ELSEVIER

Clinical Biomechanics 14 (1999) 434–437

CLINICAL BIOMECHANICS

Digital image analysis: improving accuracy and reproducibility of radiographic measurement

Michael Bould, Stuart Barnard, Ian D. Learmonth, James L. Cunningham, John R.W. Hardy*

Department of Orthopaedics, Southmead Hospital, Bristol BS10 5NB, UK

Received 1 April 1998; accepted 8 December 1998

Abstract

Objective. To assess the accuracy and reproducibility of a digital image analyser and the human eye, in measuring radiographic dimensions.

Design. We experimentally compared radiographic measurement using either an image analyser system or the human eye with digital caliper.

Background. The assessment of total hip arthroplasty wear from radiographs relies on both the accuracy of radiographic images and the accuracy of radiographic measurement.

Methods. Radiographs were taken of a slip gauge (30 ± 0.00036 mm) and slip gauge with a femoral stem. The projected dimensions of the radiographic images were calculated by trigonometry. The radiographic dimensions were then measured by blinded observers using both techniques.

Results. For a single radiograph, the human eye was accurate to 0.26 mm and reproducible to ± 0.1 mm. In comparison the digital image analyser system was accurate to 0.01 mm with a reproducibility of ± 0.08 mm. In an arthroplasty model, where the dimensions of an object were corrected for magnification by the known dimensions of a femoral head, the human eye was accurate to 0.19 mm, whereas the image analyser system was accurate to 0.04 mm.

Conclusions. The digital image analysis system is up to 20 times more accurate than the human eye, and in an arthroplasty model the accuracy of measurement increases four-fold. We believe such image analysis may allow more accurate and reproducible measurement of wear from standard follow-up radiographs.

Keywords: Radiographic measurement; Accuracy; Reproducibility; Total hip arthroplasty; Digital image analysis

1. Introduction

The unparalleled success of Sir John Charnley's low friction arthroplasty popularised total hip replacement with the inevitable sequelae of prosthetic wear and arthroplasty failure. Studies have shown that rapid polyethylene wear allows prediction of prosthetic hip joint failure [1]. Therefore accurate and reproducible wear measurement is important to predict arthroplasty failure before gross osteolysis occurs.

Charnley himself proposed two methods of measuring wear from radiographs. The uniradiographic method subtracted the narrowest polyethylene thickness in the weight-bearing zone from the widest polyethylene thickness in the non-weight-bearing zone, and then divided by two [2]. The duoradiographic method, measured the position of the femoral ball in the cup following implantation, and then again at future follow-up [3]. The difference in these two positions represented wear when corrected for the magnification factors. Subsequent laboratory studies by Clarke et al. [4] showed these methods to be inaccurate. This was principally due to radiographs being a two-dimensional representation of a three-dimensional structure. Consequently the plane of the radiograph was not always in the plane of maximum prosthetic wear. However, Wroblewski [5] has validated Charnley's radiographic wear methods by assessing prosthetic wear in retrieval studies.

The assessment of arthroplasty wear from radiographs relies on the accuracy of the radiographic images and the method of radiographic measurement. We tested the hypothesis that measurement of radiographic

* Corresponding author. E-mail: 106347, 1243@compuserve.com

0268-0033/99/$ – see front matter
PII: S0268-0033(98)00113-2

dimensions is more accurate and reproducible using digital image analysis than manual techniques.

Assessment of accuracy relied on comparing the measurement data with a known dimension. The reproducibility was assessed by the distribution of data around the mean.

2. Methods

To accurately measure radiographic dimensions elimination of errors is required. Therefore the object imaged by a radiograph must have a dimension known to a high degree of accuracy. Only if the object directly overlaid the film and the X-ray source was at infinity, would magnification be eliminated. Therefore magnification must be minimised and its effect on image size calculated. The reference dimension used in this study was a commercially available slip-gauge (ECG. Gauging, Gloucester, UK). The slip-gauge used was a metallic block with a width of 30 ± 0.00036 mm at a temperature of 20°C.

Radiographic images of the slip gauge were taken at 20°C using a single X-ray machine. The film focal distance was set at 1000 mm using the calibration inherent in the X-ray machine. The position of the X-ray source, the tube voltage and the exposure time were then left unaltered for all the subsequent radiographs. The same radiograph cartridge, with intensifying screens, was used for all exposures. The distance between the exposed surface of the cartridge and the film within was measured at the end of the experiment with a digital caliper with a resolution of 0.01 mm (Mitutoyo, absolute digimatic caliper; RS Components, Bristol, UK).

Four radiographs of the slip gauge were taken with the source centred over the slip gauge to model statistical fluctuation. A further set of four radiographs were taken of the slip gauge alongside a femoral prosthesis (Charnley: DePuy International, Leeds, UK), with the X-ray beam centred over the femoral head (Fig. 1). In each of these radiographs the slip gauge was placed adjacent to the femoral prosthesis at its circumference with its 30 mm dimension lying adjacent to the spherical head. This latter model was designed to mimic the relationship of a femoral cup and femoral head in a radiograph of a hip arthroplasty (arthroplasty model). There was no attempt to model the soft tissue that surrounds a real hip arthroplasty, and as a consequence the object was closer to the film than is possible in vivo.

The identities of these radiographs were blinded to the observers and they were further disguised amongst eight similar radiographs taken at different film focal distances. The radiographs were digitised using the VXR12 digitizer (Vidar Systems, Herndon, VA, USA) with 4096 levels of grey resolution. The images were subsequently analysed using the PC Radscan software

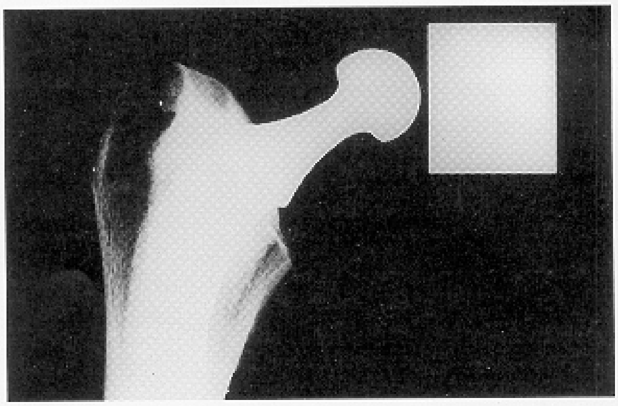

Fig. 1. Radiograph of arthroplasty model; horizontal width of slip gauge 30 mm.

(Howmedica, Rutherford, NJ, USA). Measurement of dimensions required operator pixel selection, and the centre of a circle was found by use of computer-generated concentric circles. This process was enhanced by magnification of the image and colour-density-aided edge detection. Other more complex image analyser functions were not utilised.

From each of these radiographs the width of the slip gauge was measured five times separately with both the digital image analysis system and the human eye using the digital caliper which was certified accurate to ±0.02 mm. All these measurements occurred at a constant temperature of 20°C. For the arthroplasty model radiographs, the diameter of the femoral head was also measured using an overlay of concentric circles with a resolution of 1 mm [6]. The value of each measurement was not seen by the measurer but recorded independently to prevent observer bias. The process was then repeated with a second observer to assess to the effect of inter-observer error. From these measurements, the means and standard deviations were calculated.

To calculate the accuracy of both techniques of measurement, the expected radiographic width of the upper surface of the slip gauge was calculated using trigonometry to correct for magnification. As the dimensions in the trigonometric equation were measured with a finite error, three calculations were performed as below:

1. Assuming no errors of measurement.
2. Assuming all errors set to give maximum magnification.
3. Assuming all errors set to give minimum magnification.

The distances and errors involved in the trigonometric calculation are set out in Fig. 2. From these calculations the expected width of the upper surface of the slip gauge was found to be 30.367 ± 0.006 mm. Since the slip gauge had a depth of 9.08 ± 0.02 mm, the lower surface of the slip gauge appeared as smaller

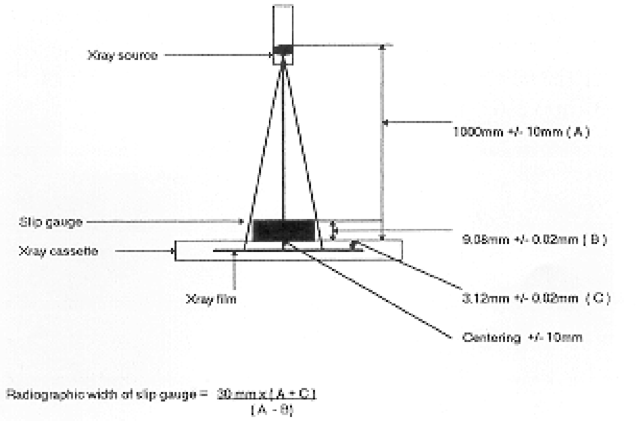

Fig. 2. Dimensions and errors for radiographic set-up.

radiographic projection. Thus a penumbra in-between both radiographic projections was cast, which might have potentially affected the precision of measurement. Therefore every effort was made to measure the outer edge, which corresponded with the upper surface of the slip gauge, with both measurement techniques.

In clinical practice the thickness of the femoral cup is corrected for magnification using the known diameter of the metal femoral head. This allows sequential follow-up radiographs to be compared, with correction for a variable magnification factor. For the arthroplasty model, the measured diameter of the femoral head was used to correct the magnification of the slip gauge using the following equation:

$$\text{slip gauge width} = \frac{\text{radiographic slip gauge width} \times \text{femoral head diameter (22.25 mm)}}{\text{radiographic femoral head diameter}}.$$

Since the equator of the femoral head was 11.125 mm from the top of the X-ray cassette and the slip gauge was only 9.08 mm, a further small magnification correction had to be made to allow for the greater relative magnification of the femoral head.

The accuracy for each of the two techniques, was assessed by subtracting the observed width of the slip gauge from the expected width. The reproducibility of each of the two techniques was assessed as plus or minus 1.96 times the standard deviation (i.e. 95% confidence interval).

Data analysis was performed on a personal computer using Student *t*-test. The data were analysed in the following subgroups for both the image analyser and the human eye:

1. One observer, one slip gauge radiograph.
2. One observer, four identical radiographs of the slip gauge.
3. Two observers, four identical radiographs of the slip gauge.
4. Two observers, four identical radiographs of the arthroplasty model.
5. Therefore for each consecutive subgroup, a further source of error is added.

To assess the validity of the digital image analyser technique against the pre-existing human eye technique, the data from only one observer was used. Five measurements on both the four slip gauge radiographs together with five measurements on the eight radiographs at various film focal distances were compared for the two measurement techniques. The method for assessing agreement between two methods of clinical measurement as proposed by Bland and Altman [7] was used.

3. Results

The accuracy and reproducibility for both techniques in each subgroup of radiographs is shown in Table 1. A Student *t*-test of the values for accuracy obtained for each of the two techniques showed that there were differences at the $P < 0.005$ significance level, for all of the data subgroups.

For the single observer, single radiograph subgroup the only sources of error were an error of measurement and intra-observer error. The image analyser allowed accuracy to 0.01 mm. However the error of measurement with the human eye was 0.26 mm. When intra-radiograph and intra-observer errors were added the accuracy remained unchanged. However the measurements became less reproducible. The error of correction for magnification decreased the accuracy of the image

Table 1
Accuracy and reproducibility for measurement using the human eye and digital image analyser

Data subgroup	Human eye		Image analyser	
	Accuracy (mm)	Reproducibility (mm)	Accuracy (mm)	Reproducibility (mm)
1 Observer/1 radiograph	0.26*	±0.10	0.01	±0.08
1 Observer/4 radiographs	0.23*	±0.25	0.03	±0.13
2 Observer/4 radiographs	0.26*	±0.25	0.01	±0.19
2 Observer/4 radiographs (Arthroplasty model)	0.19*	±0.43	0.04	±0.18

* Denotes significant difference compared with accuracy for digital image analyser in *t*-test at $P < 0.005$ level.

analyser measurement method. Nevertheless, the digital image analyser technique remained significantly more accurate than the human eye ($P < 0.005$).

To assess the validity, the agreement between the two measurement techniques was calculated for one observer measuring twelve radiographs. In this way the digital image analyser is compared with the best previous technique. The mean difference between the two techniques was 0.248 mm, with the limits of agreement being −0.004 mm below and +0.500 mm above.

4. Discussion

The clinical problem of assessing wear from radiographs requires reproducible radiographs and an accurate and reproducible method of measuring dimensions from these radiographs. Simple AP radiographs will only image wear accurately if the plane of the radiograph is in the same plane as the maximum wear. Therefore a three dimensional imaging technique is required. CT and MRI are not suitable because of surface scatter from metal implants [8]. Recently three-dimensional reconstruction from AP and lateral radiographs have become the gold standard for accurate radiographic imaging for wear measurements [9]. Such techniques are not yet widely available, and most follow-up is performed using the Duoradiogaphic technique as described by Charnley in 1973. The stated accuracy of this technique is ±0.5 mm [10].

Part of the inaccuracy of the duoradiographic technique is the inaccuracy of human eye measurement. In this simple experiment we have shown that this inaccuracy is of the order of 0.26 mm. This inaccuracy can be reduced by using a digital image analyser, which in our experiment has an accuracy of 0.01 mm. The superior accuracy and reproducibility of the digital image analyser can be attributed to its greater magnification, contrast-aided edge detection and a resolution of 0.085 mm per pixel. In addition the digital image analyser utilises information from many image pixels, whereas the human eye method utilises information from a smaller image.

These accuracy results are from a purely experimental model, and no effort has been made to reproduce the soft tissue around a hip arthroplasty and the effect this may have on the radiographic imaging.

5. Conclusions

We have shown that a digital image analysis system is up to twenty times more accurate than the human eye when measuring simple radiographic dimensions. In an arthroplasty model the Image Analyser increases the accuracy of measurement over four-fold and doubles its reproducibility. We believe such image analysis may allow more accurate and reproducible measurement of wear from standard post-operative radiographs.

References

[1] Kobayashi S, Takaoka K, Saito N, Hisa K. Factors affecting aseptic failure of fixation after primary Charnley total hip arthroplasty. JBJS 1997; 79-A:1618–27.
[2] Charnley J, Cupic Z. The nine and ten year results of the low friction arthroplasty of the hip. Clin Orthop 1973;95:9–25.
[3] Charnley J,Halley DK. Rate of wear in total hip replacement. Clin Orthop 1973;112:170–9.
[4] Clarke IC, Black K, Rennie C, Amstutz HC. Can wear in total hip arthroplastics be assessed from radiographs? Clin Orthop 1973;112:170–9.
[5] Wroblewski, BM. Direction and rate of socket wear in Charnley low-friction arthroplasty. JBJS 1985 67-B:757–61.
[6] Livermore J, Ilstrup D,Morrey B. Effect of femoral head size on wear of the polyethylene acetabular component. JBJS 1990;72-A:518–28.
[7] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;8:307–10.
[8] Mankovich NJ, Robertson DR, Cheeseman AM. Three-dimensional image display in medicine. J Digit Imaging 1990;3:200–203.
[9] Devane PA, Bourne RB, Rorabeck CH, Hardie RM, Horne KG. Measurement of Polyethylene wear in Metal-backed acetabular cups. CORR 319:303–316.
[10] Griffith MJ, Seidenstein Mk, Williams D, Charnley J. Socket wear in Charnley low friction arthroplasty of the hip. Clin Orthop 1978;137:37–47.