

Development of a new radiographic scoring system using digital image analysis

J E DACRE, J S COPPOCK, K E HERBERT, D PERRETT,
AND E C HUSKISSON

From the Department of Rheumatology, St Bartholomew's Hospital, West Smithfield, London EC1A 7BE

SUMMARY A scoring system using computerised analysis of digital stored images of knee radiographs has been developed. Measurement is based on the assessment of joint space size. It is sensitive, rapid, and reproducible. Plain radiographs are positioned on an acetate grid; a computer generated grid is superimposed on a digital image of the radiograph viewed on a closed circuit television monitor and the joint space measured automatically. Area and distance have been assessed; area measurements are more reproducible. Application of microcomputer based digital image analysis to radiological scoring systems is an important step in understanding the nature and progression of arthritis.

An estimation of joint damage and its progression is important in assessing the outcome of arthritis. Slow acting antirheumatic drugs may decrease the rate of radiological progression in rheumatoid arthritis and improve markers of disease activity.^{1,2} Non-steroidal anti-inflammatory drugs may worsen joint destruction as has been suggested by animal studies³ and reported in man.^{4,5} Radiological scoring systems may be used to substantiate these findings. Current systems, based either on verbal descriptions (such as normal, mild, moderate, and severe) or matching to standard *x* rays,^{6,7} are inherently insensitive. Sharp and colleagues have improved these methods with a complex, time consuming system of analysis of erosions.⁸ There is increasing interest in computer based *x* ray analysis as shown by Buckland-Wright *et al* who used microfocal radiography⁹ and Gaydecki and colleagues who assessed changes in bone shape using digital image analysis.¹⁰

Our hypothesis is that joint space narrowing is the most direct marker of joint destruction in osteoarthritis; osteophyte formation and periarticular sclerosis are related but secondary phenomena.¹¹ This is supported by Altman *et al*, who measured disease progression in knee radiographs.¹²

An ideal radiological scoring system should be sensitive to facilitate early detection of change. Assessment of joint destruction depends on the quality of the *x* ray and the measurement. Here our primary concern is the quality of the measurement.

Radiographic quality is limited by such factors as the position of the patient, rotation, the distance from the *x* ray source, the penetration of the film, and the signal to noise ratio of the image.

Quality of measurement in a digital analysis system is limited by the minimum data capture area—that is, the size of a pixel. Measurement of areas is likely to be more accurate than measurement of distance as more information is analysed. Variability is likely to be small and reproducibility good.

The ideal system should be inexpensive, operator independent, and easy to use. It should rapidly evaluate routine radiographs from any rheumatology or radiology department. In an attempt to realise these objectives we examined computerised analysis of digitised images of knee *x* rays to measure joint space.

Methods

RADIOGRAPHS

Posteroanterior *x* rays of the knee were used to evaluate the system. Films of a radiologically normal knee and one with definite radiological evidence of osteoarthritis—that is, asymmetrical joint space narrowing with osteophyte formation, were used.

HARDWARE

Knee radiographs were viewed on a conventional *x* ray viewing box (type No 4076-7-84, Exal X ray accessories Ltd, Mark Road, Hemel Hempstead) which was sited in a fixed position away from direct

Accepted for publication 2 July 1988.

Correspondence to Dr J E Dacre, Department of Rheumatology, St Bartholomew's Hospital, West Smithfield, London EC1A 7BE.

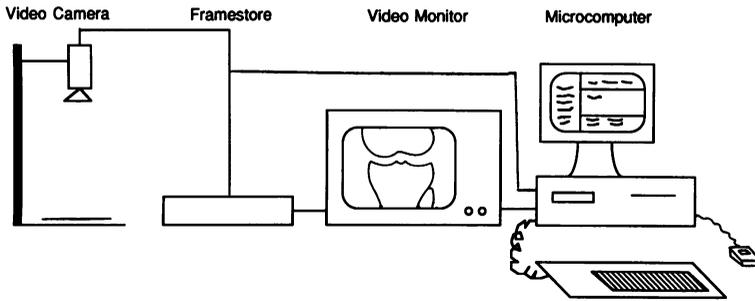


Fig. 1 Diagrammatic representation of the digital analysis system.

light, beneath a high quality black and white VHS video camera (Hitachi cctv model No HV720K). A spatial resolution of 512×512 pixels was available with six bit grey level resolution (64 levels). The analogue video signal was converted to digital form in a frame store device (micro eye) and displayed on a closed circuit Ikegami television monitor (model No PM-127). The stored image was edited and

analysed by a Victor 286 microcomputer (Fig. 1) (512K Ram 20 megabyte hard disk with mouse).

All video, computer equipment, and software were supplied by Digithurst Ltd, Royston, Herts, UK.

SOFTWARE

Modifications were made to a digital analysis

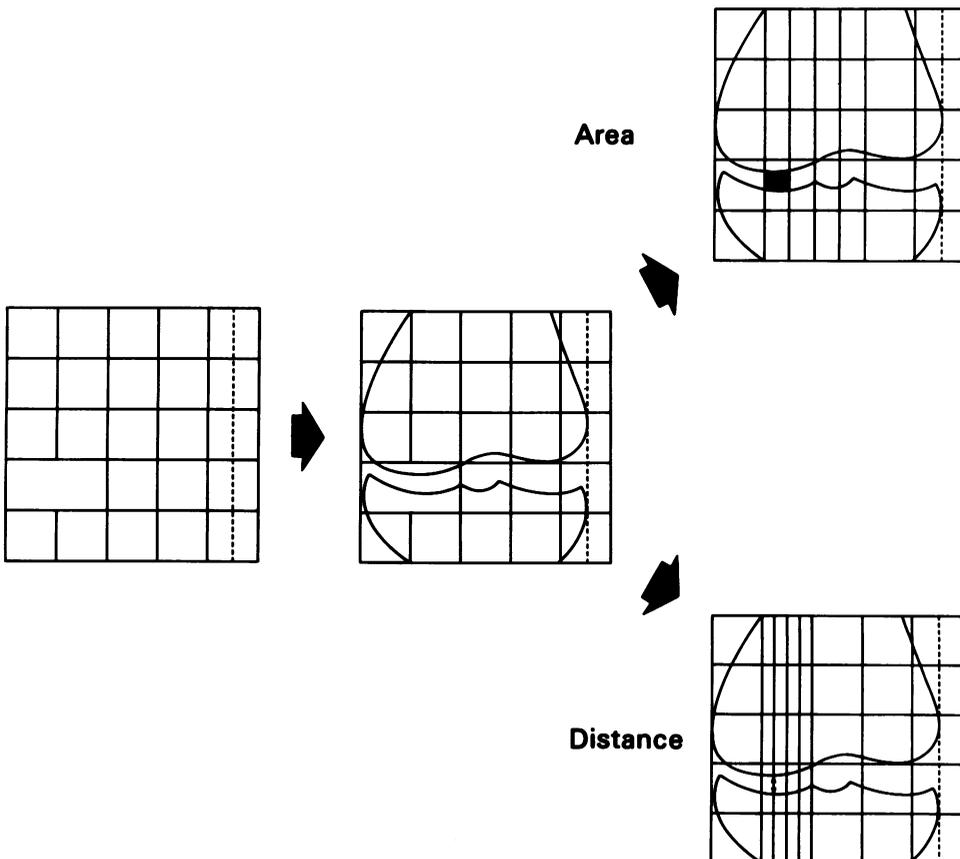


Fig. 2 Template grid showing positioning of the x ray, with measurement of area and distance.

program. A series of measuring grids consisting of a set of parallel vertical lines, each one labelled with a letter, was created.

SCORING SYSTEM

A standard grid of 5×5 4 cm squares, carefully drawn on a clear acetate sheet, was used (the 'template'). The radiograph was positioned over this template (Fig. 2). To measure areas a computer generated grid with lines 1 cm apart was superimposed on the template image over the lateral joint space. An area of joint space 1 cm wide within the grid was outlined using a mouse to define the edge. A minimum of seven points was used. The number of pixels within the area outlined was summed.

Distances across the lateral joint space were measured at a fixed point after superimposition of a computer generated grid composed of vertical lines 0.5 cm apart. Measurements were made along the vertical lines of the grid by positioning the mouse at the top and bottom of the joint space margin. The number of pixels—that is, the distance between the two joint margins, was summed (Fig. 2).

COMPARISON WITH STANDARD FILMS

Comparisons were made with our own standard films of osteoarthritis of the knee. These are graded from 0 to 5, ranging from normality (score 0) to severe joint destruction (score 5), by considering joint space narrowing.¹³ Area and linear measurements were made on each standard *x* ray.

Results

Complete assessment of a film, including positioning, image capture, measurement of area, and distance, took less than two minutes.

THRESHOLD

To assess an area within a digital image the whole of the area must be counted as black or white. This was assured by measuring the area of joint space at an increasing threshold value—that is, the value above which the intensity of the image is counted as black and below which, white. Figure 3 shows the increase in measured area with increasing threshold. With a relative threshold of above 20 the whole area of joint space outlined was measured.

CALIBRATION

The digital image on the screen was calibrated by placing a sheet of high quality graph paper beneath the camera and measuring a series of distances and areas. The number of pixels per cm was 42.3, so the pixel size was estimated at 0.24 mm. The variability of one point positioning was determined by

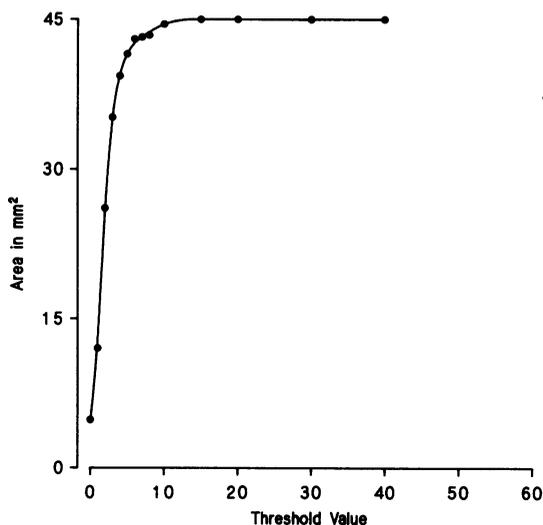


Fig. 3 Area measurement related to increasing threshold value.

measurement of the coordinates of the cursor on repeated attempts to superimpose a dot on the screen. The mean value for the coordinates was $x=179.9$ (coefficient of variation 0.17%), $y=244.9$ (coefficient of variation 0.30%), $n=10$. Mean distance measurements differed from the nominal length by a maximum of 1%, which was constant over the range studied, with a coefficient of variation of less than 5.5%. Mean area measurements differed from the nominal area by a maximum of 1.3%; the percentage deviation was constant over the range studied. Reproducibility of determination of square areas had coefficients of variation of less than 5% in all cases. The area of an actual joint space is within this range. Table 1 gives the results of the calibration experiments.

Table 1 Calibration of area and linear measurement ($n=10$)

Actual area (mm^2)	Mean area measured (mm^2)	Coefficient of variation (%)
25	25.6	4.3
50	50.3	2.7
100	101.1	1.7
200	202.6	1.8

Actual distance (mm)	Mean distance measured (mm)	Coefficient of variation (%)
2	2.02	5.5
4	3.97	4.9
6	5.99	2.1
8	7.99	2.4
10	10.02	1.1
12	11.96	1.8

Table 2 Comparison of methods for evaluation of joint space (n=10)

Method	Normal knee		Knee with osteoarthritis	
	Mean value (mm ²)	Coefficient of variation (%)	Mean value (mm ²)	Coefficient of variation (%)
No grid	44.1	7.0	26.5	10.8
Computer generated grid	44.6	3.5	26.4	2.8
Computer generated grid with 4 square template	45.5	4.0	34.8	3.2
Computer generated grid with 16 square template	44.2	1.8	25.2	3.5
Computer generated grid with 20 square template	45.1	1.8	24.6	1.6

Table 3 Comparison of area measurement of joint space with image inversion

Inversion value	Normal knee		Knee with osteoarthritis	
	Mean value (mm ²)	Coefficient of variation (%)	Mean value (mm ²)	Coefficient of variation (%)
50	44.1	2	26.0	8
40	43.4	3	25.0	8
30	45.7	4	25.5	6
20	43.8	8	25.9	5
10	43.3	3	23.6	3
8	43.3	3	24.0	6
6	44.0	3	24.7	5
4	44.3	2	24.9	6
2	46.7	3	24.9	3

DEVELOPMENT OF THE SCORING SYSTEM

Initially, measurements of joint space were made from the unadulterated image. To improve repeatability a series of grids was investigated. The first

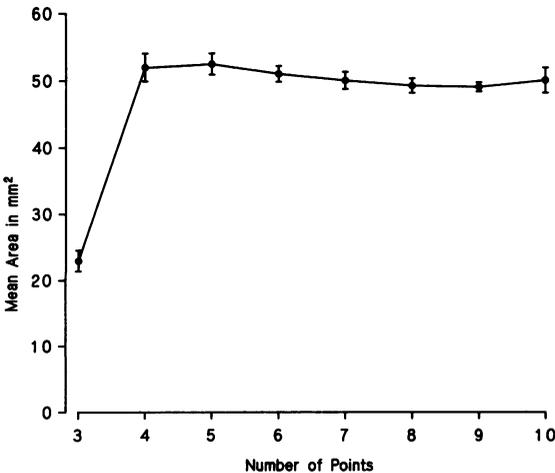


Fig. 4 Number of points required for accurate area measurement. Bars indicate one SD.

measurement used a computer generated grid alone—a series of labelled vertical lines set 1 cm apart. As the image was frozen accurate positioning of the film was a problem. Further assessment was carried out by superimposition of the computer generated grid onto a digitised image of a coarser grid, drawn on a clear acetate sheet, placed beneath the x ray and viewed through it. The x ray was aligned before frame storage. Several different sizes of computer generated grid and template were evaluated. The final version comprised a 5x5 grid of 4 cm squares and a computer generated grid with lines 1 cm apart. Areas of lateral joint space were measured within this. Each measurement was repeated 10 times on a normal radiograph and on one showing moderate joint space narrowing. Table 2 shows the results of the mean area measurement and the coefficient of variation using different templates. Figure 2 includes a diagrammatic representation of the template.

In addition, the image was evaluated inverted and in several degrees of partial inversion to assess enhancement. Joint space size was reported (n=10) and the mean and coefficient of variation calculated (Table 3). There was no advantage in image inversion.

Table 4 Interobserver and intraobserver reproducibility of joint space measurement (n=60)

Interobserver		Intraobserver	
Correlation coefficient (r)	p Value*	Correlation coefficient (r)	p Value*
<i>Area measurement</i>			
0.97	NS	0.98	NS
0.96	NS	0.99	NS
<i>Linear measurement</i>			
0.95	0.03	0.98	NS
0.96	0.02	0.95	NS

*Paired *t* test; NS=not significant.

ACCURACY OF POSITIONING

The radiograph was placed on the viewing box with the medial border of the femur over a vertical dotted line near the edge of the template and the middle of the lateral tibial spine over the third line away (Fig. 2). To assess positioning and repositioning the x ray was removed from the viewing box and resited for each measurement (n=10). The coefficient of variation of measurements taken in this way was 1.3% for area measurement and 1.4% for distance measurement.

AREA MEASUREMENT

Area measurement was dependent on the number of points used to outline the area. This was calculated by repeated measurement (n=10) using an increasing number of points (Fig. 4). A minimum of seven points was needed.

INTEROBSERVER AND INTRAOBSERVER RELATIONS

Sixty knee radiographs were measured independently by two observers on two occasions. Interobserver and intraobserver correlations were calcu-

lated. Table 4 shows significant correlations between and within observers in all cases, with *r* values of above 0.9. There were no significant differences in area measurement between or within observers with a paired *t* test. Distance measurement showed that correlations were significant, with high *r* values in all cases ($r > 0.09$), but there were significant differences between two observers when a paired *t* test was used. This suggests that area measurement is more reproducible.

LEARNING TIME

An untrained observer, unfamiliar with x ray analysis, was shown how to use the computer. After one hour's tuition 60 repeated measurements were made on one radiograph. A mean value of 58.9 mm² with a coefficient of variation of 1.3% was achieved for area measurement and a value of 5.6 mm with a coefficient of variation of 1.3% for linear measurement. These mean results were not significantly different from those of an operator familiar with the technique (area 59.0 mm²; distance 5.5 mm).

COMPARISON WITH STANDARD FILMS

In our first study of joint space narrowing in osteoarthritis we selected a series of standard films to show worsening grades of the disease. These standard films were measured with our new technique as an illustration of its usefulness in the clinical setting. Figure 5 shows the results obtained.

Discussion

We have described a radiological scoring system using computerised analysis of a stored digital image of knee radiographs. It was accurate, rapid, and easy to use, with good reproducibility. It was necessary to use a hand drawn template grid to position the film reproducibly. Using this, and a

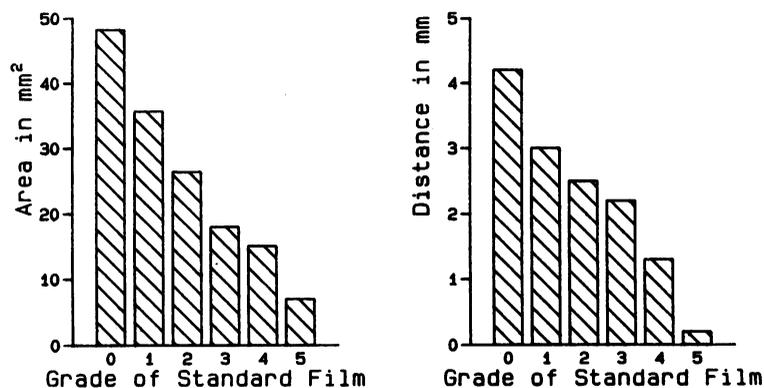


Fig. 5 Measurement of our own standard films of osteoarthritis of the knee.

computer generated grid of vertical lines, we found small coefficients of variation in the measurement of normal and osteoarthritic joint spaces. To measure an area of joint space a minimum of seven points was plotted. Although an increased number of points increased the sensitivity of measurement, larger errors were introduced when 10 points were used, perhaps due to operator fatigue. The edge of the joint space for area measurement was detected by eye. Although the edge usually appears clearly on the screen, its measurement varies as a result of observer interpretation. This is more marked in *x* rays showing moderate to severe osteoarthritis, where the edge may be asymmetrical or obscured by osteophytes.

This scoring system appears to be a substantial improvement on the comparative systems. It is difficult, however, to compare it directly as it is based on a widely different scale. The comparative scoring systems have between 0 and 6 points. A one point increase is a major deterioration, and in osteoarthritis, particularly, may take many years to occur. A scoring system with a larger scale is bound to be useful, especially in early disease. Figure 5 illustrates that our measurements are applicable to a range of *x* rays with osteoarthritis, and it is not a direct comparison of the methods. Digitised processing of *x* ray analysis was pioneered by Buckland-Wright.¹³ He used microfocal radiography to improve resolution and to look at small areas in great detail. He employed a cursor to outline and measure areas of interest.⁹ Our measurement technique is similar, and the results of our calibration experiments are comparable. The resolution of our system, however, will always be limited by the quality of the *x* ray, but we can examine a larger field, which may compensate. Other advantages of our system are that it employs standard radiographs, which could originate from any radiology department. It is fast, easy to use, readily learned, and operator independent.

Digital analysis with subtraction has been applied to serial measurement of bone contour in single small joints of the hand.¹⁰ This technique is sensitive to change and has good reproducibility. So far it has only been used in the assessment of the rheumatoid hand. Although change in bone shape is important, the major abnormalities in arthritis occur within the joint and may not be appreciated if only half of the surface is measured.

Direct measurement of joint space size is advantageous in *x* ray scoring as it can be used to create a normal range, which is essential in the study of populations. This was considered by Hollingsworth *et al*, who developed a complex manual method of measuring joint space height.¹⁴ Results showed

good correlations, but Hollingsworth and colleagues did not look at the differences between or within observers. We looked at this measurement using our digitised system and found good correlations, but differences between observers were significant. Area measurement showed good correlations with no significant differences between or within observers. Absolute area measurement is a more reliable measure of joint space loss than height.

The results of the present study using this semiautomatic method suggest that the software could be enhanced to improve ease of use and accuracy of measurements. Speed of measurement would be increased by complete automation. Incorporation of edge detection would mean that the only observer intervention was placement of the *x* ray onto a grid on the screen and initiation of the program. This would achieve rapid and accurate joint space measurement at the press of a single button.

To examine the role of this scoring system several further experiments are being undertaken, including assessment of the normal joint space and comparison with that in other arthritides, and examination of joint space narrowing as a marker of early disease and its progression. Although this scoring system has been developed in osteoarthritis, there is no reason why it should not be modified to accommodate any joint and any kind of arthritis.

Image handling computers are common research tools and labour saving devices. Their application to the radiology of rheumatic diseases is important in our understanding and treatment of these conditions. We hope to adapt and use this scoring system to study further the nature and progression of arthritis.

We would like to thank the Arthritis and Rheumatism Council and Roussel Pharmaceuticals for their help. Dr Dacre is an Arthritis and Rheumatism Council research fellow.

References

- 1 Davies P T, Fowler P D, Clarke S, Fisher J, Lawton A, Shedforth M F. Rheumatoid arthritis: treatment which controls the C-reactive protein and erythrocyte sedimentation rate reduces radiological progression. *Br J Rheumatol* 1986; **25**: 44-9.
- 2 Scott D L, Grindulis K A, Struthers G R, Coulton B L, Popert A J, Bacon P A. Progression of radiological changes in rheumatoid arthritis. *Ann Rheum Dis* 1984; **43**: 8-17.
- 3 Kaji Wara T, Mitsui K, Fujii K, Murota K. The effects of non-steroidal anti-inflammatory drugs in synovial fluid on proteoglycans. *New Trends in Rheumatology* 1987; **5**: 14-23.
- 4 Arora J S. Indomethacin arthropathy of hips. *Proceedings of the Royal Society of Medicine* 1968; **61**: 669.
- 5 Newman N M, Ling R S M. Acetabular bone destruction related to non-steroidal anti-inflammatory drugs. *Lancet* 1985; **ii**: 11-14.
- 6 Kellgren J H, Lawrence J S. Radiological assessment of osteoarthritis. *Ann Rheum Dis* 1957; **16**: 494-502.

- 7 Larsen A, Dale K, Eck M. Radiographic evaluation of rheumatoid arthritis and related conditions by standard reference films. *Acta Radiol [Diagn] (Stockh)* 1977; **18**: 481-91.
- 8 Sharp J T, Lidsky M D, Collins L C, Moreland J. Methods of scoring the progression of radiological changes in RA. *Arthritis Rheum* 1971; **14**: 706-20.
- 9 Buckland-Wright J C, Carmichael I, Walker S R. Quantitative microfocal radiography accurately detects joint changes in rheumatoid arthritis. *Ann Rheum Dis* 1986; **45**: 379-83.
- 10 Gaydecki P A, Browne M, Mamtara H, Grennan D M. Measurement of radiographic changes occurring in rheumatoid arthritis by image analysis techniques. *Ann Rheum Dis* 1987; **46**: 296-301.
- 11 Greenwood A, Bryans R, Coppock J S, Dacre J E, Scott D L, Huskisson E C. Measurement of joint damage in osteoarthritis: the development of a standardised assessment technique. *Br J Rheumatol* 1987; **26** (suppl): 67.
- 12 Altman R D, Fries J F, Block D A, et al. Radiographic assessment of progression in osteoarthritis. *Arthritis Rheum* 1987; **30**: 1214-25.
- 13 Buckland-Wright C J. Microfocal radiographic examination of erosions in the wrist and hand of patients with rheumatoid arthritis. *Ann Rheum Dis* 1984; **43**: 160-71.
- 14 Hollingsworth P, Melson R D, Scott J T. Measurement of radiographic joint space in the rheumatoid knee: correlation with obesity, disease duration and other factors. *Rheumatology and Rehabilitation* 1982; **21**: 9-14.