Modified Larsen Scoring of Digitized Radiographs in Rheumatoid Arthritis

STEVEN A. YOUNG-MIN, SANJEEV SHAKHAPUR, NICKY MARSHALL, IAN GRIFFITHS, TIM CAWSTON, and ANDREW GRAINGER

ABSTRACT. ***Objective.*** To facilitate storage, retrieval, and analysis of radiographic images we assessed the validity of a film digitizer and computer based system.

Methods. A total of 101 sets of radiographs of hands and feet from patients with early rheumatoid arthritis were digitized according to standards defined by the American College of Radiology. Two blinded observers applied the modified Larsen score to the digitized images and the original radiographs.

Results. There was good concordance between the scoring of digitized images and radiographic films; intraclass correlation coefficient was 0.89. Regression analysis of a Bland-Altman plot revealed that there was no significant systematic bias. It was noted that scoring the digital images took more than twice as long as scoring the radiographic films.

Conclusion. Our results indicate that modified Larsen scoring of adequately digitized films on standard computer displays is valid and shows good agreement with conventional techniques. Scoring digitized images appears to require more time, but this may be outweighed by the considerable advantages of computer storage and retrieval of radiographic images. (J Rheumatol 2003;30:238–40)

Key Indexing Terms:
RHEUMATOID ARTHRITIS RADIOGRAPHY DIAGNOSTIC IMAGING METHODS
IMAGE PROCESSING COMPUTER ASSISTED REPRODUCIBILITY OF RESULTS

Scoring of radiographs as a measure of outcome in rheumatoid arthritis (RA) is a standard procedure. To facilitate storage, retrieval, and analysis of radiographic images we assessed the validity of a film digitizer and computer based system. The specification of the equipment used is within those currently available from carefully chosen consumer electronic products, and our results are of interest to anyone setting up an inexpensive digitizer system for radiograph scoring.

MATERIALS AND METHODS

A total of 101 sets of radiographs of hands and feet were obtained from patients with RA as defined by the American College of Rheumatology criteria[1] (disease duration < 2 yrs). Each set of radiographs was anonymized and randomized. Digitization was performed on a Vidar Systems Corporation VXR-12 film scanner at 300 dots per inch (dpi) [pixel size 85 µm, 5.8 line pairs per mm (lp/mm)], at 12-bit pixel depth (4096 levels of grey). This scanner uses a charge coupled device detector (CCD) and has an optical density range of 0 to 3.3 OD. Images were stored uncompressed in tagged image file format (TIFF) on a personal computer (IBM PC compatible with a Pentium 166 MHz processor running Windows 95). The images were displayed using Adobe Photoshop (version 4.01) on a 21 inch color display (Cornerstone color 50/115) set at 1024 × 768 pixel resolution and 32 bit color depth.

Two observers, a rheumatologist (SY) and a radiologist (SS), were trained to apply the modified Larsen scoring system[2] on a selected set of radiographic films until their intraobserver and interobserver reliability as assessed by intraclass correlation coefficients[3] were satisfactory. The observers then scored the 101 sets of films to produce a total Larsen score for each set of hands and feet. The digitized images were then scored in a similar random and blinded manner. Each observer then repeated the scoring of the radiographic films and the computer images.

RESULTS

The range of modified Larsen scores obtained was between 0 and 20, reflecting the early disease of the patients. The mean of the scores obtained by the 2 observers was compared for each technique and analyzed for agreement. The overall intraclass correlation coefficient (ICC R1) was 0.89, showing good concordance (Table 1). A Bland-Altman plot[4] was performed (Figure 1) and regression analysis revealed that there was no significant systematic bias. It was noted that in addition to the time involved in digitizing the radiographs, scoring the digital images took more than twice as long as scoring the radiographic films.

DISCUSSION

The scoring system modified by Larsen for use in longterm studies[2] was selected because it has been reported to be more sensitive to change than the original Larsen and Sharp methods in scoring radiographic progression in early RA[5]. No

From the Department of Rheumatology, University of Newcastle-upon-Tyne; Department of Radiology, Royal Victoria Infirmary, Newcastle-upon-Tyne; and Department of Musculoskeletal Radiology, Leeds Teaching Hospitals, Leeds, UK.

Supported by the Arthritis Research Campaign (UK).

S.A. Young-Min, BA Hons, MRCP, ARC Clinical Research Fellow; N. Marshall, RN, Research Associate; I. Griffiths, FRCP, Honorary Senior Lecturer; T. Cawston, PhD, Professor of Rheumatology, Department of Rheumatology, University of Newcastle-upon-Tyne; S. Shakhapur, FRCR, Specialist Registrar, Department of Rheumatology, University of Newcastle-upon-Tyne; A. Grainger, MRCP, FRCR, Consultant Radiologist, Department of Musculoskeletal Radiology, Leeds Teaching Hospitals.

Address reprint requests to Dr. S.A. Young-Min, Rheumatology, University of Newcastle-upon-Tyne, NE2 4HH, UK.
E-mail: Steven.YoungMin@ncl.ac.uk

Submitted April 23, 2002; revision accepted July 18, 2002.

Table 1. Agreement as measured by the intraclass correlation coefficient R1 (ICC R1) between observers (SY and SS) and methods applying a modified Larsen score to 101 sets of patient radiographs.

Intraobserver and intramethod agreement	
Film scores by SY	0.91
Digital image scores by SY	0.92
Film scores by SS	0.97
Digital image scores by SS	0.95
Interobserver agreement	
Mean film scores SY vs SS	0.75
Mean digital image scores SY vs SS	0.79
Intermethod agreement	
SY mean film score vs mean digital image score	0.93
SS mean film score vs mean digital image score	0.73
Overall intermethod agreement	0.89

previous studies have examined the performance of this scoring method on digitized radiographs. Two studies have, however, examined other scoring systems on digitized radiographs including the unmodified Larsen score. These studies used substantially different equipment than ours.

The first of these studies applied a modified Genant scoring method with sophisticated equipment consisting of a dedicated laser scanner, multiprocessor computer, and specialized greyscale monitor[6]. Images were scanned at 12-bit depth with pixel sizes of 50 µm. Despite concerns over the statistical methods used[7], the results showed good agreement between digital and film techniques. The intraclass correlation values obtained were similar to those observed here. The second study applied the original Larsen score to images digitized at an 8-bit depth (256 levels of grey) and 75 dpi (1.5 lp/mm)[8]. Following digitization, an unacceptably high number of the images (7%) were deemed either poor or unreadable. All the images in our study were of good quality.

The American College of Radiology have produced standards for digital image data management and these state that radiographs should be digitized to a resolution of 2.5 lp/mm or greater (> 125 dpi) and to a bit depth of 10 bits per pixel or greater (> 1024 levels of grey)[9]. These standards also specify that specialized greyscale monitors should be used to view images. The digitizer used in this study met these criteria, but the display we used was a standard 21 inch color display. Despite this, our system produces scores that agree well with traditional film based methods.

A survey of scanner manufacturers' websites revealed a number of relatively inexpensive CCD based flatbed scanners fitted with large transparency units whose specified optical performance in terms of resolution, greyscale depth, and optical density match or exceed those of the aged film digitizer used in this study and the standards set by the American College of Radiology. That our system did not require an expensive greyscale monitor adds to its affordability and potential for widespread use. Computer equipment currently installed in consulting rooms may already be suitable for reading digitized radiographs. Although we found that reading digital images required more time than conventional radiographs, this is likely to be less of a problem with modern computer equipment and is outweighed by the considerable advantages of computer storage and retrieval of radiographic images.

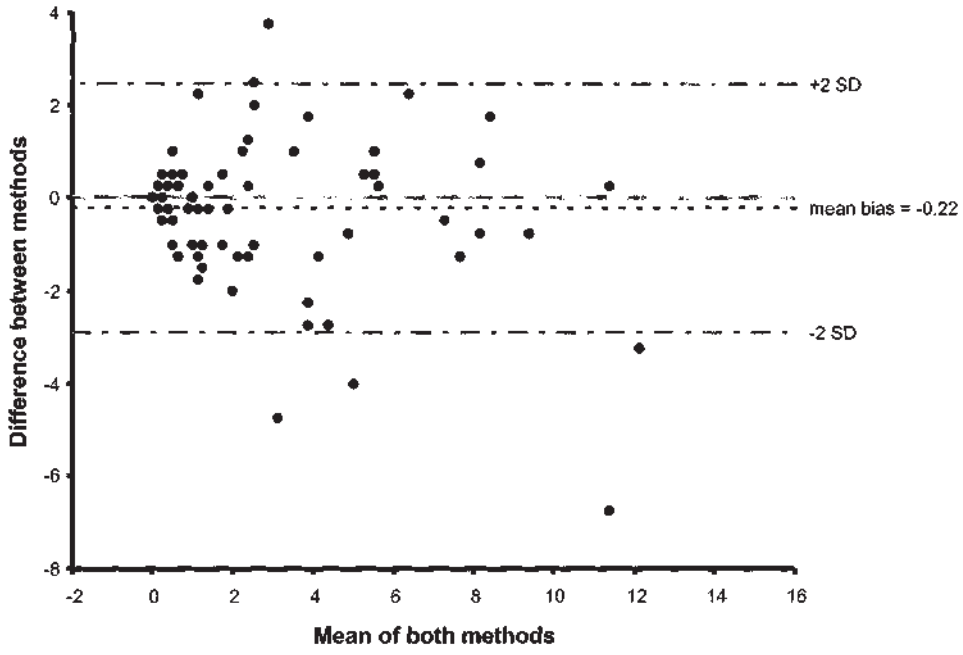

Figure 1. Bland-Altman plot of modified Larsen scores obtained by scoring radiographic films and digitized images.

REFERENCES

1. Arnett FC, Edworthy SM, Bloch DA, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315-24.
2. Larsen A. How to apply Larsen score in evaluating radiographs of rheumatoid arthritis in long-term studies. J Rheumatol 1995;22:1974-5.
3. Kramer MS, Feinstein AR. Clinical biostatistics. LIV. The biostatistics of concordance. Clin Pharmacol Ther 1981;29:111-23.
4. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1:307-10.
5. Paimela L, Laasonen L, Helve T, Leirisalo-Repo M. Comparison of the original and the modified Larsen methods and the Sharp method in scoring radiographic progression in early rheumatoid arthritis. J Rheumatol 1998;25:1063-6.
6. Genant HK, Jiang Y, Peterfy C, Lu Y, Redei J, Countryman PJ. Assessment of rheumatoid arthritis using a modified scoring method on digitized and original radiographs. Arthritis Rheum 1998;41:1583-90.
7. Giraudeau B, Ravaud P. Methodologic issues for the assessment of reproducibility: comment on the article by Genant et al [letter]. Arthritis Rheum 1999;42:1556-7.
8. Solymossy C, Dixey J, Utley M, et al. Larsen scoring of digitized X-ray images. Rheumatology (Oxford) 1999;38:1127-9.
9. Anonymous. American College of Radiology standard for digital image data management. Atlanta: American College of Radiology; 1999.