■ HIP

Monitoring the quality of total hip replacement in a tertiary care department using a cumulative summation statistical method (CUSUM)

D. J. Biau,
M. Meziane,
R. S. Bhumbra,
V. Dumaine,
A. Babinet,
P. Anract

From Hopital Cochin, Paris, France

The purpose of this study was to define immediate post-operative 'quality' in total hip replacements and to study prospectively the occurrence of failure based on these definitions of quality. The evaluation and assessment of failure were based on ten radiological and clinical criteria. The cumulative summation (CUSUM) test was used to study 200 procedures over a one-year period. Technical criteria defined failure in 17 cases (8.5%), those related to the femoral component in nine (4.5%), the acetabular component in 32 (16%) and those relating to discharge from hospital in five (2.5%). Overall, the procedure was considered to have failed in 57 of the 200 total hip replacements (28.5%). The use of a new design of acetabular component was associated with more failures. For the CUSUM test, the level of adequate performance was set at a rate of failure of 20% and the level of inadequate performance set at a failure rate of 40%; no alarm was raised by the test, indicating that there was no evidence of inadequate performance.

The use of a continuous monitoring statistical method is useful to ensure that the quality of total hip replacement is maintained, especially as newer implants are introduced.

■ D. J. Biau, MD, PhD, Orthopaedic Surgeon, Fellow
■ M. Meziane, MD, Orthopaedic Surgeon
Clinique de Domont, Dpt Orthopedie, 7 rue Andre Nouet, 95330 Domont, France.

■ R. S. Bhumbra, FRCS (Tr & Orth), PhD, Orthopaedic Surgeon, Senior Registrar
Royal National Orthopaedic Hospital, Brockley Hill, Stanmore, Middlesex HA7 4LP, UK.

■ V. Dumaine, MD, Orthopaedic Surgeon
■ A. Babinet, MD, Orthopaedic Surgeon
■ P. Anract, MD, Orthopaedic Surgeon, Professor
Hopital Cochin, Dpt Orthopedie, 27 rue du Fauboug Saint-Jacques, 75014 Paris, France.

Correspondence should be sent to Dr D. J. Biau; e-mail: djmbiau@yahoo.fr

Total hip replacement (THR) is one of the most widely used and most successful orthopaedic procedures performed in developed countries.[1] The burden of revision surgery, however, has become a major issue in terms of both volume and cost.[2] Technical errors at the time of the index operation are known to be associated with an increased rate of revision.[3-7]

Standard outcomes of joint replacement, such as function, implant survival and dislocation are often chronologically remote from the primary procedure. Post-operative radiological factors such as the position and quality of fixation of the components may serve as surrogate markers for these clinically important endpoints.[3-7] Monitoring the quality of THRs using these surrogates may allow the quality of the procedure to be assessed much earlier.

Statistical methods, which have been developed for the manufacturing industry to monitor the quality of products, have come to the attention of health-care workers as a result, for instance, of the high mortality rates among children undergoing heart surgery at the Bristol Royal Infirmary,[8] or the protracted criminal activity of the general practitioner Harold Shipman.[9] Appropriate statistical methods are critical in ensuring that an adequate level of performance is maintained over time, and that if performance is inadequate it is immediately detected.[10] In orthopaedics, these methods, such as the CUSUM test (CUmulation SUMation test), have been used retrospectively to monitor the performance of a surgeon during a total knee replacement[11] and the survival of an implant.[12] This test has also been used prospectively to monitor the performance of surgeons performing THR and reconstruction of the anterior cruciate ligament.[13,14] It has been shown to be of particular value in ensuring that performance is maintained during the introduction of a new implant, or when using a technique with which a surgeon has little experience.[13]

In this study we report our experience with the implementation of a continuous monitoring process control chart to monitor the quality of surgery during THR. We report the proportion of failed THRs and propose the use of the CUSUM test for monitoring the procedure. We also looked at whether the introduction of a new design of implant was associated with a greater risk of technical failure.

Patients and Methods

Settings, selection criteria and surgical technique. The study was carried out at a tertiary care teaching hospital between October 2008 and September 2009, having been approved by the institutional review board. The inclusion criterion was any patient undergoing a primary

Fig. 1

Radiograph showing a total hip replacement with the implantation of the femoral component (Océane; Tornier, Saint-Ismier, France) according to the French paradox with stem-cortex contact.

THR. Patients were excluded if the procedure was the revision of a hemiarthroplasty or THR. Our usual surgical approach involves a partial anterior trochanterotomy,[15] although surgeons were at liberty to use another approach based on the characteristics of the patient and the nature of the operation. Femoral components were implanted according to the French paradox,[21] where cancellous bone is removed from the femoral cavity and the largest component yielding rotational stability is cemented into the femoral canal (Fig. 1). A resorbable bone plug was positioned 1.5 cm from the tip of the stem and cement introduced by retrograde filling of the femoral canal with a syringe. A cement pressuriser was not routinely used. For cemented acetabular components, the acetabulum was prepared with curved gouges and reamers. For uncemented components, the acetabulum was sequentially reamed and the components impacted line to line in accordance with the manufacturer's recommendations. This gives a press-fit of 0.75 mm to 1.8 mm, depending on the type and size of the component. If a trainee performed the operation, they were under the direct supervision of a senior surgeon or fellow.

The following cemented femoral components were available for use during the study: MK III (Stryker Benoist Girard, Hérouville, France), Dedicace (Stryker Benoist Girard) and Océane (Tornier, Saint-Ismier, France). The following acetabular components were used: MK III (Stryker Benoist Girard), Trident (Stryker Benoist Girard), Dynacup (Tornier), RM (Mathys, Bettlach, Switzerland), Medial Cup (Aston Medical, Saint Etienne, France) and Tregor (Aston Medical). An acetabular reinforcement ring (Stryker Benoist Girard) was also available.

Assessment. The quality of the procedure was assessed at the time of discharge from hospital using ten criteria (Table I), based on data from the literature showing an association with an increased risk for aseptic loosening,[3-7] poor function[7,16,17] or dislocation.[18] Anteroposterior radiographs of the pelvis and lateral radiographs of the hip were taken before discharge and evaluated using dedicated software (Roman V.1.7[19]) by two independent orthopaedic surgeons (MM, RSB), who did not participate in the care of the patients. Both rated a small set of radiographs with the main author (DJB) before rating the radiographs of the patients in the study. The anteversion angle of the acetabular component was measured using the method described by Widmer.[20] The assessment of quality of fixation of the femoral component was based on a modified version of the method described by Barrack et al[3,14] to accommodate the French paradox;[21] the presence of a lucent line within 2 mm was considered inadequate, but stem-cortex contact was considered acceptable. Overall, if one or more of these ten criteria were not met, the procedure was considered to have failed.

Patients. Between 1 November 2008 and 31 October 2009, 200 THRs were performed in our department; 150 (75%) for primary osteoarthritis, 13 (6.5%) for avascular necrosis, 11 (5.5%) for dysplasia, ten (5%) for inflammatory arthritis, eight (4%) for post-traumatic arthritis and eight (4%) for other conditions. Senior surgeons performed 156 procedures (78%). The surgical approach involved a partial anterior trochanterotomy in 185 cases (92.5%). Concomitant hardware removal was performed in nine cases (4.5%), an acetabular reinforcement ring was used in eight (4%) and bone autograft was added in five (2.5%). Post-operative inpatient radiological evaluation was not completed for four patients, one of whom died during their hospital stay; radiological evaluation was incomplete in a further eight cases because in four the contralateral hip centre was not adequate for comparative analysis (high-riding dislocation, hip with no femoral head, loose total hip replacement with high centre of rotation and contralateral bone deformity), no reliable reference was available from the lesser trochanter in three and radiographs were not adequate in one.

Statistical methods. The CUSUM test[22] has a simple and intuitive graphical representation where the score S_t is plotted on the y-axis against the observations on the x-axis. The score increases with failures and decreases with successes. If the plot crosses the limit h, an alarm is emitted indicating that performance is not adequate. The test has a holding barrier at zero that prevents the score from drifting away from the limit h with accumulation of successes. Therefore, the test remains responsive at all times to poor performance, should it occur, and past successes do not compensate unnecessarily for present failures. In the present case, based on a preliminary report[14] and on consensus among surgeons in the department, the adequate performance level (null hypothesis, also known as the target) was set at 20% failure rate and the inadequate level of performance was set at 40% failure rate (alternative hypothesis). Based on

Table I. Description of the ten criteria considered for assessing quality

Criteria		Inadequate if…	Method of measurement
Technique	1	Fracture of the acetabulum requiring fixation or a modification in the surgical technique; intra-operative revision of the implants	As per operative note
	2	Fracture of the femur requiring fixation or a modification in the surgical technique; intra-operative revision of the implants	As per operative note
	3	-1 cm > leg-length difference > 2 cm	Radiograph anteroposterior pelvis
Stem	4	-4° > stem inclination angle > +4°	Radiograph anteroposterior pelvis
	5	Stem cement mantle modified Barrack[3] C or D inadequate in eight zones or more	Radiograph anteroposterior and lateral hip
Cup	6	+30° > cup abduction angle > +55°	Radiograph anteroposterior pelvis
	7	0° > cup version angle > +30°	Radiograph anteroposterior pelvis and hip
	8	Lucent line ≥ 1 mm in one zone or more for cemented cups or a gap ≥ 3 mm for uncemented cups	Radiograph anteroposterior pelvis
Clinical	9	Dislocation or re-operation for any reason during hospital stay	As per hospital record
	10	Death, regardless of the reason during hospital stay	As per hospital record

Table II. Monthly and overall description of patients' and surgeons' characteristics, with the number of failed total hip replacements per technique, stem, cup, clinical and overall criteria

Variable	Months												Total	NAs[*]
	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct		
Number of patients	10	21	21	21	26	20	12	20	11	8	16	14	200	0
Surgeon grade (senior)	9	14	16	16	20	15	9	16	9	6	13	13	156	0
Diagnosis (primary)	6	11	18	19	21	10	10	19	7	5	14	10	150	0
Cup (cemented)	5	12	11	8	8	11	8	9	8	3	6	6	95	0
Stem (cemented)	10	21	21	21	26	20	12	20	11	8	16	14	200	0
Intra-operative revision	0	1	1	0	0	0	1	0	0	0	0	0	3	0
Intra-operative fracture/fixation	0	1	1	0	0	0	0	0	1	0	0	0	3	0
Leg-length inequality	0	0	1	1	1	1	2	1	1	0	1	2	11	12
Stem angulation (outside -4° to +4° range)	0	0	0	0	1	0	0	0	0	1	0	0	2	6
Stem fixation (Barrack C or D[3])	1	0	0	1	0	2	0	1	2	0	0	0	7	4
Cup abduction (outside 30° to 55° range)	0	1	2	5	2	0	2	1	0	1	0	0	14	4
Cup version (outside 0° to 30° range)	0	2	2	1	0	0	0	0	0	1	0	0	6	11
Cup fixation (inadequate)	0	0	1	3	3	0	2	1	0	2	2	2	16	4
Dislocation/re-operation	0	1	0	0	0	1	1	0	0	0	0	1	4	0
Death	0	0	0	0	0	0	0	0	1	0	0	0	1	0
Technique failed[†]	0	2	3	1	1	1	3	1	2	0	1	2	17	nap
Stem failed[‡]	1	0	0	1	1	2	0	1	2	1	0	0	9	nap
Cup failed[§]	0	3	4	8	4	0	4	1	0	4	2	2	32	nap
Discharge failed[¶]	0	1	0	0	0	1	1	0	1	0	0	1	5	nap
Overall failed procedure	1	5	7	9	6	4	7	3	4	4	3	4	57	nap

* NAs, missing values; nap, not applicable

† technique failure considers any failure of criteria 1, 2 or 3 from Table I

‡ stem failure considers any failure of criteria 4 and 5 of Table I

§ cup failure considers any failure of criteria 6, 7 or 8 from Table I

¶ discharge failure considers any failure of criteria 9 or 10 from Table I

computer simulations, a limit of h = 8.8 was chosen to yield a true discovery rate (TDR) of at least 99% over 200 procedures (the estimated number of procedures performed per year). The corresponding false discovery rate (FDR) was 0.17%. Specific meetings were held every three to four months to present the results of the CUSUM test and the radiographs of failed procedures. Consensus was obtained over the failures, technical issues were discussed and solutions were proposed. In order to test whether the introduction of new implants was associated with more technical failures, a two-sided Fisher's test at a significance level of 5% was performed to compare the proportion of cups and stems that failed when a new implant was used.

Results

Overall, 57 of the 200 THRs (28.5%) were considered to have failed. Technical factors were responsible for failure in 17 procedures (8.5%) (Table II), factors relating to the femoral component in nine (4.5%), those relating to the acetabular component in 32 (16%) and factors relating to
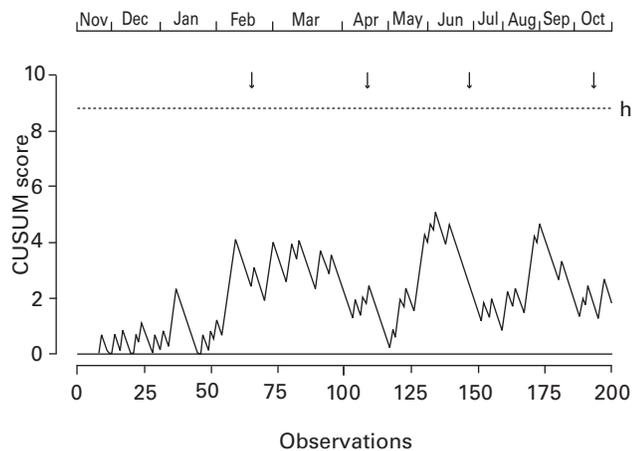
Fig. 2

Graph showing the CUSUM test over 200 total hip replacements. $p_0 = 0.2$, $p_1 = 0.4$, and $h = 8.8$. h is the decision limit; arrows indicate the dates when meetings were held in the department. No alarms were raised, indicating that performance was maintained throughout the monitoring period.

discharge from hospital in five (2.5%). Individual factors causing failure most often were fixation of the acetabular component (16 procedures (8%)), the abduction angle (14 procedures (7%)), and leg length (11 procedures (5.5%)).

Throughout the monitoring period no alarm was raised by the CUSUM test to indicate that performance was inadequate (Fig. 2). Nonetheless, technical issues were discussed during the four meetings conducted during the year. The fixation of newly introduced uncemented acetabular components was related to failure because of a gap ≥ 3 mm at the dome region (Fig. 3). It was thought that this was either due to inappropriate reaming or because of over-sizing of the component. For the abduction angle and leg-length inequality it was emphasised that care be taken during the procedure to position the components correctly, and that an intra-operative radiograph be taken if there was any doubt about the position of a component or routinely when implanting newly introduced acetabular designs.

The proportion of failed acetabular components increased significantly with the use of a new design compared to those that had been used before (17 failures in 63 hips (27%) *versus* 15 failures in 137 hips (11%), Fisher's exact test; p = 0.0064). However, the proportion of failed femoral components of a new design did not differ significantly from that of femoral components with a known design (one failure in 51 hips (2%) *versus* eight failures in 149 hips (5%), Fisher's exact test; p = 0.45).

Discussion

The emergence of quality control methods to monitor the performance of surgeons is driven by health-care authorities, patients and doctors themselves.[8,23,24] The need for these methods is reflected in both intra- and inter-unit performance variation at any given time,[25] as well as over a

period of time.[14] Therefore, one should not assume that the level of care provided to patients is adequate based on the publication of reports from other centres or other providers at any particular time. Indeed, only a continuous evaluation of performance will actually inform patients, doctors and health authorities of the true level of care provided. Continuous monitoring of the quality of THRs was introduced in our department in 2006 to ensure that the quality was maintained over time,[14] especially when new implants were introduced.

Based on previous work, a target failure rate of 20% was chosen.[4,16,26-29] In a similar previous study[14] we targeted a 10% failure rate, but this was later felt too difficult to attain: too many alarms were raised and the alarms lost their capacity to inform. Consequently, we decided to aim for a more conservative approach. However, a 20% failure rate may, in fact, be particularly difficult to achieve, given that the final outcome is a composite of ten criteria and that previous authors have reported a significant rate of failure for some of these criteria alone. To the best of our knowledge, no other centre has reported the proportion of combined technical errors. Other authors have looked at only one or two factors. Bosker et al,[26] in a prospective study of 200 consecutive THRs performed by various surgeons at a single centre, reported a 17% probability that the version angle of the acetabular component would be outside the +5°/+25° target, a 15% probability that the abduction angle would be outside the +30°/+50° target, and a combined 29% probability that the component would be outside one of these two targets. Chambers et al,[4] in a case-control comparative study of femoral components which had failed due to aseptic loosening and those that had survived, reported that the prevalence of an inadequate cement mantle (Barrack C or D) was 19% in a randomly chosen sample of 60 patients, and 69% in the 22 failed cases. García-Cimbrelo et al,[27] in a retrospective study of acetabular components after 452 Charnley THRs performed by various surgeons, reported that the prevalence of a demarcation line in one zone or more on the immediate post-operative radiographs was 48%. Kobayashi et al,[30] in a retrospective study of 328 acetabular components performed by various surgeons at different institutions, reported that the prevalence of a demarcation line in one zone or more on the one-year post-operative radiographs was 30%. White and Dougall,[31] in a prospective multicentre study of 200 THRs performed freehand by various surgeons, reported that the prevalence of leg-length inequality outside the -10 mm/+20 mm target was 10%.

The most common reasons for failure were inadequacy of fixation of the acetabular component and an inappropriate abduction angle and leg length. Fixation of the acetabular component was the most common cause of failure during the monitoring period, and this was seen mostly with the use of newly introduced press-fit acetabular components. Despite fixation being achieved at the periphery, we believe that a gap of > 3 mm demonstrates inadequate
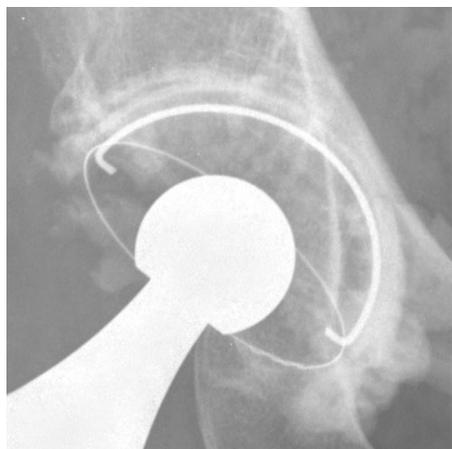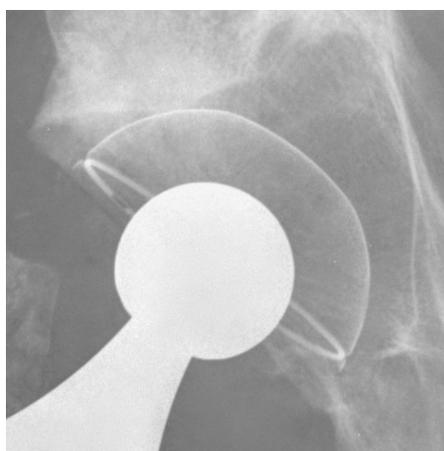
Fig. 3a

Fig. 3b

Fig. 3c

Examples of post-operative radiographs with inadequate fixation of the acetabular component (see Table I, criterion 8), showing a) the MK III acetabular component (Stryker Benoist Girard, Hérouville Saint Clair, France), b) the Dynacup component (Tornier, Saint-Ismier, France), and c) the RM component (Mathys, Bettlach, Switzerland).

technique. Inadequate reaming, due either to the inaccuracy of acetabular reamers or to the implantation of oversized components, is the usual cause for a gap at the dome.[32-34]

In our study, important gaps at the dome could also be explained by inadvertently reaming too far with small reamers (Fig. 3). These factors certainly played a role in the failures. During the study it was therefore recommended after the third meeting that intra-operative radiographs should be routinely obtained when using one of the new acetabular components (Dynacup, Tornier; and RM, Mathys). Compared to our previous report, in which we had a failure rate of 34%, our results have improved.[14] This may be due to improvements in our technique for cementing the femoral component which was a significant problem during the last period of monitoring. Nonetheless, technical issues, such as positioning the acetabular component appropriately, remain, and there may be a limit to the accuracy and consistency that one can reach with a freehand technique. The use of computer-assisted surgery may be necessary to improve performance further.[35-37]

The interpretation of some of the radiological criteria, such as the abduction angle of the acetabular component or leg length, and extrapolating those criteria to outcome, might be considered to be a limitation of this study. We decided to set limits at the widest acceptable range to avoid penalising surgeons excessively. For instance, a large range of anteversion angles was allowed because the effect of this criterion on the post-operative risk of dislocation depends also on the anteversion of the femoral component, which was not measured.[18] Other centres or surgeons may choose different criteria, so as to accommodate for uncemented femoral components. Errors in interpretation associated with radiological measurements, missing radiographs and inadequate operation notes may affect the proportion of inadequate hips reported. Significant efforts have been made to minimise the number of patients discharged without having radiographs (1.5%).

Although the overall failure was above the targeted performance rate, the CUSUM test did not emit an alarm. There are two possible explanations: the first is that the true performance was worse than the target and that this

was a type II error (the null hypothesis is not rejected when it is false); the second is that the true performance was equal to or better than the target, and the performance observed reflects some natural variation in the process. A second limitation to the study is that the CUSUM test was not adjusted for case mix[38] because of the inability to attribute weights to difficult cases without having a logistic regression model available. Lastly, the CUSUM test is sensitive to the height of the limit h, to the criteria chosen; a sensitivity analysis, conducted but not reported, demonstrated that an alarm would have been raised if the range allowed for criteria 3, 4, 6 and 7 was narrowed by 15%, and to the adequate performance targeted. It is therefore paramount that this 'fine-tuning' of the test is performed before the study is started.

In conclusion, monitoring the quality of implantation of THRs is useful in maintaining and improving surgical performance, as it allows the detection of technical issues in due time. Those who wish to monitor and define their performance accurately on a continuous basis should consider using such statistical methods.

Further opinion

A further opinion by Dr J. Kuiper is available with the electronic version of this article on our website at www.jbjs.org.uk/education/further-opinions

Supplementary material

Appendices and tables describing in detail the CUSUM test, TDR, FDR, and h, as well as a rationale for the choice of the ten performance criteria, are available with the online version of this article on our website at www.jbjs.org.uk

Listen live

Listen to the abstract of this article at www.jbjs.org.uk/interactive/audio

References

1. **Learmonth ID, Young C, Rorabeck C.** The operation of the century: total hip replacement. *Lancet* 2007;370:1508–1519.

2. **Bozic KJ, Kurtz SM, Lau E, et al.** The epidemiology of revision total hip arthroplasty in the United States. *J Bone Joint Surg [Am]* 2009;91-A:128–133.

3. **Barrack RL, Mulroy RD Jr, Harris WH.** Improved cementing techniques and femoral component loosening in young patients with hip arthroplasty: a 12-year radiographic review. *J Bone Joint Surg [Br]* 1992;74-B:385–389.

4. **Chambers IR, Fender D, McCaskie AW, Reeves BC, Gregg PJ.** Radiological features predictive of aseptic loosening in cemented Charnley femoral stems. *J Bone Joint Surg [Br]* 2001;83-B:838–842.

5. **Hirakawa K, Mitsugi N, Koshino T, et al.** Effect of acetabular cup position and orientation in cemented total hip arthroplasty. *Clin Orthop* 2001;388:135–142.

6. **Kennedy JG, Rogers WB, Soffe KE, et al.** Effect of acetabular component orientation on recurrent dislocation, pelvic osteolysis, polyethylene wear, and component migration. *J Arthroplasty* 1998;13:530–534.

7. **Maloney WJ, Keeney JA.** Leg length discrepancy after total hip arthroplasty. *J Arthroplasty* 2004;19(Suppl 1):108–110.

8. **The Bristol Royal Infirmary Inquiry.** Learning from Bristol: the report of the public inquiry into children's heart surgery at the Bristol Royal Infirmary 1984-1995. http://www.bristol-inquiry.org.uk/ (date last accessed 19 May 2011).

9. **The Shipman Inquiry.** Shipman: the final report. 2005. http://www.shipman-inquiry.org.uk/finalreport.asp (date last accessed 19 May 2011).

10. **Spiegelhalter D, Grigg O, Kinsman R, Treasure T.** Risk-adjusted sequential probability ratio tests: applications to Bristol, Shipman and adult cardiac surgery. *Int J Qual Health Care* 2003;15:7–13.

11. **Nizard RS, Porcher R, Ravaud P, et al.** Use of the Cusum technique for evaluation of a CT-based navigation system for total knee replacement. *Clin Orthop* 2004;425:180–188.

12. **Hardoon SL, Lewsey JD, Gregg PJ, Reeves BC, van der Meulen JH.** Continuous monitoring of the performance of hip prostheses. *J Bone Joint Surg [Br]* 2006;88-B:716–720.

13. **Biau DJ, Landreau P, Graveleau N.** Monitoring surgical performance: an application of industrial quality process control to anterior cruciate ligament reconstruction. *Knee Surg Sports Traumatol Arthrosc* 2010;18:1263–1268.

14. **Biau DJ, Milet A, Thévenin F, Anract P, Porcher R.** Monitoring surgical performance: an application to total hip replacement. *J Eval Clin Pract* 2009;15:420–424.

15. **Nezry N, Jeanrot C, Vinh TS, et al.** Partial anterior trochanteric osteotomy in total hip arthroplasty: surgical technique and preliminary results of 127 cases. *J Arthroplasty* 2003;18:333–337.

16. **Konyves A, Bannister GC.** The importance of leg length discrepancy after total hip arthroplasty. *J Bone Joint Surg [Br]* 2005;87-B:155–157.

17. **Ranawat CS, Rodriguez JA.** Functional leg-length inequality following total hip arthroplasty. *J Arthroplasty* 1997;12:359–364.

18. **Lewinnek GE, Lewis JL, Tarr R, Compere CL, Zimmerman JR.** Dislocations after total hip-replacement arthroplasties. *J Bone Joint Surg [Am]* 1978;60-A:217–220.

19. **Orthopaedic Institute at Oswestry.** Roman V.1.70 software: technical report. http://www.orthopaedic-institute.org.

20. **Widmer KH.** A simplified method to determine acetabular cup anteversion from plain radiographs. *J Arthroplasty* 2004;19:387–390.

21. **Langlais F, Kerboull M, Sedel L, Ling RS.** The 'French paradox'. *J Bone Joint Surg [Br]* 2003;85-B:17–20.

22. **Page E.** Continuous inspection schemes. *Biometrika* 1954;4:100–115.

23. **Kohn LT, Corrigan JM, Donaldson MS, eds.** *To err is human: building a safer health system.* Washington: National Academy Press, 2000.

24. **de Leval MR, François K, Bull C, Brawn W, Spiegelhalter D.** Analysis of a cluster of surgical failures: application to a series of neonatal arterial switch operations. *J Thorac Cardiovasc Surg* 1994;107:914–924.

25. **Biau DJ, Halm JA, Ahmadieh H, et al.** Provider and center effect in multicenter randomized controlled trials of surgical specialties: an analysis on patient-level data. *Ann Surg* 2008;247:892–898.

26. **Bosker BH, Verheyen CC, Horstmann WG, Tulp NJ.** Poor accuracy of freehand cup positioning during total hip arthroplasty. *Arch Orthop Trauma Surg* 2007;127:375–379.

27. **García-Cimbrelo E, Diez-Vazquez V, Madero R, Munuera L.** Progression of radiolucent lines adjacent to the acetabular component and factors influencing migration after Charnley low-friction total hip arthroplasty. *J Bone Joint Surg [Am]* 1997;79-A:1373–1380.

28. **Gruen TA, Poggie RA, Lewallen DG, et al.** Radiographic evaluation of a monoblock acetabular component: a multicenter study with 2- to 5-year results. *J Arthroplasty* 2005;20:369–378.

29. **Kobayashi S, Takaoka K, Saito N, Hisa K.** Factors affecting aseptic failure of fixation after primary Charnley total hip arthroplasty: multivariate survival analysis. *J Bone Joint Surg [Am]* 1997;79-A:1618–1627.

30. **Kobayashi S, Eftekhar NS, Terayama K, Iorio R.** Risk factors affecting radiological failure of the socket in primary Charnley low friction arthroplasty: a 10- to 20-year followup study. *Clin Orthop* 1994;306:84–96.

31. **White TO, Dougall TW.** Arthroplasty of the hip: leg length is not important. *J Bone Joint Surg [Br]* 2002;84-B:335–338.

32. **Macdonald W, Carlsson LV, Charnley GJ, Jacobsson CM, Johansson CB.** Inaccuracy of acetabular reaming under surgical conditions. *J Arthroplasty* 1999;14:730–737.

33. **MacKenzie JR, Callaghan JJ, Pedersen DR, Brown TD.** Areas of contact and extent of gaps with implantation of oversized acetabular components in total hip arthroplasty. *Clin Orthop* 1994;298:127–136.

34. **Thiemann DR, Coresh J, Powe NR.** Quality of care at teaching and nonteaching hospitals. *JAMA* 2000;284:2994–2995.

35. **Fukunishi S, Fukui T, Imamura F, Nishio S.** Assessment of accuracy of acetabular cup orientation in CT-free navigated total hip arthroplasty. *Orthopedics* 2008;31:987.

36. **Leenders T, Vandevelde D, Mahieu G, Nuyts R.** Reduction in variability of acetabular cup abduction using computer assisted surgery: a prospective and randomized study. *Comput Aided Surg* 2002;7:99–106.

37. **Manzotti A, Cerveri P, De Momi E, Pullen C, Confalonieri N.** Does computer-assisted surgery benefit leg length restoration in total hip replacement?: navigation versus conventional freehand. *Int Orthop* 2011;35:19–24.

38. **Steiner SH, Cook RJ, Farewell VT, Treasure T.** Monitoring surgical performance using risk-adjusted cumulative sum charts. *Biostatistics* 2000;1:441–452.