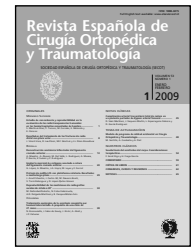
Rev esp. cir. ortop. traumatol. 2009;**53(1)**:1-7

Revista Española de Cirugía Ortopédica y Traumatología

www.elsevier.es/rot



ORIGINAL PAPERS

Concordance and reproducibility in the assessment of humeral radiolucencies in cemented shoulder hemiarthroplasties

S. Martínez-Díaz*, C. Torrens, M. Corrales, E. Melendo and E. Cáceres

Department of Orthopedic and Trauma Surgery, IMAS Hospitals, del Mar and La Esperanza Hospitals, Barcelona, Spain

Received May 16, 2007; accepted March 26, 2008

KEYWORDS:
Shoulder hemiarthroplasty;
Radiolucency;
Aseptic loosening;
Radiological analysis;
Intraobserver variability;
Interobserver variabilit

Abstract

Purpose: To assess concordance and reproducibility in the observation of radiolucent lines of the humeral prosthetic component in cemented shoulder hemiarthroplasties.

Materials and methods: Five observers evaluated 128 X-rays obtained from 32 cemented shoulder hemiarthroplasties at 2 points of time during follow-up in anteroposterior and scapular Y views; a re-evaluation was made after 3 weeks. Radiolucent lines were studied in 7 areas around the implant at the cement-bone and cement-implant interfaces. Radiolucencies were classified into 4 categories according to their size in millimeters. Results were assessed with the kappa statistical method.

Results: In the study of intraobserver reproducibility kappa indices of 0 to 0.6 were obtained for the implant-cement and cement-bone interfaces. In the study of interobserver concordance for each of the areas and interfaces, 65% of kappa indices obtained range between 0 and 0.4. Intraobserver capacity to follow up one same patient in time was assessed on the basis of the percentage of possible responses by each observer, which ranged between 85 and 90%

Conclusions: A low rate of concordance and reproducibility is obtained when analyzing radiolucencies both at the bone-cement and at the cement-implant interfaces. There is also a low error index when analyzing the evolution of radiolucencies in 2 X-rays of the same implant obtained at 2 different points in time.

* Corresponding author.
E-mail: 94064@imas.imim.es (S. Martínez-Díaz).

1888-4415/$ - see front matter

PALABRAS CLAVE
Hemiartroplastia de hombro;
Radiotransparencia;
Aflojamiento aséptico;
Análisis radiológico;
Variabilidad intraobservador;
Variabilidad entre observadores

Estudio de concordancia y reproducibilidad en la evaluación de las radiotransparencias humerales en las hemiartroplastias cementadas de hombro

Resumen
Objetivo: evaluar la concordancia y la reproducibilidad en las observaciones de las líneas de radiotransparencia del componente protésico humeral en las hemiartroplastias cementadas de hombro.
Material y método: cinco observadores han evaluado 128 radiografías obtenidas de 32 hemiartroplastias cementadas de hombro en dos momentos del seguimiento evolutivo en proyecciones anteroposterior y perfil de escápula, con una reevaluación a las 3 semanas. Se han evaluado las líneas de radiotransparencia en 7 zonas alrededor del implante en las interfaces cemento-hueso y cemento-implante. Se han clasificado las radiotransparencias en cuatro categorías según el tamaño en milímetros. Se han evaluado los resultados con el método estadístico kappa.
Resultados: en el estudio de la reproducibilidad intraobservador se han obtenido para las interfaces implante-cemento y cemento-hueso índices kappa de 0 a 0,6. En el estudio de la concordancia entre observadores para cada una de las zonas e interfaces el 65%de los índices kappa obtenidos oscilan entre 0 y 0,4. La capacidad intraobservador de seguimiento en el tiempo de un mismo paciente se ha evaluado por el porcentaje de respuestas posibles de cada observador que ha oscilado entre el 85 y el 90%
Conclusiones: hay un bajo índice de concordancia y reproducibilidad al analizar las radiotransparencias tanto en la interfaz hueso-cemento como en la de cemento-implante. También hay un bajo índice de errores al analizar la evolución de las radiotransparencias en dos radiografías del mismo implante obtenidas en dos momentos diferentes en el tiempo.

Introduction

Cemented shoulder hemiarthroplasty is a common treatment for glenohumeral arthritis and 3 and 4-part humeral head fractures[1].

Clinical and radiological study of glenoid component loosening has become one of the main problems of total shoulder replacement in the last 2 décades[2-5]. Radiolucent lines, both around the glenoid and the humeral components, in particular those appearing in the immediate postoperative period, have been attributed to a poor cementing technique[6,7]. Progression of these radiolucent lines with time, as well as the appearance of new ones, has led to the conclusion that there are other factors involved in their development such as implant design, surgical technique, bone quality, individual tissue reactions, glenohumeral stability and each patient's activity level[7,8].

Evaluation of radiolucent lines in the postoperative radiologic follow-up provides reference images that will subsequently allow the monitoring and diagnosis of both septic and aseptic implant loosening. The study of radiolucent lines in prosthetic follow-up was first reported for the hip joint almost 3 decades ago in the work of Gruen et al[9] on cemented femoral component loosening, and subsequently, by Barrack et al[10].

The main problem with any image-based method of classification or diagnosis is the ability of different observers or a single observer on different occasions (intra-observer evaluations) to reproduce it. In this paper we aim to find out whether radiological evaluation of radiolucent lines in shoulder hemiarthroplasty is reproducible by both inter and intra observer evaluations at two different points in time and whether ii is a reliable method for conveying information on prosthetic loosening and following up cases of loosening (figs. 1 y 2).

The purpose of the present paper is to evaluate the concordance and reproducibility in the observation of radiolucent lines around the humeral component in shoulder hemiarthroplasties.

Materials and methods

A concordance and reproducibility study was performed with 5 observers that had attended the Surgery Department of our hospital for a minimum of 6 months to be trained in the assessment of radiolucent lines according to the criteria laid down in the present paper. These 5 observers evaluated a group of 32 patients implanted with a shoulder hemiarthroplasty further to osteoarthritis or fracture, who had been subjected to 2 follow-up x-ray studies with two different views in each study (total: 128 radiographs). These 128 x-rays were evaluated by the 5 observers on 2 occasions with an interval of 3 weeks between them. The evaluation was made without observers knowing which x-ray corresponded to which patient or being aware that each time they were evaluating all 4 follow-up x-rays of a single patient (2 anteroposterior and 2 scapular Y views).

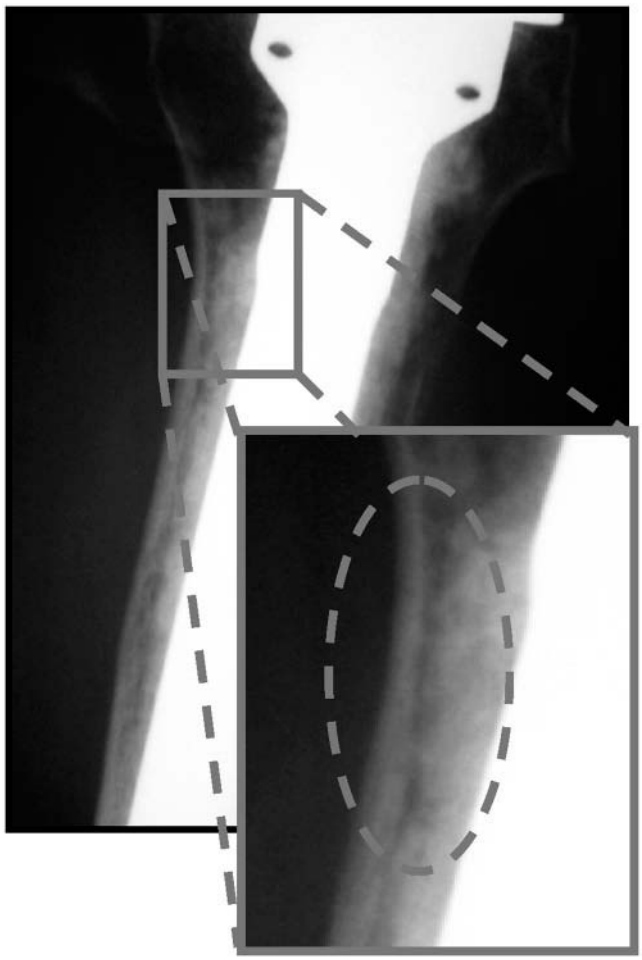


Figure 1 Radiolucent line at the bone-cement interface. Close-up of a case of shoulder hemiarthroplasty.

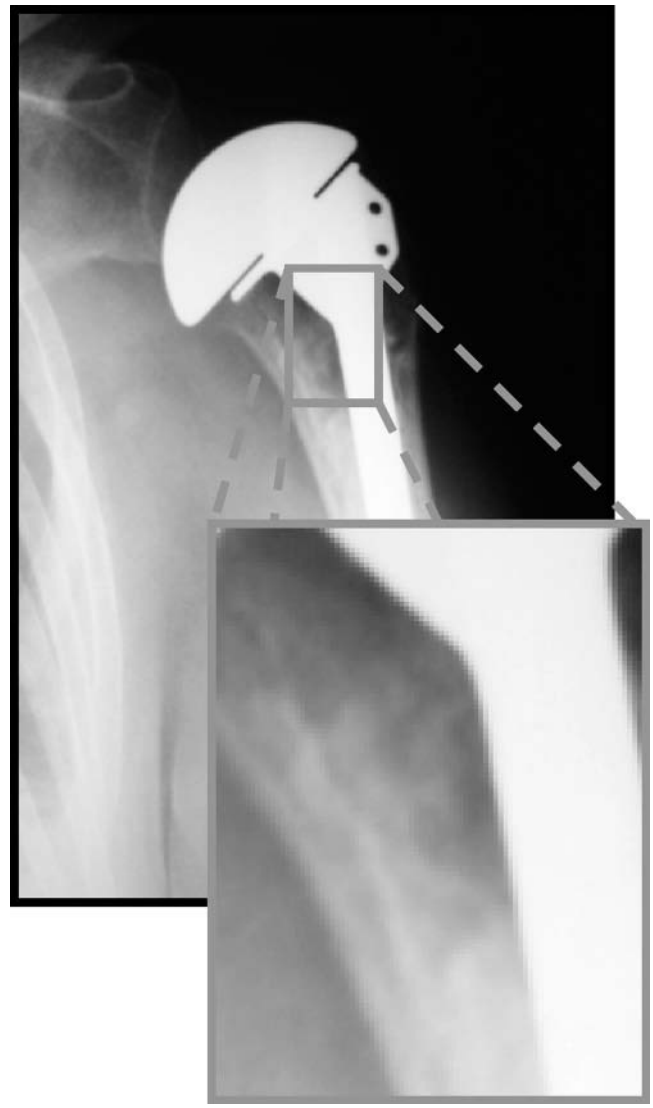


Figure 2 Radiolucent line at the bone-cement interface. Close-up of another case of shoulder hemiarthroplasty.

Radiolucent lines were evaluated in the 2 anteroposterior and scapular Y views, with 7 zones being analyzed in each x-ray view. Each zone and each of the 2 interfaces (bone-cement and cement-bone) were classified into 4 categories according to the size of the radiolucency (fig. 3). A value of 1 was given if there was no radiolucency present, 2 if it was larger than 1 mm, 3 if the radiolucency was between 1 and 3 mm and 4 if it was larger than 3 mm[11]. Thus, at each evaluation, each of the 5 observers assessed 128 views, 896 zones and 1,792 interfaces, which added up to a total 17,920 interfaces evaluated in the study.

An analysis was made of intraobserver reproducibility, i. e. the number of times an observer agreed with his or her own prior evaluations when assessing the same zone and the same interface at two different points in time, separated by 3 weeks. In addition, inter-observer concordance was also determined, i.e. the Lumber of times 2 observers agreed in one same evaluation. For that purpose, a comparison was made between the responses that each observer had given for each of the zones and interfaces with those given by the other 4 observers.

In order to assess the capacity of an observer to follow the evolution of the radiolucencies of one same patient, each observer was required to conduct, for each series, a blind evaluation of 2 x-rays (with 2 views per x-ray; one anteroposterior and the other a scapular Y view) obtained at 2 different points in time during follow-up of a single patient. The percentage of possible replies given by each observer in the course of this follow-up was determined. Only 2 possible responses were considered with respect to radiolucencies, for the same zone and interface and for the same patient: "the radiolucency has remained stable" and "the radiolucency had increased with time and was unlikely to improve."

The analysis of results was made by means of the kappa statistical method, which measures the degree of agreement between 2 observers, with a range between 0 and 1. Following Landis and Koch, different categories were established (table 1)[12]. The kappa index was also used to analyze inter-observer concordance. Since all possible comparisons were made between the observers, a decision

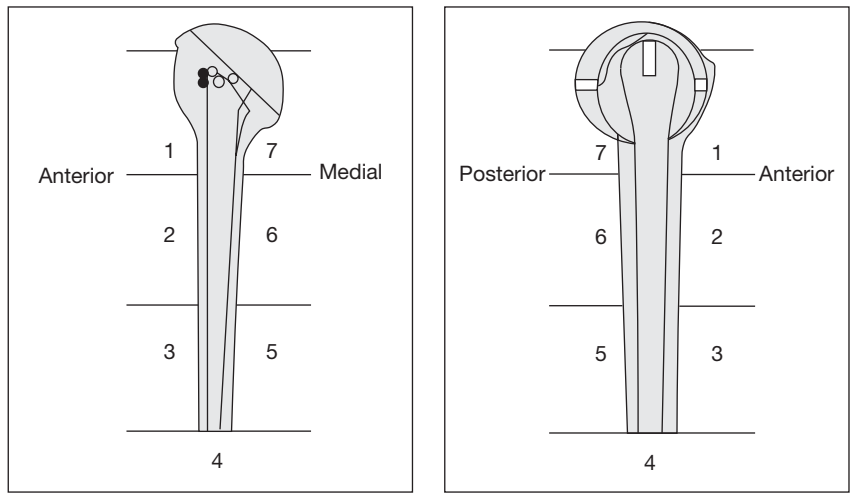


Figure 3 Data collection matrix with the 7 radiolucency zones in each of the x-ray views and interfaces.

Table 1 Value of the kappa index and the levels of concordance (Landis and Koch)

Kappa index	Level of concordance
<0	No concordance
0–0.2	Negligible
0.2–0.4	Low
0.4–0.6	Moderate
0.6–0.8	Good
0.8–1	Very good

was made to use alpha <0.001 as the significance threshold in order to reduce the error that could arise from the multiple comparisons made.

Results

Results were divided into three groups: intraobserver reproducibility, interobserver concordance and follow-up of a single patient over time.

Intra-observer reproducibility

Intra-observer reproducibility is the number of times an observer's evaluation coincides with a previous evaluation. We obtained the kappa indices for the implant-cement interface, both for the anteroposterior view and the scapular Y view, which ranged between 0 and 0.6; these are considered negligible/moderate scores (table 2). Kappa indices for the cement-bone interface in the analyses of anteroposterior and scapular Y views also ranged between0 and 0.6, i-e. negligible/moderate concordance scores (table 3). Both for the implant-cement and the cement-bone interfaces we obtained kappa indices ranging between 0 and 0.6, i-e. negligible, low and moderate scores (tables 2 and 3).

Table 2 Kappa indices related to reproducibility at the implant-cement interface in the 2 x-ray views

	Anteroposterior view	Scapular Y view
Zone 1	0.6221 (0.469–0.6958)	0.6555 (0.47–0.87)
Zone 2	0.405 (0.263–0.623)	0.5041 (0.15–0.73)
Zone 3	0.5052 (0.1423–0.7643)	0.5277 (0.12–0.81)
Zone 4	0.2855 (0.0498–0.4476)	0.2864 (0.01–0.41)
Zone 5	0.0174 (0.012–0.0456)	0.5791 (0.14–0.84)
Zone 6	0.2489 (0.072–0.2972)	0.5215 (0.15–0.79)
Zone 7	0.5712 (0.41–0.68)	0.6607 (0.45–0.89)

Data express the kappa index (95%confidence interval).

Table 3 Kappa indices obtained when studying reproducibility at the cement-bone interface in both x-ray views

	Anteroposterior view	Scapular Y view
Zone 1	0.6531 (0.52–0.72)	0.6217 (0.46–0.78)
Zone 2	0.5151 (0.21–0.64)	0.5552 (0.26–0.74)
Zone 3	0.6016 (0.30–0.74)	0.6274 (0.33–0.81)
Zone 4	0.2068 (0.11–0.66)	0.0342 (–0.10 a 0.281)
Zone 5	0.5972 (0.25–0.82)	0.5742 (0.26–0.79)
Zone 6	0.4897 (0.18–0.69)	0.5937 (0.29–0.79)
Zone 7	0.6249 (0.48–0.70)	0.6221 (0.469–0.6958)

Data express the kappa index (95%confidence interval).

Inter-observer concordance

The kappa method measures the degree of concordance between 2 observers. However, we have not found in the literature any method that allows comparison of multiple kappa indices for 5 observers. Therefore, in the inter-

Table 4 Comparison of concordance between the 5 observers for each of the x-ray views and interfaces

	Anteroposterior view, implant-cement	Scapular Y view, implant-cement	Anteroposterior view, cement-bone	Scapular Y view, cement-bone
1/2	0.064±0.022	0.077±0.020	0.343±0.021	0.304±0.022
1/3	0.171±0.026	0.043±0.024	0.307±0.020	0.149±0.020
1/4	0.091±0.019	0.041±0.023	0.299±0.020	0.261±0.022
1/5	0.171±0.019	0.049±0.012	0.290±0.018	0.204±0.017
2/3	0.684±0.027	0.049±0.012	0.518±0.020	0.149±0.020
2/4	0.540±0.026	0.662±0.028	0.539±0.020	0.261±0.022
2/5	0.336±0.020	0.289±0.028	0.448±0.019	0.204±0.017
3/4	0.520±0.025	0.630±0.029	0.529±0.020	0.455±0.022
3/5	0.358±0.021	0.262±0.018	0.519±0.020	0.387±0.020
4/5	0.364±0.021	0.286±0.019	0.499±0.020	0.438±0.020)

The data reflects the kappa index ± standard error.

Table 5 Percentage of possible inter-observer responses during follow-up for the bone-cement interface in each of the x-ray views and zones

	Anteroposterior view	Scapular Y view
Zone 1	88.1%	90.6%
Zone 2	88.8%	92.8%
Zone 3	93.5%	93.1%
Zone 4	96.6%	97.2%
Zone 5	94.1%	91.9%
Zone 6	91.9%	92.2%
Zone 7	88.8%	90.9%

Tabla 6 Percentage of possible inter-observer replies during follow-up for the cement-bone interface in each of the 2 views and zones

	Anteroposterior view	Scapular Y view
Zone 1	85.9%	85.0%
Zone 2	69.7%	69.9%
Zone 3	72.5%	81.5%
Zone 4	97.5%	96.3%
Zone 5	77.4%	77.8%
Zone 6	70.9%	73.1%
Zone 7	87.5%	85.0%

observer concordance study, i.e. number of times 2 observers agree on an evaluation, each observer was compared with the other 4 for each of the zones and interfaces. Note that comparisons with observer 5 are included in the comparisons of the other 4 observers (table 4). After reviewing these results, we concluded that 65%of the kappa indices obtained reflected negligible or low levels of inter-observer concordance, i.e. between 0 and 0.4. Twenty-five percent of kappa indices showed moderate concordance and only 10% a good concordance score. No kappa index showed a very good inter-observer concordance score.

Follow up of a single patient over time

We calculated the percentage of possible replies given by each observer during follow-up for one single patient, as well as the percentage of possible replies at the implant-cement interface in anteroposterior and scapular Y views (table 5) and the replies for the cement-bone interface in anteroposterior and scapular Y views (table 6). The percentage of possible replies was high, i.e. 85-90%for all zones and interfaces (tables 5 and 6).

Discussion

Studies of periprosthetic radiolucent lines to assess both septic and aseptic loosening rates in arthroplasties have been carried out for almost 30 años[9,10]. Numerous analyses have been conducted to try and establish some kind of connection between radiolucent lines (and their progression) with glenoid component loosening in total shoulder prostheses, which occurs in up to 90% of patients[5,13–15]. Radiolucent lines in the glenoid have been evaluated according to the types of fixation[13,14,16] and to the differences between the radius of curvature of the glenoid and the humeral head[15]. Studies have been carried out even of the variability in the observation of glenoid radiolucent lines depending on the degree of version when the x-ray was taken. Results have been discouraging regarding the potential validity of these measurements for following-up the loosening process.[17].

Numerous papers have attempted to correlate radiolucent lines with the clinical and functional results of prosthetic components, mostly in the glenoid component of total shoulder implants[4,5,13–17] but rarely in the humeral component[3,6,8]. None of these studies has validated intra-

and inter-observer variability in the study of radiolucent lines.

From the early days of prosthetic cementing techniques with the Sir Charnley's hip implants over 40 years ago, the study of radiolucent lines around the implant has been the subject of debate. These radiolucent lines have been evaluated in order to determine survivorship and, therefore, both septic and aseptic loosening of total hip arthroplasty during follow-up[9,10,18-20]. Other authors have tried to validate inter-observer variability for evaluating radiolucent lines both around the femoral[21-23] and the acetabular components[24], obtaining low, moderate or good degrees of inter- and intra-observer concordance in the different studies. However, these may be considered insufficient to warrant a comparison of the results of the different series. Work like that by McCaskie et al[22] and Kneif et al[24] do show a trend toward a higher degree of intra-observer concordance, similar to that obtained in our series in the follow-up of one single patient.

The study of radiolucent lines for the follow-up and post-operative assessment of total knee arthroplasties is also quite profuse in the literature. Attempts have been made to correlate femoral and tibial radiolucent lines around total and unicompartmental knee prostheses, both with the different surgical techniques and with implant survivorship, without reaching hard-and-fast conclusions[25,26]. The study of reproducibility and concordance for the evaluation of radiolucent lines around knee arthroplasties has also produced discouraging results according to the literature[27,28]. These results, with low and moderate intra- and inter-observer scores, are similar to those obtained in our series. More recent studies evaluating radiolucent lines in total knee prostheses, such as the paper by Bach et al[29], as modified by the Knee Society Total Knee Arthroplasty Radiographic Evaluation and Scoring System (introduced in 1989 to promote unity and concordance in the postoperative x-ray assessment of the results of total knee arthroplasties), obtain excellent intra-observer concordance for all components as well as excellent inter-observer concordance for tibial and patellar components, but not for inter-observer evaluations of the femoral components, where concordance is low.

We thought that, for this study, assessment of radiolucent lines by Experts was unnecessary since an observational and descriptive study has already been conducted, which did not obtain a clinical diagnosis of loosening, which is one of the criticisms leveled at some of the prior studies of other aerthroplasties[18,19,25,26].

To conclude we could say that our results indicate that an analysis of radiolucent lines in shoulder hemiarthroplasty shows concordance and reproducibility levels, both at the bone-cement and at the cement-implant interfaces in all zones evaluated. This means that the evaluation of radiolucent lines in shoulder hemiarthroplasties is reproducible neither between different observers nor within one single observer.

On the other hand, we obtained a low error rate when analyzing the evolution of radiolucencies in two x-rays of the same implant obtained at two different points in time, which means that radiolucent line evaluation has a very high predictive value for the evaluations made by each observer of one single patient.

If we transfer the results obtained to clinical practice, we could conclude that the assessment of radiolucent lines in shoulder hemiarthroplasties do not show concordances across different specialists or between different patients evaluated by the same Observer. However, such an assessment may be of use for the follow-up carried out by each specialist of a single patient over time.

Conflict of interests

The authors have not received any financial support in the preparation of this article. Nor have they signed any agreement entitling them to receive benefits or fees from any commercial entity. Furthermore, no commercial entity has paid or will pay any sum to any foundation, educational institution or other non-profit-making organization to which they may be affiliated.

References

1. Pfahler M, Jena F, Neyton L, Sirveaux F, Mole D. Hemiarthroplasty versus total shoulder prosthesis: results of cemented glenoid components. J Shoulder Elbow Surg. 2006;15:154-63.
2. Wirth MA, Rockwood CA. Complications of shoulder arthroplasty. Clin Orthop. 1994;307:47-69.
3. Torchia ME, Cofield RH, Settergren CR. Total shoulder arthroplasty with the Neer prosthesis: long-term results. J Shoulder Elbow Surg. 1997;6:495-505.
4. Wallace AL, Phillips RL, MacDougal GA, Walsh WR, Sonnabend DH. Resurfacing of the glenoid in total shoulder arthroplasty: a comparison, at a mean of five years, of protheses inserted with an without cement. J Bone Joint Surg Am. 1999;81:510-8.
5. Gupta S, Van der Helm FCT, Van Keulen F. Stress analysis of cemented glenoid prostheses in total shoulder srthroplasty. J Biomechanics. 2004;37:1777-86.
6. Neer CS, Watson KC, Stanton FJ. Recent experience in total shoulder replacement. J Bone Joint Surg Am. 1982;64:319-37.
7. Boyd AD, Thomas WH, Scott RD, Sledge CB, Thornhill TS. Total shoulder versus hemiarthroplasty: indications for glenoid resurfacing. J Arthroplasty. 1990;5:329-36.
8. Kelly IG, Foster RS, Fisher WD. Neer total shoulder replacement in rheumatoid arthritis. J Bone Joint Surg Br. 1987;69:723-6.
9. Gruen T, McNeice G, Amstutz H. Modes of failure of cemented stem-type femoral components: a radiographic analysis of loosening. Clin Orthop. 1979;141:17-27.
10. Barrack R, Mulroy R, Harris W. Improved cementing techniques and femoral component loosening in young patients with hip arthroplasty: a 12-year radiographic review. J Bone Joint Surg Am. 1992;74:385-9.
11. Sperling JW, Cofield RH, O'Driscoll SW, Torchia ME, Rowland CM. Radiographic assessment of ingrowth total shoulder arthroplasty. J Shoulder Elbow Surg. 2000;9:507-13.
12. Joshep L. Statistical methods for rates and proportions. J Wiley, 1981.
13. Rahme H, Mattsson P, Larsson S. Stability of cemented all-polyethylene keeled glenoid components: a radiostereometric study with a two-year follow-up. J Bone Joint Surg Br. 2004;86: 856-60.
14. Lazarus M, Jensen K, Southworth C, Matsen F. The radiographic evaluation of keeled and pegged glenoid component insertion. J Bone Joint Surg Am. 2002;84:1174-82.
15. Walch G, Edwars B, Boulahia A, Boileau P, Molé D, Adeleine P. The influence of glenohumeral prosthetic mismatch on glenoid

radiolucent lines. Results of a multicenter study. J Bone Joint Surg Am. 2002;84:2186-91.
16. Boileau P, Avidor C, Krishnan SG, Walch G, Kempf JF, Mole D. Cemented polyethylene versus uncemented metal-backed glenoid components in total shoulder arthroplasty: a prospective, double-blind, randomized study. J Shoulder Elbow Surg. 2002;11:351-9.
17. Havig M, Kumar A, Carpenter W, Seiler J. Assessment of radiolucent lines about the glenoid. J Bone Joint Surg Am. 1997;79:428-32.
18. Tigges S, Stiles R, Roberson R. Complications of hip arthroplasty causing periprosthetic radiolucency on plain radiographs. AJR Am J Roentgenol. 1994;162:1387-91.
19. Iwaki H, Scott G, Freeman MAR. The natural history and significance of radiolucent lines at a cemented femoral interface. J Bone Joint Surg Br. 2002;84:550-5.
20. Ebramzadeh E, Normand P, Sangiorgio S, Llinás A, Gruen TA, McKellop H. Long-term radiographic changes in cemented total hip arthroplasty with six designs of femoral components. Biomaterials. 2003;24:3351-63.
21. Harvey EJ, Tanzer M, Bobyn JD. Femoral cement grading in total hip arthroplasty. J Arthroplasty. 1998;13:396-401.
22. McCaskie AW, Brown AR, Thompson JR, Gregg PJ. Radiological evaluation of the interfaces after cemented total hip replacement: interobserver and intraobserver agreement. J Bone Joint Surg Br. 1996;78:191-6.
23. Kramhoft M, Gehrchen PM, Bodtker S, Wagner A, Jensen F. Inter- and intraobserver study of radiographic assessment of cemented total hip arthroplasties. J Arthroplasty. 1996;11:272-80.
24. Kneif D, Downing M, Ashcroft P, Gibson P, Knight D, Ledingham W. Peri-acetabular radiolucent lines: inter- and intra-observer agreement on post-operative radiographs. Int Orthop. 2005;29: 152-5.
25. Ritter M, Herbst S, Keating E, Faris P. Radiolucency at the bone-cement interface in total knee replacement: the effects of bone-surface prepartion and cement technique. J Bone Joint Surg Am. 1994;76:61-5.
26. Klemme WR, Galvin EG, Petersen SA. Unicompartimental knee arthroplasty: sequential radiographic and scintigrahic imaging with an average five-year follow-up. Clin Orthop. 1994;301:233-8.
27. Reckling FW, Asher MA, Dillon WL. A longitudinal study of the radiolucent line at the bone-cement interface following total joint-replacement procedures. J Bone Joint Surg Am. 1977;59: 355-8.
28. Vyskocil P, Gerber C, Barmert P. Radiolucent lines and component stability in knee arthroplasty. J Bone Joint Surg Br. 1999;81:24-6.
29. Bach CM, Bieremann R, Goebel G, Mayer E, Rachbauer F. Reproducible assesment of radiolucent lines in total knee arthroplasty. Clin Orthop. 2005;433:183-8.