EBRA improves the accuracy of radiographic analysis of acetabular cup migration

Thomas Ilchmann, Uldis Kesteris and Hans Wingstrand

EBRA (Ein Bild Röntgen Analyse) is a new computerized method measuring migration and wear of the acetabular cup, suggested to improve measurement accuracy. We evaluated possible errors of measurement and compared EBRA with standard methods.

1. We did repeated measurements on a single radiograph using the same reference lines. The reliability of the input procedure with standard measurements was significantly better than repeated digitization with EBRA.

2. In a more clinical test, a group of 10 patients was studied. 5 radiographs were taken of the same patient on the same day. EBRA improved the reliability of repeated radiographic examination significantly for migration measurements in the vertical direction.

3. To assess the inter- and intraobserver variations, repeated measurements were performed on the clinical series of pelvic radiographs of 10 patients. EBRA was significantly better than standard methods.

With EBRA, errors of wear and migration measurements could be reduced, as compared to standard methods. The major improvement with EBRA was found for migration measurements in the vertical direction.

Department of Orthopedics, Lund University Hospital, SE-221 85 Lund, Sweden. Tel +46 46-171510. Fax -130732.
Correspondence: Dr. Thomas Ilchmann, BG Unfallklinik, Schnarrenbergstr. 95, DE-72076 Tübingen, Germany
Tel +49 7071-6062308. Fax -6062233
Submitted 96-10-27. Accepted 97-10-27

Radiographic assessment of prosthetic migration and wear is well established in follow-up studies of results after total hip arthroplasty. When the assessment is accurate, it is of prognostic value (Ryd et al. 1993, Stocks et al. 1995, Krismer et al. 1996). Computerized image analysis may become important because of the increasing use of digitized roentgen techniques. The reproducibility of such measurements is high (Hardinge et al. 1991), but the accuracy is not proven. Roentgen stereophotogrammetry (RSA) (Selvik 1989) is a very accurate method for 3-dimensional analysis. It requires special roentgen-equipment, implantation of markers, a special digitizer and it is time-consuming. Therefore only smaller prospective studies are possible.

An objective and accurate method is needed for comparing results. A retrospective method would permit studies of the huge amount of clinical data currently available. This might lead to better understanding of any shortcomings in present designs and routines.

The Department of Orthopedics and the Institute of Mathematics and Geometry at the University of Innsbruck, Austria, have developed the EBRA (Ein Bild Roentgen Analyse) method for measuring migration of acetabular cups by means of a computer (Russe 1988, Krismer et al. 1995). As suggested by one of the authors (TI), an analysis of wear was added to the program. In previous studies, EBRA was found to differ less from corresponding RSA values, thus being more accurate than standard non-computerized methods for measuring migration and wear (Ilchmann et al. 1992, 1995). In this study, we have evaluated possible errors in measurements using EBRA and compared it with standard methods for measuring migration and wear.

Material and methods

EBRA

A pencil, a ruler, a light board, technical drawing equipment, an ordinary digitizing table and a personal computer with Windows® are needed. For each patient at least 4 anteroposterior radiographs of the pelvis have to be collected, best evenly distributed during the period of follow-up (Russe 1988). Retrospective studies are possible if the patient has had regular radiographic examinations. In later follow-ups, further radiographs can be added to the same analysis using the input and data of the initial study. Thus a database of EBRA measurements is created. All cups

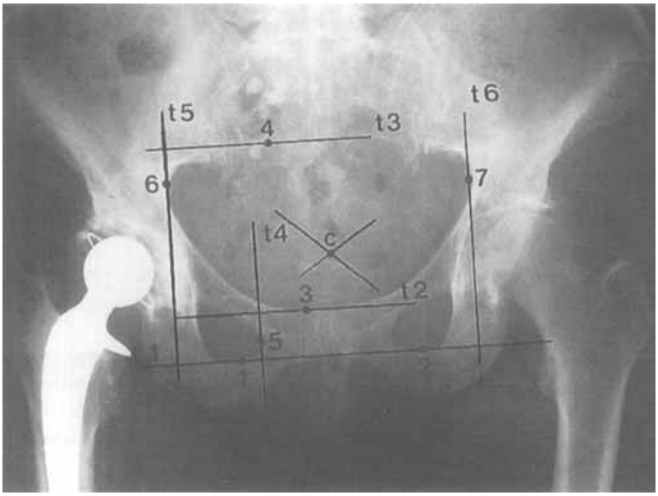

Figure 1. The tangents t1–t6 are defined and digitized by 7 points 1–7. The point c marks the approximate center of the radiograph.

with a perfectly round contrast wire, equatorially or coronally placed, can be measured. Metal cups must have a spherical surface. The size of the femoral head or the acetabular cup has to be known.

The center of the radiograph (crossing-point of the diagonals, point c, Figure 1) is marked with a ruler. Six tangents (t1–t6, Figure 1) must be drawn accurately on prominent pelvic landmarks. Their positions were selected because they were easy to detect and little affected by any pelvic tilt. T1 is placed on the distal border of the projection of the foramina obturatoria or else on the distal border of the tuber ischiadicae. T2 and t3 are parallel to t1. They are placed on the cranial border of the symphysis and on the most clearly visible foramen in the sacrum. T4–t6 are perpendicular to t1–t3. T4 is located on the medial border of the foramen obturatum on the side of the replacement, while t5 and t6 are located on both insides of the large pelvic aperture. It is important that identical pelvic structures are marked on all radiographs of the same patient.

Patient data, type of implant, head or cup size, the date of the radiograph and, if known, the focus-film distance are fed into the computer. The center of the radiograph (c) and the seven points (1–7) defining the position of the 6 tangents (t1–t6) must be digitized (Figure 1).

The software provides a menu for selecting the type of implant. A drawing of the corresponding prosthesis appears on the screen, showing the reference points to be digitized. The femoral head and cup are digitized with 3–7 points, distributed about equally on the contours. From the digitized points, the best-fitting circle for the head, the best-fitting ellipse for the contrast wire of the cup and the circle for the surface of the metal-backed cups are calculated by the software. A circle is defined by 3, an ellipse by 5 points. The remaining digitized points are used to check the input. The maximal differences between the digitized points and the calculated circle or ellipse are calculated. If a point differs by more than 0.2 mm from the calculated values, a warning sign is shown and a new input proposed (radial error).

The horizontal position of the implant is calculated as the distance to a plane defined by an approximately calculated focus point and tangent t2, the vertical position to a plane defined by the focus point and tangent t4. The background mathematics are complicated, but Russe (1988) and Krismer et al. (1995) state that, for geometrical reasons, such measurements are less affected by possible pelvic tilt than are measurements taken in the film plane.

For each radiograph, the radiographic magnification is calculated from the digitized femoral head (f = calculated size/known size). All digitized coordinates are corrected for magnification. The distances between the tangents t1–t6 are also calculated for each radiograph. Projections are defined as comparable if the distances between the tangents are similar. The maximal accepted difference ("limit of comparability, L") between the distances should not exceed 3 mm. Radiographs with tangent distances differing by more than "L" mm among the repeated examinations are excluded from further analysis. Subgroups with similar distances between the reference lines are identified. For a single patient several subgroups of comparable projections may be found. The position of the head and cup in relation to the reference plane (migration) and the relation of the head to the cup center (wear) are calculated for each subgroup. In the case of several subgroups, the results are averaged for each interval. The migration and wear curve over the total

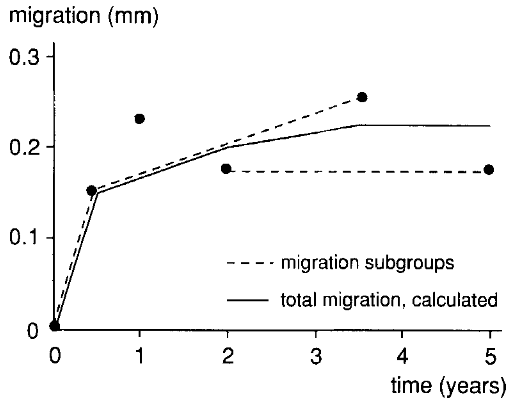

Figure 2. Imaginary construction of a mean migration curve based on 6 radiographs. 2 subgroups with similar projections, consisting of 2 and 3 radiographs, were found by the software. For the radiograph at 1 year, no radiograph with a corresponding projection could be found, and therefore it was excluded from analysis. In the interval from 2–3.5 years, there was an overlap of 2 subgroups. The mean trend of both curves was calculated, representing the gradient in the total migration curve for that period.

period of analysis are created by adding the partial trends (Figure 2). In case of only one subgroup of comparable radiographs, this subgroup represents the pattern of migration (wear) in the period, covered by the subgroup. The remaining radiographs are excluded as not comparable, thus avoiding errors of measurement. Additional radiographs in later follow-ups may be comparable to the remaining radiographs and these may be included in later analyses also.

The results are presented in migration-time and wear-time diagrams and in tables that can be used with other software. The migration of the femoral head (head x), the acetabular cup (cup x) and the wear (wear x) in the horizontal and vertical directions (head y, cup y, wear y) can be studied.

The total wear (EBRA wear) was calculated from the EBRA wear results in the horizontal (wear x) and vertical (wear y) directions by vectorial addition, to make the results comparable to the other methods. Furthermore, the total wear was calculated as the differences between migration of the head and cup in the horizontal and vertical directions (EBRA h-c).

Non-computerized methods

Standard measurements were made on all radiographs, disregarding possible exclusion by the EBRA method. Templates were used to define the center of the femoral head, distances were drawn and marked with a well-sharpened pencil. Measurements were made with a slide caliper having a 0.1 mm scale. All results were corrected for radiographic magnification (f = measured head size/true head size).

For migration measurements, the teardrop line and a perpendicular line through the center of the teardrop were used as the reference lines (Nunn et al. 1989), the cup was defined by the center of the contrast wire. Migration was measured in the horizontal (Nunn x) and vertical (Nunn y) directions. The wear, measured with the Livermore method, was the narrowest distance between the border of the femoral head and the prostheses/cement interface (Livermore et al. 1990).

All measurements were made by one of the authors (TI), who was very familiar with the methods used. The measurements of the interindividual difference were made by 2 of the authors (UK, TI). In all EBRA measurements, the limit of comparability was L = 3 mm. If the input of the prosthesis proved to be inaccurate regarding the radial error calculated by the software, it was repeated up to 5 times. A maximal radial error of 0.25 mm was accepted.

Methods of evaluation

Reliability of the input procedure

A given clinical radiograph was digitized 5 times with EBRA, using the same reference lines. The radiograph was not moved on the digitizer between the examinations. Nunn and Livermore measurements were made 5 times on the same radiographs, using the same reference lines. With EBRA, the first radiograph is set as zero automatically. Thus, in a similar way, the first measurement with the standard methods was taken at zero, the differences from the first radiograph being considered as errors of input. Five radiographs from 5 different patients were analyzed. We calculated the maximal error, the measurement error (standard deviation) and the 95% confidence interval of the measurement error in all corresponding measurements. Statistical significance corresponds to no overlap of the confidence intervals of the measurement error.

Reliability of radiographic re-examination

This study was approved by the local Swedish committee on ethics. 5 radiographs of the pelvis were taken of a given patient on the same day. The 10 patients included were mobilized with full weight bearing after THA had been performed. They were requested to stand up and walk between each radiographic examination. No special care was taken to position the patients, no changes were made in the settings of the radiographic equipment between the examinations. No migration or wear should have occurred between the repeated examinations.

The implant position was measured with the EBRA and the Nunn methods. The thickness of the plastic

Table 1. Repeated input procedures of the same radiograph. Migration measurements in the horizontal (x) and vertical (y) directions with the EBRA (EBRA head, cup) and Nunn methods. Wear measurements with the EBRA (EBRA wear) and Livermore methods. Differences of head and cup migration with EBRA (EBRA h-c). No radiograph was excluded by the EBRA software. Absence of overlap of the 95% confidence intervals of the measurement error indicates a significant difference between the methods used for comparison

Method	n	Maximal error (mm)	Measurement error (mm) [a]	
EBRA head x	25	0.2	0.13	0.10–0.18
EBRA cup x	25	0.5	0.25	0.20–0.35
Nunn x	25	0.2	0.11	0.09–0.15
EBRA head y	25	0.4	0.19	0.15–0.26
EBRA cup y	25	0.9	0.31	0.24–0.43
Nunn y	25	0.3	0.11	0.09–0.15
EBRA wear	25	0.4	0.19	0.15–0.26
EBRA h-c	25	0.8	0.23	0.18–0.32
Livermore	25	0.1	0.06	0.05–0.08

[a] Error and 95% confidence interval of error

was measured with the EBRA and the Livermore methods. The differences, as compared to the first radiographic examination, were considered as errors of repeated examination and measurement. We calculated the maximal error, the measurement error (standard deviation) and the 95% confidence interval of the measurement error for all corresponding measurements.

Intraobserver variation

10 patients were operated on with THA (Scanhip®, MitAB, Sjöbo, Sweden), published in a previous study (Mjöberg 1986). A series of 5 (4–8) radiographs of the pelvis was taken at regular intervals over a period of 56 (48–68) months, no special care being taken to position the patients. With EBRA, all 60 radiographs were analyzed, the early postoperative radiographs were considered to have been detected by the program in case of possible tilt. With the standard methods, the early postoperative radiographs were excluded to avoid errors of measurement, 48 radiographs were left. All measurements were made twice by the same person, using all methods in question. The measurements were made on different occasions, all lines being thoroughly erased between the examinations. We calculated the maximal error, the measurement error (standard deviation) and the 95% confidence interval of the measurement error in all corresponding pairs of measurements.

Interobserver variation

The same radiographs from the study of the intraobserver variation were re-analyzed with all methods by another of the authors, after erasing all lines thoroughly. The results were compared with the corresponding values from the first measurement of the intraobserver difference. We calculated the maximal error, the measurement error (standard deviation) and the 95% confidence interval of the measurement error in all corresponding pairs of measurements.

Results

The analysis of a single radiograph took about 10–15 minutes, using the EBRA method. The contour of the pubic symphysis for the construction of tangent t2 was sometimes difficult to discern. In some patients, it was difficult to identify the teardrop line in the Nunn method and the prosthesis/cement interface in the Livermore method.

Reliability of the input procedure

The measurements could be made on all radiographs and none of the repeated inputs was excluded by the EBRA software, because of differences in projection. The input was significantly more precise for the head than for the cup. The EBRA wear measurements seemed to be more precise than the EBRA h-c measurements (Table 1).

Nunn and Livermore measurements were significantly better than EBRA measurements. Direct measurements of well-defined structures or lines on single radiographs, using the slide caliper, were more precise than repeated digitization, using EBRA (Table 1).

Reliability of radiographic reexamination

The software of the EBRA method excluded some radiographs for measurements in the vertical direction because of an abnormal projection or inaccurate input of the implant (number of measurements, Table 2). With the standard methods, measurements could be made on all radiographs. The EBRA measurements of the head were significantly better than those of the cup (Table 2).

With the Nunn method, measurements in the horizontal direction were significantly better than in the vertical direction (Table 2). Migration measurements with EBRA were significantly better than with Nunn, especially in the vertical direction. There was no significant difference between wear measurements with EBRA and Livermore. EBRA wear measurements seemed better than EBRA h-c measurements (Table 2).

Intraobserver variation

The EBRA software excluded some radiographs for measurements in the horizontal and several in the ver-

Table 2. Repeated radiographic examinations on a given day. Some radiographs were excluded by the EBRA software (as in Table 1)

Method	n	Maximal error (mm)	Measurement error (mm) [a]	
EBRA head x	50	0.6	0.25	0.21–0.31
EBRA cup x	50	0.8	0.35	0.29–0.44
Nunn x	50	1.1	0.35	0.29–0.44
EBRA head y	46	0.6	0.20	0.17–0.25
EBRA cup y	47	1.3	0.38	0.32–0.48
Nunn y	50	1.8	0.75	0.63–0.93
EBRA wear	46	0.7	0.17	0.14–0.21
EBRA h-c	46	1.4	0.24	0.20–0.30
Livermore	50	0.4	0.13	0.11–0.16

[a] Error and 95% confidence interval of error

Table 3. Assessment of the intraobserver variation. Several radiographs were excluded by the EBRA software, especially for analysis in the vertical direction and for wear assessment (as in Table 1)

Method	n	Maximal error (mm)	Measurement error (mm) [a]	
EBRA head x	44	2.4	0.55	0.45–0.70
EBRA cup x	39	2.0	0.51	0.42–0.66
Nunn x	48	2.0	0.82	0.68–1.03
EBRA head y	20	1.3	0.64	0.49–0.93
EBRA cup y	18	1.7	0.57	0.43–0.85
Nunn y	48	3.5	1.02	0.85–1.28
EBRA wear	16	0.4	0.15	0.11–0.23
EBRA h-c	16	0.5	0.25	0.18–0.39
Livermore	48	0.9	0.29	0.24–0.36

[a] Error and 95% confidence interval of error

Table 4. Assessment of the interobserver variation. Several radiographs were excluded by the EBRA software (as in Table 1)

Method	n	Maximal error (mm)	Measurement error (mm) [a]	
EBRA head x	50	2.2	0.70	0.58–0.87
EBRA cup x	41	2.5	0.79	0.65–1.01
Nunn x	48	4.5	0.91	0.76–1.14
EBRA head y	23	1.3	0.43	0.33–0.61
EBRA cup y	20	2.6	0.75	0.57–1.10
Nunn y	48	3.5	1.05	0.87–1.32
EBRA wear	17	0.4	0.18	0.13–0.27
EBRA h-c	17	0.9	0.39	0.29–0.59
Livermore	48	0.9	0.40	0.33–0.50

[a] Error and 95% confidence interval of error

tical directions because of an abnormal projection or inaccurate input of the implant (number of measurements, Table 3). All radiographs were measured with the standard methods. With EBRA (EBRA cup), the intraobserver variation was significantly lower for migration measurements in the horizontal and vertical directions than with the Nunn method. Wear measurements were significantly better with EBRA than with the Livermore method. EBRA wear measurements seemed to be better than EBRA h-c measurements (Table 3).

Interobserver variation

The results were similar to the intraobserver variation. For migration measurements in the medial direction, we found no significant difference between EBRA and Nunn. Migration measurements (EBRA head) in the vertical direction and EBRA wear measurements were significantly better than the corresponding standard measurements. EBRA wear measurements were significantly better than EBRA h-c measurements (Table 4).

Discussion

We evaluated the EBRA method to obtain more information about possible advantages of the use of a computerized method for assessing migration and wear, as compared to standard non-computerized methods.

The use of a digitizer and PC did not improve the precision of the data input. The data input was time-consuming and a cause of error. This might be remedied by using automatic image analysis, a very reliable method (Hardinge et al. 1991). On the other hand, if the distance between two well-defined points or lines is carefully measured with a slide caliper, the value is also very precise.

When assessing the reliability of the radiographic examination, errors in input of data and in projection differences may have occurred. Measurements of implant position (and calculation of possible migration in time series) with EBRA were more accurate than those with standard methods, probably because of the tilt analysis. The main improvement was in the assessment of position in the vertical direction. Measurements of wear were not improved by using the EBRA method.

The reproducibility of the measurements was tested by assessing the intra- and interobserver errors. EBRA was better than the standard measurements, especially for measurements of migration in the vertical direction. Furthermore, the concordance was better for measurements of wear.

We found a maximal error of 2.5 mm in migration and 0.4 mm in wear measurements with EBRA. The maximal error of the standard methods was 4.5 mm for migration and 0.9 mm for wear. Krismer et al. (1995) found a lower interobserver error. The radiographs in their study were probably more standardized. Single measurements should be interpreted with caution before clinical decisions are made. More reliable methods, such as RSA, should be used to evaluate the individual patient.

We found a discrepancy between wear measurements using the software (EBRA wear) and calculation of wear as the difference between head and cup migration (EBRA h-c). This may be due to a difference between the routines in the software for assessing migration and wear. The EBRA measurements of wear are very reliable (Ilchmann et al. 1995). Therefore the measurements of migration of the head or cup may have some systematic error. It could also be caused by differences between drawings of the reference lines, which were sometimes difficult to make.

The goals of analyzing tilt and excluding unusual projections seem to have been achieved (Ilchmann et al., unpublished data). A new version of the software allows marking of digitized radiographs directly on the screen with automatic input of the prosthetic component data. The analysis then takes about five minutes per radiograph and errors in data input are avoided. Additional radiographs during follow-up, highly standardized radiographic techniques and the implantation of metal markers in the pelvic bone may further improve the accuracy, if the method is used in prospective studies (Krismer et al. 1995, Malchau et al. 1995). More studies are needed concerning errors in measurements of implant designs, using different input routines.

The main improvement is made in migration analysis in the vertical direction. This is the main direction of possible cup migration. Cup migration has predictive value and can be detected earlier with more accurate methods (Mjöberg 1986, Önsten et al. 1993, Stocks et al. 1995).

The accuracy of RSA as the "golden standard" cannot be attained by the EBRA method. RSA should be used for first clinical trials of new implants and techniques in pilot studies of a small number of patients. As it is technically demanding and time-consuming, it might be more appropriate in larger research centers. However, EBRA is more reliable than standard methods for radiographic studies in a larger number of patients.

Acknowledgments

This study was supported by grants from the Swedish Medical Research Council (MFR) project 09509, King Gustaf V 80th Birthday Fund, Medical Faculty of the University Lund, Lunds sjukvårdsdistrikt, Stiftelsen för bistånd åt vanföra i Skåne, and Science and Technology Procordia AB.

References

Hardinge K, Porter M L, Jones P R, Hukins D W L, Taylor C J. Measurement of hip prostheses using image analysis. The MAXIMA hip technique. J Bone Joint Surg (Br) 1991; 73: 724-8.

Ilchmann T, Franzén H, Mjöberg B, Wingstrand H. Measurement accuracy in acetabular cup migration. A comparison of four radiologic methods versus roentgen stereophotogrammetry. J Arthroplasty 1992; 7 (2): 121-7.

Ilchmann T, Mjöberg B, Wingstrand H. Measurement accuracy in acetabular cup wear. Three retrospective methods compared with roentgen stereophotogrammetry. J Arthroplasty 1995; 10 (5): 636-42.

Ilchmann T, Kesteris U, Wingstrand H. The effect of pelvic tilt on radiographic measurements after total hip arthroplasty. An experimental analysis of the computerized EBRA method. Hip International. In press.

Krismer M, Bauer R, Tschupik J, Mayrhofer P. EBRA: a method to measure migration of acetabular components. J Biomechanics 1995; 28: 1225-36.

Krismer M, Stöckl B, Fischer M, Bauer R, Mayrhofer P, Ogon M. Early migration predicts late aseptic failure of hip sockets. J Bone Joint Surg (Br) 1996; 78: 422-6.

Livermore J, Ilstrup D, Morrey B. Effect of femoral head size on wear of the polyethylene acetabular component. J Bone Joint Surg (Am) 1990; 72: 518-28.

Malchau H, Kärrholm J, Wang X Y, Herberts P. Accuracy of migration analysis in hip arthroplasty. Digitized and conventional radiography, compared to radiostereometry in 51 patients. Acta Orthop Scand 1995; 66 (5): 418-24.

Mjöberg B. Loosening of the cemented hip prosthesis: the effect of heat injury. Acta Orthop Scand (Suppl 221) 1986; 57: 1-40.

Nunn D, Freeman M A R, Hill P F, Evans S J W. The measurement of migration of the acetabular component of hip prostheses. J Bone Joint Surg (Br) 1989; 71: 629-31.

Önsten I, Bengnér U, Besjakov J. Socket migration after Charnley arthroplasty in rheumatoid arthritis and osteoarthritis: a roentgen stereophotogrammetric study. J Bone Joint Surg (Br) 1993; 75: 677-80.

Russe W. Röntgenphotogrammetrie der künstlicen Hüftgelenkspfanne. Aktuel Probl Chir Orthop 32. Huber, Berne 1988.

Ryd L, Albrektsson B E J, Carlsson I, Toksvig-Larsen S, Herberts P, Lindstrand A. On the clinical significance of micromotion of joint implants. 39th Annual Meeting, Orthopedic Research Society, San Francisco 1993.

Selvik G. Roentgen stereophotogrammetry. A method for the study of the kinematics of the skeletal system. Acta Orthop Scand (Suppl 232) 1989; 60: 1-51.

Stocks G W, Freeman M A R, Evans S J W. Acetabular cup migration. Prediction of aseptic loosening. J Bone Joint Surg (Br) 1995; 77: 853-61.