Acta Orthop Scand 1995; 66 (5): 418–424

Accuracy of migration analysis in hip arthroplasty

Digitized and conventional radiography, compared to radiostereometry in 51 patients

Henrik Malchau[1], Johan Kärrholm[1], Yu Xing Wang[2] and Peter Herberts[1]

We assessed the accuracy of migration measurements on conventional and digitized radiographs of total hip arthroplasties by comparing the results with radiostereometry (RSA). 4 stem and 3 acetabular designs were studied. 2 of the stem designs and 1 cup design were uncemented. 180 manual and 202 digitized measurements were done on 120 conventional radiographic examinations.

The readings on digitized radiographs did not differ from the manual measurements on the same radiographs. A comparison of the measurements from plain radiographs and with RSA of the femoral stems revealed an accuracy (absolute mean + 2 SD) of between 3.9 and 12.3 mm, depending on the choice of landmarks. The greatest accuracy was obtained by using tantalum markers inserted into the greater or the lesser trochanter and the shoulder of the stem. The most medial point of the lesser trochanter proved to be the best bony landmark. Measurements, including both the center of the femoral head and the greater trochanter, were associated with poor accuracy. The accuracy as regards horizontal cup migration varied from 4.4 to 6.5 mm and the accuracy as regards vertical migration varied between 4.4 and 6.3 mm.

The intraobserver error (2 SD) ranged from 1.6 mm to 5.6 mm, The corresponding figures for the interobserver error were 2.6 mm and 6.6 mm, respectively. One of the cemented cup designs was associated with inferior accuracy. Stem design did not affect the accuracy.

Departments of [1]Orthopedics and [2]Diagnostic Radiology, Sahlgrenska University Hospital, S-413 45 Gothenburg, Sweden
Tel +46 31-602076. Fax -825599
Submitted 95-02-15. Accepted 95-05-20

We evaluated the accuracy of migration measurements on conventional radiographs of cemented and cementless total hip arthroplasty, using digitized and conventional manual measurements. These measurements were compared to those obtained by radiostereometry (RSA).

Patients and methods

2 types of uncemented (Rippen®, Valdemar Link and Ti-fit®, Smith and Nephew Richards) and 2 cemented stems (Lubinus SP I®, Valdemar Link and Spectron EF®, Smith and Nephew Richards) were evaluated. 2 types of cups were cemented (Lubinus®, Valdemar Link and Spectron EF®, Smith and Nephew Richards) and 1 was uncemented (Harris-Galante II®, Zimmer). The cemented cups were supplied with circular steel wires, but no metal-backing. All patients were investigated with repeated conventional radiographic and RSA examinations at 2 departments (Sahlgrenska University Hospital in Gothenburg and Northern University Hospital in Umeå). The mean follow-up time was 4.1 (0.5–5) years. 51 patients were included in the study. In 18 patients, 2 follow-up examinations were used, corresponding to a total of 120 radiographic examinations (Table 1).

RSA

Preoperatively, 3 of the stem designs (Ribbed stem, Tifit and Spectron stems) were marked with 5–9 0.8 tantalum markers, (Kärrholm and Snorrason 1993, Kärrholm et al. 1994, Thanner et al. 1995). The relative motion of the gravitational center of the prosthetic markers in the proximal–distal direction was used

Table 1. Number of examinations included in the study

Postoperative	2-year follow-up	5-year follow -up	Total
51	51	18	120

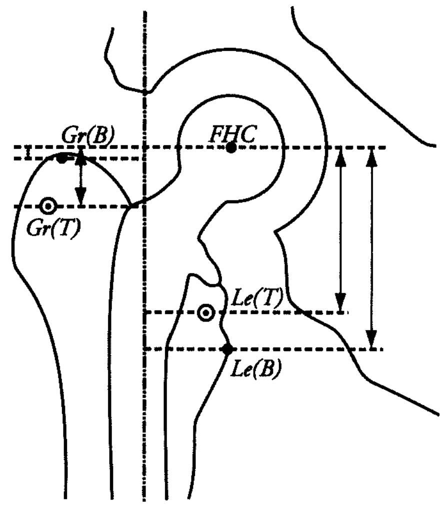

Figure 1. The femoral and femoral head center (FHC) landmarks. The 4 distances measured are shown.

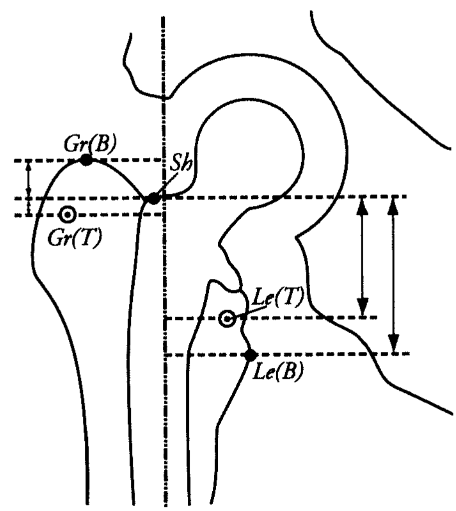

Figure 2. The femoral and stem shoulder (Sh) landmarks. The 4 distances measured are shown.

to represent migration in these cases. In the Lubinus stems, the corresponding migration of the femoral head center (Baldursson et al. 1979) was used to represent migration. The discrepancy in the choice of prosthetic landmarks in the RSA evaluation was supposed to have a negligible influence on the results. At surgery, 3–7 tantalum markers inserted into the greater and 1–3 into the lesser trochanter were implanted to obtain bony landmarks.

During surgery, 6–9 0.8 mm tantalum markers were inserted into the polyethylene liner of the porous-coated HG cups or the polyethylene Spectron cups. The Lubinus cups were supplied with tantalum markers by the manufacturer. The pelvis was prepared for RSA by implantation of 4–9 markers, before the insertion of the acetabular prosthesis. RSA examinations were performed postoperatively (4–6 days) and, thereafter, at regular intervals. All patients were scheduled for follow-up examinations at 6 months, 1, 2, and 5 years after the operation. Proximal–distal and medial–lateral migration of the cup center were measured. The error of the RSA examinations varied, depending on the prosthetic design, between 0.15 and 0.35 mm, 99 percent confidence limits (Kärrholm and Snorrason 1992, 1993, Snorrason et al. 1993, Kärrholm et al. 1994, Thanner et al. 1995), but was disregarded in this study.

Measurements of conventional radiographs

All patients underwent plain radiographic examination on the same day as the RSA examination. For each examination, an AP pelvis view centered on the symphysis, and AP and lateral views centered on the proximal part of the stem were taken. The conventional radiographs were measured using a pencil and ruler. They were also video-captured and through a frame grabber card in a personal computer digitized with a resolution of 1024 × 1024 pixels and 256 grayscales. Software was developed for the correction of planar distortion. The digitized image was displayed on a TV monitor. An image analysis software (Research Metrics®, OrthoGraphics Inc., Salt Lake City, Utah, U.S.A.) computed distances based on landmarks defined by the examiner.

Femoral landmarks

2 prosthetic landmarks were evaluated to measure the proximal–distal migration of the stem (Figures 1 and 2).

1. Femoral–head center (FHC). The head center was determined by using 3 points on the periphery of the femoral head. On the digitized radiographs, the center was computed by the software. In the manual evaluation, concentric circle templates were laid over the femoral head to determine its center (Livermore et al. 1990). In 1 of the stems (Tifit) we were unable to identify the periphery of the ceramic head with sufficient accuracy on the digitized radiographs. This reference point was therefore omitted in the digital evaluation of the Tifit stem which explains the difference in numbers of examinations for the different landmarks.

2. The shoulder of the stem corresponds to its proximal–lateral edge (Sh).

The long axis of the femur was defined and the per-

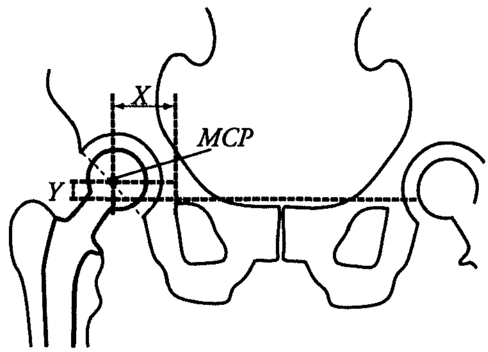

Figure 3. The pelvic and cup landmarks. Horizontal MCP (mid-center of cup)–teardrop distance (x) and vertical MCP teardrop line distance (y) are illustrated.

Table 2. Absolute mean + 2 standard deviations for manual and digital measurements of the stem

	Accuracy (mean + 2SD)	
Landmarks	Manual	Digital
Le(B)-Sh	5.3	4.1
Le(T)-Sh	3.9	4.1
Le(B)-FHC	6.1	6.6
Le(T)-FHC	5.9	6.4
Gr(B)-Sh	9.1	8.4
Gr(T)-Sh	6.9	6.3
Gr(B)-FHC	11.9	12.2
Gr(T)-FHC	9.4	12.3

pendicular projections of 4 femoral landmarks on this line were identified as follows:

1 Le(B). The most medial point on the lesser trochanter.

2 Gr(B). The tip of the greater trochanter.

3 Le(T). The most medial tantalum marker in the lesser trochanter region.

4 Gr(T). The most cranial tantalum marker in the greater trochanter.

4 different distances between bone and implant landmarks were calculated. Correction for out-of-plane angulation between the stem and the film plane was made according to Jones et al. (1988). No correction for magnification was made.

Pelvic landmarks

The teardrop and the midpoint of the cup diameter (MCP) represented the bony and the prosthetic landmarks, respectively. The shortest proximal–distal distance (vertical height, y-distance) between the teardrop line and the mid-cup point and the corresponding horizontal distance (x-distance) between a vertical line through the base of the teardrop (Figure 3) and the MCP were measured. No corrections were made for magnification, pelvic tilt or rotation (Sutherland et al. 1982).

Intra- and interobserver variability

62 radiographs were measured manually and on digitized radiographs by 3 of the authors (HM, JK, YXW). In addition, 1 of the authors (YXW) measured the radiographs twice at an interval of 3 weeks. The intraobserver error was represented by 2 standard deviations from 0. The interobserver error was represented by 2 pooled standard deviations, based on the measurements made by the 3 observers.

Statistics

All calculations were based on differences between the 2 methods (migration according to conventional radiography, minus migration according to RSA). The paired samples t-test, repeated measurement ANOVA for multiple comparisons and multivariate ANOVA were used. In total, differences based on 180 manually measured and 202 digitized radiographs and a corresponding number of radiostereometric examinations were included. The patients were selected from records in Umeå and Gothenburg on the basis of designs of the prostheses used and the availability of complete radiographic and stereoradiographic follow-up. Some of the conventional radiographic examinations were accessible during a limited period allowing only for video-capturing. This slightly reduced the number of manually measured radiographs.

Results

Manual vs. digitized measurements

The digital and the manual measurements did not differ (femoral components, including the Tifit prostheses without femoral head center measurements: p 0.2, excluding the Tifit prostheses—both femoral head and shoulder measurements: p 0.9, acetabular component: p 0.5).

Stem measurements (Table 2)

A comparison of the 4 stem designs revealed that none of them was associated with superior or inferior accuracy of the radiographic measurement (p 0.5–0.9).

The accuracy (absolute mean value + 2 standard deviations), including all 4 stem designs, varied from 3.9 to 12.3 mm, depending on which of the prosthet-

Table 3. Intraobserver error (n 14–19) (2 standard deviations) and interobserver error (n 29–39) (2 pooled standard deviations) for manual measurements of the stem

Landmarks	Intraobserver error	Interobserver error
Le(B)-Sh	2.6	4.4
Le(T)-Sh	2.8	2.6
Le(B)-FHC	2.8	4.2
Le(T)-FHC	2.9	4.2
Gr(B)-Sh	4.9	4.6
Gr(T)-Sh	1.6	5.2
Gr(B)-FHC	2.9	6.6
Gr(T)-FHC	5.6	3.6

ic and bony landmarks were chosen to represent the proximal–distal migration.

The shoulder landmark proved to be the most accurate ($p < 0.0001$) of the 3 designs in which both the femoral head center and the shoulder could be used as landmarks.

As regards the bony landmarks, the lesser trochanter was associated with better accuracy than the greater trochanter (p 0.001). Implanted tantalum markers improved the accuracy on both locations (greater trochanter: p 0.04, lesser trochanter: p 0.02). The migration of the femoral component could be most accurately measured if the prosthetic shoulder marker and the tantalum marker in the lesser trochanter were used (lesser vs. greater trochanter marker: $p < 0.0001$). The accuracy obtained with these 2 landmarks was 3.9 mm.

Cup measurements

The accuracy obtained for manual cup measurements was 4.4 for the x-distance and 4.4 for the y-distance (n 43). The corresponding values for the digital measurements were 4.5 and 4.4, respectively (n 43). We found the accuracy significantly better for the Spectron and Harris-Galante designs, compared to the Lubinus design when the teardrop was used as a bony landmark (p 0.02). We therefore further evaluated 3 baselines on these radiographs, the Kohler line, the obturator line and the tuber ischii line, but this did not change the result.

The accuracy (mean + 2 SD) of migration in the medial–lateral direction was 6.5 mm for the Lubinus cup (n 16). The corresponding value for migration in the proximal–distal direction was 6.3 mm.

Intra- and interobserver variability

The intraobserver error (2 SD) varied from 1.6 to 5.6 mm as regards the stem (Table 3), and 2.7 (x-distance) to 2.4 (y-distance) for the cup (n 21), respectively.

The interobserver error (2 pooled standard deviations) varied from 2.6 to 6.6 mm (stem values, Table 3). As regards the cup measurements, the corresponding error was 3.6 for the x-distance and 3.2 for the y-distance (n 47).

There were no statistical differences between the 3 observers.

Discussion

Correct interpretation of radiographs after total hip replacements is important because clinical score systems are insensitive for detection of mechanical failures (Andersson 1972, Callaghan et al. 1990). The early development of radiolucencies and migration indicates an increased risk of later clinical loosening (Harris et al. 1982, Carlsson and Gentz 1984, Braud and Freeman 1990, Kärrholm et al. 1994b, Strömberg et al. 1995). The interval between the appearance of these changes and clinical symptoms of loosening or catastrophic failures may extend to several years, which emphasizes the importance of an accurate radiographic evaluation. However, manual radiographic measurements are time-consuming and laborious. Several systems using measurements on a digitizing table or direct measurements of digitized radiographs have been introduced (Hardinge et al. 1991, Jones et al. 1988, 1992).

Using repeated measurements on the same radiographs, Loudon and Charnley (1980) estimated the precision of migration measurements below 2 mm. Ilchmann et al. (1992) evaluated the EBRA method by comparing the results with radiostereometry (RSA) and conventional radiography. Carlsson et al. (1993) analyzed tibial component migration in knee prostheses, comparing measurements by using tantalum markers on conventional radiographs (MIRA) and RSA. According to our knowledge, no comparison has been made between conventional radiography and a high precision method, such as RSA in order to evaluate the accuracy of measuring the migration of the femoral component in total hip arthroplasty.

The introduction of new technology and evaluation tools, such as digitized image analysis systems, demands careful documentation. New prostheses and the cementless designs, in particular, may change the premises for previous calculations and estimates of the precision or accuracy of measurements by apparent changes in prosthetic landmarks.

Many factors are responsible for errors in measuring prosthetic migration on conventional radiographs. At the postoperative examination, pain and any remaining extension defect will jeopardize a standardized positioning of the hip. When studies are extended over long periods of time, the radiographic equipment and the personnel will often not be the same. The point of centering of the radiography beam is subjected to variations as also is film quality.

Variations in radiographic exposure and in the formation of ectopic bone are most certainly also important reasons for the poor accuracy of the measurements, particularly when the greater trochanter is involved. Braud and Freeman (1990) found a very small measurement error when using the greater trochanter as a bony landmark. The corresponding value (interobserver error) was 10–15 times higher in our study, despite the fact that identical prophylaxis against ectopic bone formation was used. 2 different surgical approaches were employed in our series. The Lubinus and the Ribbed stems were operated through a posterior and the Spectron and the Tifit stems through a lateral transgluteal approach. More heterotopic calcifications may be expected after the transgluteal approach, but the accuracy of the measurements did not differ between the various stem designs. A difference in the reproducibility of the centering technique, implant design and definition of the interobserver error may, at least partly, explain the discrepancy between our results and those of Braud and Freeman (1990).

The interpretion of radiographs is, however, associated with several types of uncertainties and, in addition, the true magnification may be difficult to establish. In the digitized image analysis, we used the femoral head diameter to estimate the magnification, which was found to be about 18 percent. Thus, all distances should be reduced by multiplying by a factor of about 0.85. Since the measurements of the femoral head were not accurate in all the cases, this procedure was omitted (Sutherland et al. 1982, Nunn et al. 1989, Ilchmann et al. 1992).

In our study, the use of small tantalum markers improved the accuracy. In a previous study, Loudon and Charnley (1980) reported a very small intraobserver variability (mean + 2 SD 0.8 mm). Such high precision could be obtained only by using an artificial landmark on the femur, the cerclage wire. Sutherland et al. (1982) reported a considerably poorer detection limit for stem migration when using the hip center and the tip of the greater trochanter as landmarks. Braud and Freeman (1990) reported an interobserver error of 0.3 mm when using a distinct tip on the shoulder of the Freeman prosthesis. In our study, optimum accuracy was obtained when the prosthetic shoulder and the tantalum marker in the lesser trochanter were used, which also indicates the need for distinct landmarks. The intra- and the interobserver variability, however, was about 1–6 mm less than the true accuracy (as compared to RSA), indicating that repeated measurements on the same radiographs by one or several observers are not sufficient to determine the true error. A reference method is preferable. If this is not possible, then 2 examinations performed on the same day by different personnel provides a better estimate of the error than repeated measurements of the same radiographs.

Many studies (Harris et al. 1982, Sutherland et al. 1982) have not addressed the problem of determining the accuracy of conventional radiography of stem prostheses, but have instead chosen an arbitrary limit for significant migration in the absence of exact determinations. The definitions of femoral component loosening proposed by Harris et al. (1982) are widely accepted, but are not based on the determination of accuracy or on the validity of migration analysis. These studies did not comment on which bony and implant landmarks were used to measure prosthetic migration. If our results are valid in general, it does not seem reasonable to regard subsidence as established, if it does not exceed 5–7 mm, provided that the measurements are based on bony landmarks. The conventional radiographic examinations in our study were done at departments with a special interest in skeletal radiology. If the technique were further standardized, preferably with the help of an image intensifier, and if all radiographs were displayed simultaneously, the accuracy would probably improve.

Manual radiographic analysis is the standard method for making radiographic measurements, but it has the disadvantage of being time-consuming. The digitized image analysis we used is 5–10 times faster, but is associated with inferior resolution. In terms of accuracy the 2 methods do not differ, perhaps because the inferior resolution of the digitiszed system was compensated for by the exactness of a superior setting (Hardinge et al. 1991, Ragnarsson et al. 1992). In future, the use of improved scanning techniques for conventional radiographs or directly digitized radiographic images may further improve the accuracy.

Accuracy in measuring cup migration is generally better than that obtained for stem migration. Sutherland et al. (1982) reported a detectable cup migration limit of 3 mm with the Kohler line, teardrop line and hip center as landmarks. Sutherland and Bresnina (1992) noted a minimum projection and magnification error of 2 mm and found no means of mathematically compensating for that error. Ilchmann et al. (1992) detected no difference between the Sutherland, Wetherell and Sulzer methods, but found significantly better agreement between the EBRA and RSA results. The accuracy increased when standardized projections were used. Amstutz et al. (1986) described a system in which the migration analysis was based on manually superimposed serial radiographs. The system was supposedly easy and it was assumed to have an accuracy of 1–3 mm, but hitherto it does not seem to have come into general use.

In our evaluation, one of the cup designs was less accurate than the others. The Lubinus series was the only one that included revised cases (n 5), in which a previous loosening might have eroded the medial acetabular wall, making the identification of the teardrop difficult. Supplementary measurements using alternative reference lines, however, did not support this theory. Another explanation might be that all Lubinus prostheses were examined at 1 of the 2 radiological departments involved in this study, in which a more stringent examination of the pelvis was introduced between the postoperative and the follow-up examinations.

Our results indicate that cup migration is detectable when it exceeds 3–4 mm. The accuracy of measurements of stem migration on conventional radiographs was shown to vary between 4 mm and up to more than 1 cm, depending on the choice of landmarks. In the prosthetic designs that we evaluated, the prosthetic shoulder and a tantalum marker inserted into the lesser trochanter were the best means of obtaining accurate measurements. Further optimizing of conventional radiography can certainly be done but, to be effective, the radiographic technique has to be consistently applied during periods corresponding to the lifetime of a modern hip prosthesis.

Acknowledgements

Financial support was given by the Swedish Medical Research Council, grant numbers B91-17x-08306-04, B94-17x-07941-08A, IngaBritt and Arne Lundberg's Research Foundation, Neubergh Foundation, Orthopaedic Research, Gothenburg's Medical Association,and the Greta and Einar Askers Foundation.

References

Amstutz H C, Ouzounian T, Grauer D, Flink C, Kirkpatrick J, Bassett L. The grid radiograph. J Bone Joint Surg (Am) 1986; 68 (7): 1052-6.

Andersson G B J. Hip assessment: A comparison between nine different methods. J Bone Joint Surg (Br) 1972; 54 : 621-5.

Baldursson H, Egund N, Hansson L I, Olsson T H, Selvik G. Instability and wear of total hip prostheses determined with roentgen stereophotogrammetry. Arch Orthop Trauma Surg 1979; 95: 257-63.

Braud P, Freeman M A R. The effect of retention of the femoral neck and of cement upon the stability of a proximal femoral prosthesis. J Arthroplasty (Suppl) 1990; 5: 5-10.

Callaghan J J, Stanley D H, Carlton F S, Hopkinson W J. Assessing the results of hip replacement. A comparison of five different rating systems. J Bone Joint Surg (Br) 1990; 72 (6): 1008-9.

Carlsson L V, Albrektsson B E J, Freeman M A R, Herberts P, Malchau H, Ryd L. A new radiographic method for detection of tibial component migration in total knee arthroplasty. J Arthroplasty 1993; 8 (2): 117-23.

Carlsson S, Gentz C-F. Radiographic versus clinical loosening of the acetabular component in non-infected total hip arthroplasty. Clin Orthop 1984; 185: 145-50.

Hardinge K, Porter M L, Jones P R, Hukins D W L, Taylor C J. Measurement of hip prosthese,s using image analysis. The Maxima hip technique. J Bone Joint Surg (Br) 1991; 73 (5): 724-8.

Harris W H, McCarthy J C, O'Neill D A. Femoral component loosening using contemporary techniques of femoral cement fixation. J Bone Joint Surg (Am) 1982; 64 (7): 1063-7.

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

Jones P R, Taylor C J, Hukins D W L, Hardinge K, Porter M L. Prosthetic hip failure: preliminary findings of retrospective radiograph image analysis. Eng Med 1988; 17: 119-25.

Jones P R, Hukins D W L, Porter M L, Davies K E, Hardinge K, Taylor C J. Bending and fracture of the femoral component in cemented total hip replacement. J Biomed Eng 1992; 14: 9-15.

Kärrholm J, Snorrason F. Migration of porous-coated acetabular prostheses fixed with screws. J Orthop Res 1992; 10: 826-35.

Kärrholm J, Snorrason F. Subsidence, hump and tip micromovements of noncoated ribbed femoral prostheses. Clin Orthop 1993; 273: 50-60.

Kärrholm J, Malchau H, Snorrason F, Herberts P. Micromotion of femoral stems in total hip arthroplasty. A randomized study of cemented, hydroxyapatite-coated and porous-coated stems, using roentgen stereophotogrammetric analysis. J Bone Joint Surg (Am) 1994a; 76 (11): 1692-1705.

Kärrholm J, Borssén B, Löwenhielm G, Snorrason F. Does early micromotion of femoral stem prostheses matter? 4-7 year stereoradiographic follow-up of 84 cemented prostheses. J Bone Joint Surg (Br) 1994b; 76 (6): 912-7.

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 (4): 516-27.

Loudon J R, Charnley J. Subsidence of the femoral prosthesis in total hip replacement in relation to the design of the stem. J Bone Joint Surg (Br) 1980; 62 (4): 450-3.

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 (4): 629-31.

Ragnarsson J I, Eliasson P, Kärrholm J, Lundström B. The accuracy of measurements of femoral neck fractures. Conventional radiography versus roentgen stereophotogrammetric analysis. Acta Orthop Scand 1992; 63 (2): 152-6.

Snorrason F, Kärrholm J, Holmgren C. Fixation of cemented acetabular prostheses. The influence of preoperative diagnosis. J Arthroplasty 1993; 8 (1): 83-93.

Strömberg C, Herberts P, Palmertz B, Garellick G. Prediction of outcome after cemented THA, based on radiographic changes. An evaluation of risk signs for mechanical loosening. Acta Orthop Scand 1995. In print.

Sutherland C J, Wilde A H, Borden L S, Marks K E. A ten-year follow-up of one hundred consecutive Müller curved-stem total hip-replacement arthroplasties. J Bone Joint Surg (Am) 1982; 64 (7): 970-82.

Sutherland C J, Bresina S J. Measurement of acetabular component migration, using two-dimensional radiography. J Arthroplasty (Suppl) 1992; 7: 377-9.

Thanner J, Freij-Larsson C, Kärrholm J, Malchau H, Wesslén B. Evaluation of Boneloc®. Chemical and mechanical properties, and a randomized clinical study of 30 total hip arthroplasties. Acta Orthop Scand 1995; 66 (3): 207-14.