. *A treatise on rheumatic gout*. London: Churchill, 1857.
- 4 Fox RF, Van Breeman J. *Chronic rheumatism: causation and treatment*. London: J A Churchill, 1934:191.
- 5 Goldthwaite JE. The differential diagnosis and treatment of the so-called rheumatoid disease. *Boston Med Surg J* 1904;151:529–34.
- 6 Kellgren JH, Lawrence JS. Radiological assessment of osteoarthritis. *Ann Rheum Dis* 1957;16:494–501.
- 7 Kellgren JH, Lawrence JS. *Atlas of standard radiographs: the epidemiology of chronic rheumatism*. Vol 2. Oxford: Blackwell Scientific Publications, 1963.
- 8 Kallman DA, Wigley FM, Scott WW, Hochberg MC, Tobin JD. New radiographic grading scales for osteoarthritis of the hand. *Arthritis Rheum* 1989;32:1584–91.
- 9 Lane NE, Nevitt MC, Genant HK, Hochberg MC. Reliability of new indices of radiographic osteoarthritis of the hand and hip and lumbar disc degeneration. *J Rheumatol* 1993;20:1911–18.
- 10 Guenther KP, Stürmer T, Sauerland S, Zeissig I, Sun Y, Kessler S, et al. Prevalence of generalized osteoarthritis in patients with degenerative disease of hip and knee joints: the Ulm osteoarthritis study. *Ann Rheum Dis* 1998;57:717–23.
- 11 Egger P, Cooper C, Hart D, Doyle DV, Coggon D, Spector T. Patterns of joint involvement in osteoarthritis of the hand: the Chingford study. *J Rheumatol* 1995;22:1509–13.
- 12 Altman RD, Hochberg M, Murphy W, Wolfe F, Lequesne M. Atlas of individual radiographic features in osteoarthritis. *Osteoarthritis Cartilage* 1995;3 (suppl A):3–70.
- 13 Hart DJ, Spector TD, Egger P, Coggon D, Cooper C. Defining osteoarthritis of the hand for epidemiological studies: the Chingford Study. *Ann Rheum Dis* 1994;53:220–3.
- 14 Hirsch R, Lethbridge-Cejku M, Scott WW Jr, Reichle R, Plato CC, Tobin J. Association of hand and knee osteoarthritis: evidence for a polyarticular disease subset. *Ann Rheum Dis* 1996;55:25–9.
- 15 Stürmer T, Sun Y, Sauerland S, Zeißig I, Günther KP, Brenner H. Serum cholesterol and osteoarthritis. The baseline examination of the Ulm Osteoarthritis Study. *J Rheumatol* 1998;25:1827–32.