Biomechanical, morphological, and histological analysis of early failures in hip resurfacing arthroplasty

M M Morlock[1]*, **N Bishop**[1], **W Rüther**[2], **G Delling**[3], and **M Hahn**[4]

[1] Biomechanics Section, Hamburg University of Technology, Hamburg, Germany
[2] Institute of Orthopaedic Surgery, University Hospital Eppendorf, Hamburg, Germany
[3] Institute of Osteopathologie, University Hospital Eppendorf, Hamburg, Germany
[4] Biomechanics Section, University Hospital Eppendorf, Hamburg, Germany

Abstract: The present revival of hip resurfacing arthroplasty may be related to an increase in early failures owing to the challenging technique of the procedure. Fifty-five retrieved implants were analysed with respect to wear, cement mantle and cement penetration, fracture and head morphology, as well as standard histology. Femoral neck fractures occurred in median after 102 days. The time to failure was shorter for older women. Major deviations from the suggested cement mantle thickness and cement penetration were found. Indications for high trauma during implantation leading to early failure due to weakening of the femoral neck were also observed. Some failures had signs of pseudarthrosis beneath the implant. Four different fracture patterns with different mean survival times were identified. Observed wear was minor with the exception of that due to alignment mistakes (rim loading). The cups were not damaged by the failures. Histological results indicate that avascular necrosis is not necessarily connected with this kind of endoprosthetic surgery. Most of the failures analysed can probably be attributed to the 'learning curve' effect, which is an unsatisfactory situation.

Keywords: wear, deformation, necrosis, failure mechanism, metal-on-metal, surface replacement, learning curve, cementing technique

1   INTRODUCTION

Hip resurfacing arthroplasty is presently experiencing an unexpected revival. After the bad experiences with Wagner resurfacing arthroplasty, which were attributed mostly to bearing wear, aseptic osteolysis [1, 2], and implant design [3], a new approach was taken using newly designed metal-on-metal bearings. The first results with follow-up periods of 7–10 years were promising [4, 5] and, consequently, the number of companies and surgeons interested in hip resurfacing arthroplasty increased. In 1997 three different implant types were established on the market. Today more than nine different implants are available and the number of hip resurfacing procedures is continuously increasing, from some 36 000 in 2005 to 45 000 in 2006 (industry estimates).

Presently, no long-term results for the second- and third-generation hybrid resurfacing implants are available. Some studies indicate survival rates of over 97 per cent with follow-up times ranging from 2 to 8 years [5–7]. Fracture of the femoral neck is reported to be one of the most common reasons for revision [8, 9]. The fractures occur within the first 3–4 months after surgery and are attributed to different factors: uncovered bone, leaving the component proud, notching the neck, osteopenia and cysts, impingement, and trauma [10]. One factor not mentioned in this list is avascular necrosis, which is deemed quite important by other authors [9]. All these results are from controlled studies in hospitals with substantial experience in the area of hip resurfacing. At present, this technique, however, is applied by many less-experienced surgeons for whom hip resurfacing is novel. Since hip resurfacing is a challenging procedure, a high initial failure rate has to be expected. These failures can be attributed either to patient selection or problems with surgical technique.

Corresponding author: Biomechanics Section, TU Hamburg-Harburg, Denickestrasse 15, D-21073 Hamburg, Germany. email: morlock@tuhh.de

Problems regarding patient selection can best be assessed in controlled studies. Problems regarding surgical technique, however, should be investigated by looking at cases from as many different surgeons as possible. This is particularly valid for a surgical procedure which in its classical format is highly successful and presently shows survival rates of 94 per cent after 10 years [11].

As a design inherent limitation, X-rays do not give any information about the situation beneath the femoral component in hip resurfacing arthroplasty and, as such, the surgeon can only guess what the reason for failure could be. In order to keep the 'learning curve' to a minimum, early and direct feedback to the surgeon for each failure case is required. The purpose of this study is the analysis of failed hip resurfacing arthroplasties from a random sample of institutions. The goal is not to establish survival percentages based on a controlled clinical study but to analyse the failed cases with regard to possible failure mechanisms. Failure rates determined in controlled clinical studies represent the best case scenario since, typically, a small number of surgeons, careful patient selection, and large procedure numbers are involved. Failure rates outside such studies are expected to be higher. Currently, it is speculated that early failures are not heavily influenced by implant design but mostly by surgical technique and patient selection. Consequently, implants from different manufacturers are included in the study but the specific implant type is not explicitly mentioned as it is expected to be of minor importance. Late failures, which have not yet been analysed, may show differences between designs. This will have to be investigated in the future.

2 METHODS

The study was made known to the surgeons through the companies involved as well as through scientific presentations at conferences. Surgeons were asked to send retrieved femoral heads together with revised acetabular cups wrapped in gauze immersed in 4 per cent Formalin solution to Hamburg. Patient demographics, medical history, pre- and post-op X-rays as well as information on cementing technique were also requested.

2.1 Wear estimation

The surface geometry of the retrieved bearings was assessed with a coordinate measurement machine (Mitutoyo BHN 805, Fig. 1). Sixteen planar scans were
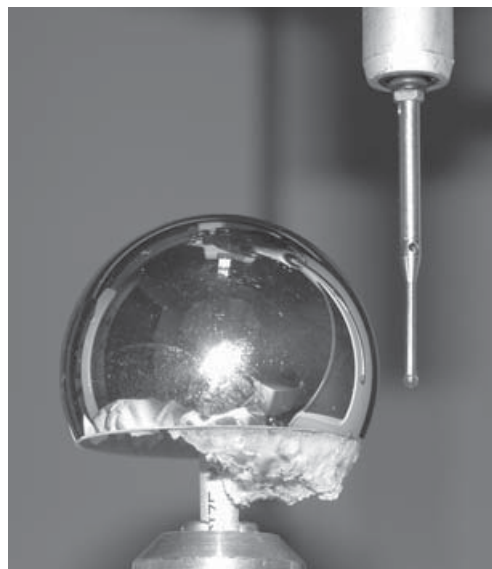

Fig. 1 Digitization of the retrieved femoral head using a coordinate measurement machine

performed through the pole at $11.25°$ intervals with a spacing of 0.5 mm between points using a 1 mm diameter ruby head starting and ending at the equator. To determine wear, the shape and size of the original bearing have to be estimated. As the manufacturing tolerances on bearing sphericity can be assumed to be sufficiently low, a perfect sphere can be fit to the measured data [12, 13]. This scheme, however, results in errors when all points in the surface measurement of the retrieved bearings are considered because the worn areas bias the size and position of the best-fit sphere away from regions of the original unworn surface (Fig. 2(a)). Therefore, points on the worn regions should be eliminated from the best-fit calculation (Fig. 2(b)). Another factor which could introduce similar errors is overall deformation of the bearing, for example, due to press-fit (Fig. 3(a)). This may be particularly relevant to uncemented components, which often exhibit an ellipsoidal deformation, particularly at the open (equatorial) perimeter.

Consequently, an ellipsoidal surface was used for fitting rather than a sphere (Fig. 3(b)), in which the optimization variables were the coordinates of the ellipsoidal centre (c_x, c_y, c_z) and its orthogonal radii (r_x, r_y, r_z) in the coordinate system of the measurement machine. Rotation of the ellipsoid in the plane of the open end was also incorporated as another variable. Thus, seven degrees of freedom were varied using a Matlab function ('fminunc') to minimize the sum of squared radial residuals (distances) between the measured points (x_i, y_i, z_i) and the ellipsoidal
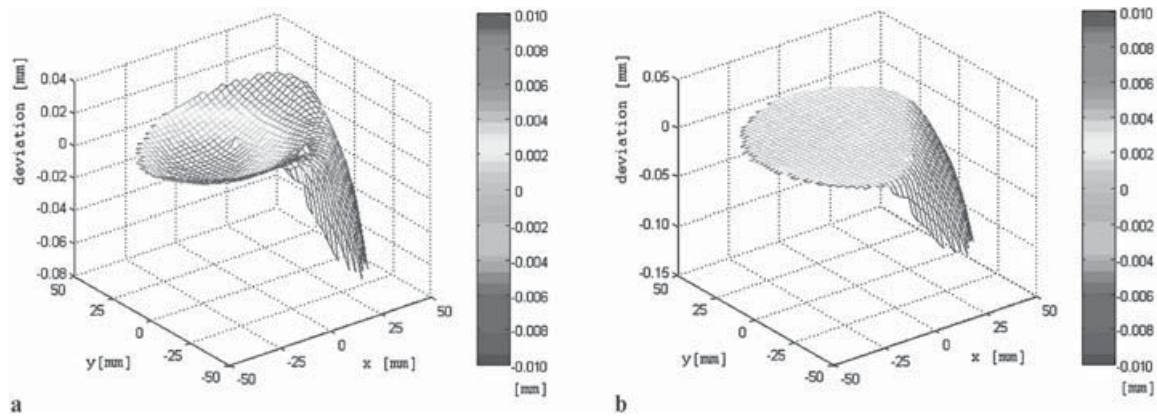
Fig. 2 Deviations of a worn cup surface from a best fit surface. (a) All points considered: the curved shape of the plot indicates that the centre of the best fit surface is not at the centre of the original cup, but biased by the worn region. This leads to erroneous estimation of wear (estimated depth = 81 μm, volume = 9.15 mm³). (b) Iterative removal of 20 per cent of the cup surface points with the highest residual allow the actual wear to be better estimated (estimated depth = 119 μm, volume = 16.97 mm³)
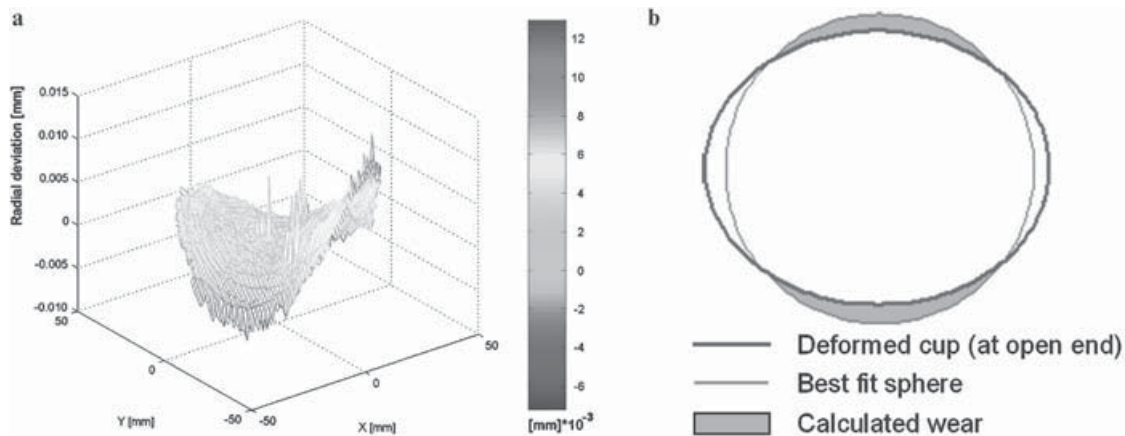
Fig. 3 (a) Radial deviations of a deformed head from a best fit sphere. The saddle shape indicates a non-circular deformation which is greatest at the open end of the cup. (b) Schematic showing how elliptical deformation of the head (e.g. due to press fit) could lead to a false estimation of the wear magnitude using a spherical best fit: view of the open end

surface defined by

$$\left(\frac{x_i - c_x}{r_x}\right)^2 + \left(\frac{y_i - c_y}{r_y}\right)^2 + \left(\frac{z_i - c_z}{r_z}\right)^2 = 1$$

The proportion of points to be used was varied parametrically from 100 to 60 per cent. The points to be used in the routine were then determined by eliminating the given fraction of points with the greatest radial distance from the best-fit surface, which was revised for each iteration of the optimization procedure.

Figure 4 shows the effect of varying the proportion of points used to determine the best-fit ellipsoid for the example of a resurfacing head with flattened pole and a flattened side (achieved by grinding). The

central negative peak represents the polar flat and the peripheral negative peak represents the side flat. With all points considered (0 per cent removed), only very few of the points lie on the surface of the best-fit ellipsoid. Removing the largest 10 per cent of residuals, the majority of points move towards the zero residual plane, as the points on the negative peaks are no longer used for determination of the ellipsoid. Omitting 20 per cent or more of the points for the determination of the best-fit ellipsoid results in a stable solution, with all points apart from those corresponding to the flats showing a very low residual (Fig. 4(d)).

The surface of the best-fit ellipsoid was triangulated between measured points using a Delaunay
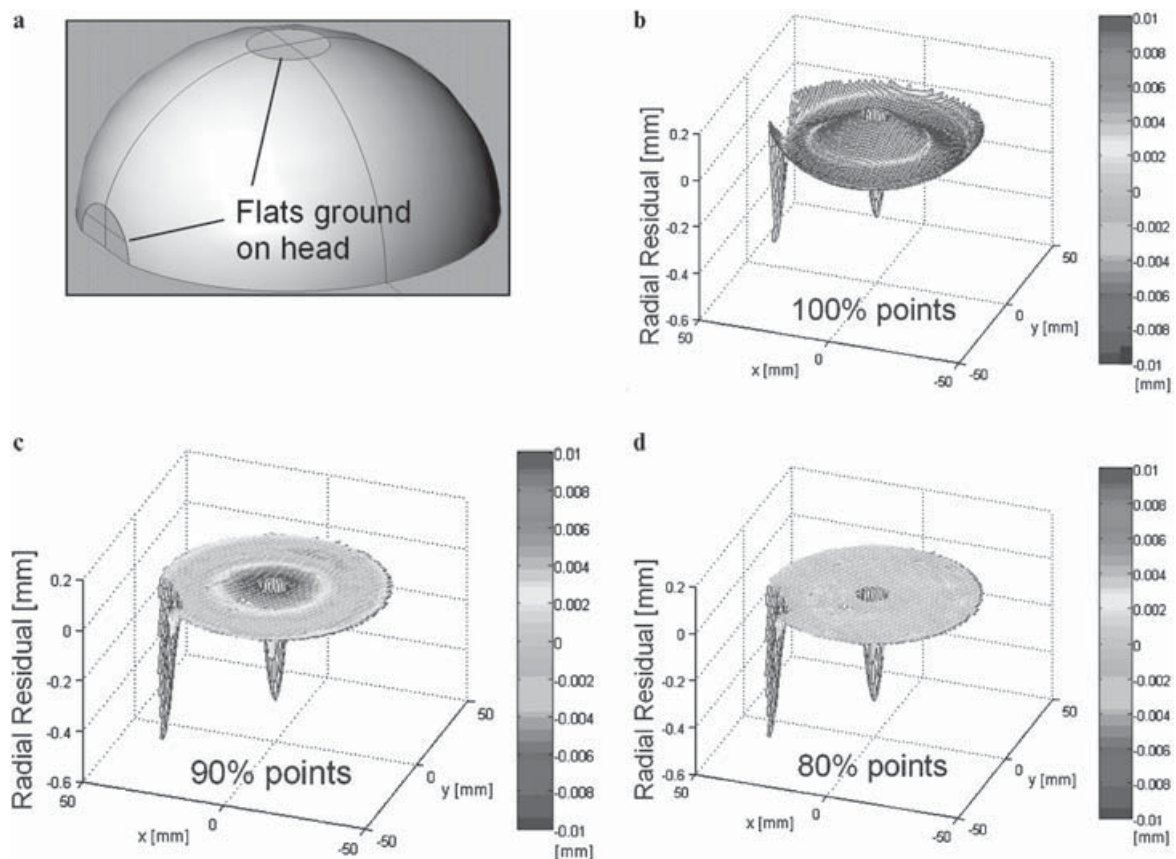
Fig. 4 A new (spherical) resurfacing head with two ground flats to simulate wear (a), was used to estimate the effect of elimination of certain proportions of measured points with the largest residuals from the best fit calculation. (b) shows that the best fit surface using all points lies at some distance from the actual surface. This improves when 10 per cent of the most distant points is neglected (c), and a good fit to the original surface is found with 20 per cent of the largest residuals neglected (d)

routine and the volume, area, and distance between the centroid of each triangle and its radial projection onto the best-fit surface were calculated. Area and volume wear magnitudes were summed up for all triangles with centroidal distances from the best-fit surface greater than 3 μm (the calibrated accuracy of the measurement machine) and the largest distance was also recorded.

This method was applied to the resurfacing head with flattened pole and side described above in an undeformed and a deformed state (press-fitted onto a conically reamed femoral head). Spherical and ellipsoidal best fits were compared. Consideration of all points resulted in an error in volume estimation (not shown) as well as a magnitude of deviation of the surface area (Fig. 5). Steady state volume, area, and depth magnitudes resulted when 20 per cent or more of the points with peak residuals were not considered, particularly for the undeformed head, with either ellipsoidal or spherical fits. For the deformed state the wear area was estimated much better by

the ellipsoidal fit than by the spherical fit (Fig. 5). Based on these results the best-fit ellipsoid approach was used in this study and the parameters wear area (area with more than 3 μm deviation from the best-fit ellipsoid), wear volume (integrated over the wear area), wear depth (largest centroidal distance), area (total digitized area), and area ratio (wear area divided by total area) were calculated.

The scheme was also tested on simulated point data sets, representing a perfect sphere, with varying regions of simulated wear of 10 μm depth. Polar, equatorial, and segmental regions of 'wear', of 10 and 20 per cent of the total hemispherical area, were found to result in simulated area and volume measurement errors of less than 8 per cent.

2.2 Morphological methods

After the wear analysis a central 4 mm thick slice was cut from the heads in the femoral neck plane using a diamond saw (EXAKT 310). Orientation of the
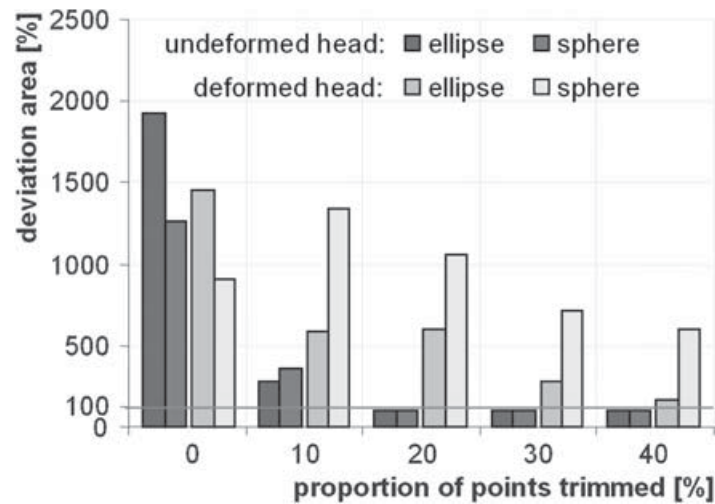
Fig. 5 Estimation of the 'wear' area using a best fit sphere and best fit ellipsoid for an undeformed and a deformed (press-fit) resurfacing head flattened at the pole and on one side (Fig. 2). The estimated true area of the flats is represented by 100 per cent

specimens prior to cutting was difficult as the anatomical orientation was not known. Consequently, errors with regard to the cutting plane were to be expected. The slice was X-rayed, photographed, then embedded in Technovit 7200 (Kulzer) and ground down to a thickness of 1 mm (Fig. 6). The techniques are described in detail elsewhere [**14**].

The cement mantle was assessed visually using a projected polar grid at 22.5° intervals, similar to the method used by Howie [**15**] but with a higher reso-

lution (Fig. 7). The centre of the grid was placed at the centre of the head. Parameters determined were the cement mantle thickness and the depth of cement penetration into the bone. A combination of both parameters yielded the cement thickness. The cement thickness, according to manufacturer's instructions, should be approximately 2–3 mm for the mantle and about the same amount for penetration. Mantle thickness above 5 mm is referred to as 'excessive cement', cement penetration above 5 mm
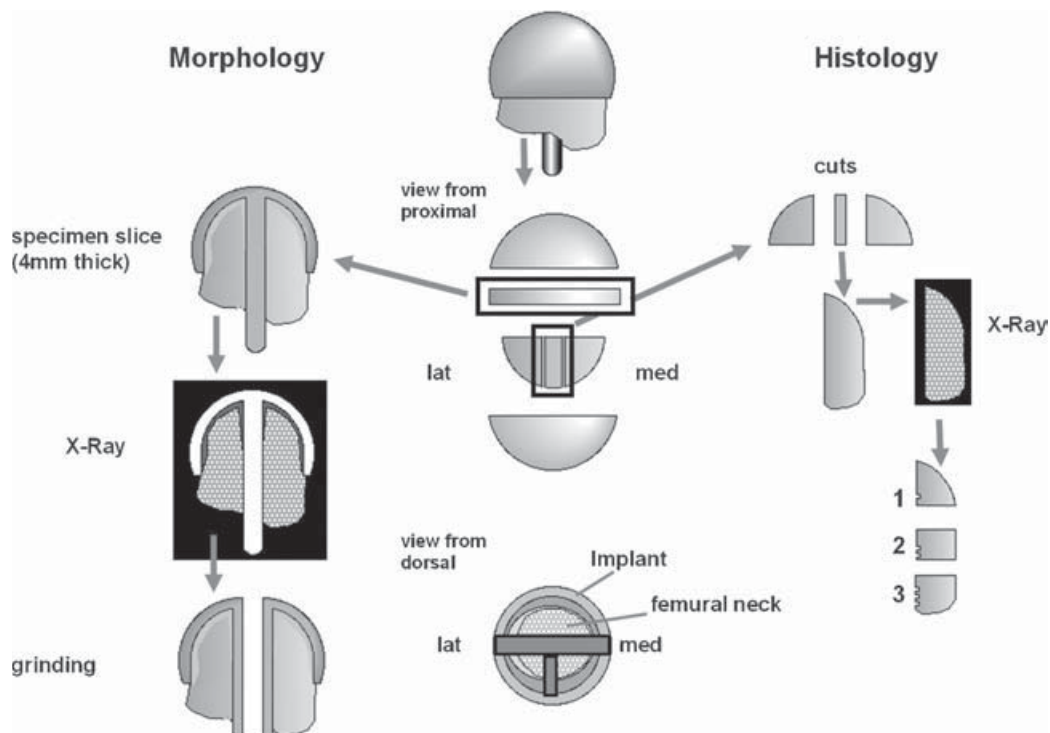
Fig. 6 Schematics of the preparation for the morphological and histological analysis of the head
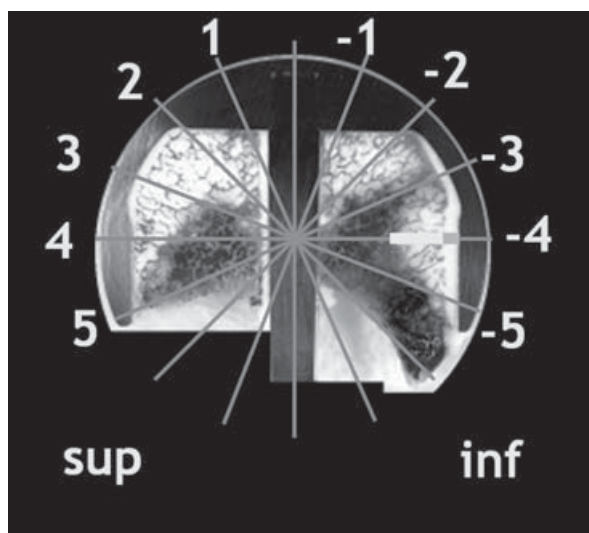
Fig. 7 Sections of the head for the determination of cement mantle thickness (indicated in darker grey) and cement penetration depth (indicated in lighter grey) from the ground central slice

is referred to as 'excessive penetration'. These values were chosen arbitrarily as no scientific evidence regarding either parameter was available.

The morphology of the fracture line was assessed visually based on the central slice and was categorized into five groups:

1. fracture line completely within the implant ('inside head');
2. fracture line between both edges of the implant ('edge to edge');
3. fracture line from the edge of the implant to the outside ('edge to out');
4. fracture completely outside the implant ('outside');
5. femoral neck cut for revision ('no fracture').

2.3 Histological methods

The remaining anterior half of the head was used for the histological analysis. The metallic implant was removed using acetone (Fig. 6). The middle section (4 mm thick) was used for the cutting sections (slice thickness 5 μm) using a Microtom. Toluidine blue-staining for the ultra-thin (10–20 μm) and surface-stained block grindings (1 mm) was performed. The cut sections were stained with von Kossa, Goldner, and Toluidine blue.

2.4 Material

Within 13 months of the study duration, 55 implants (44 heads alone, five heads with cups, one cup alone) were received. The information supplied with the retrievals was incomplete in all but 19 cases. The number of specimens analysed varies for each of the analyses performed and is, therefore, specified as 'n_{tot}'.

2.5 Statistical analysis

The data were analysed using 'SPSS 12.1 for Windows'. Nominal variables were analysed using Chi^2-tests. Continuous variables were analysed using one-way analysis of variance or linear regression analysis. Due to the incomplete data sets, pairwise exclusion of missing data was selected. Type II error probability was set to 5 per cent (alpha$=0.05$) for all tests.

3 RESULTS

3.1 Patient and medical history

Indications for the performance of the surface replacement were primary coxarthrosis (in 69 per cent of the cases), dysplastic hip (14 per cent), post-traumatic arthritis (10 per cent), and rheumatoid arthritis (7 per cent) ($n_{tot}=29$). The average patient age was 56 years [standard deviation (STD) 8 years; $n_{tot}=31$] with slightly more males than females (56 per cent, $n_{tot}=32$). The average patient height and weight were 171 cm (STD 5.9 cm, $n_{tot}=25$) and 79.7 kg (STD 13.8 kg, $n_{tot}=27$) respectively. Revision occurred after a median of 102 days (mean 150 days, STD 151 days, $n_{tot}=33$). The time to revision increased significantly with the number of prior procedures performed by the respective surgeon ($r=0.42$, $p=0.039$, $n_{tot}=31$). The average number of prior procedures reported for the respective surgeon ranged from 0 to 460 with a median of 200. The time to revision decreased significantly with patient age, but only for the female patients ($r=-0.63$, $p=0.02$, $n_{tot}=14$). The time to revision for females older than 58 years was only 24 days (± 10 days), which was significantly less than for females younger than 58 years (236 days ± 204 days; $p=0.04$). For the male patients, age had no influence ($r=0.15$, $p=0.58$, $n_{tot}=17$). Mean age and age ranges were similar for female and male patients (male: mean 56.4 ± 7.6 years, minimum 38, maximum 74; female: mean 55.4 ± 8.7 years, minimum 45, maximum 72; $p=0.74$). Since the overall patient age distribution is unknown, these results should be treated with great caution.

3.2 Surface wear

The digitized wear area on the head increased with the height and weight of the patient, this being a

Proc. IMechE Vol. 220 Part H: J. Engineering in Medicine

JEIM70 © IMechE 2006

Table 1 Wear parameters for the retrieved heads dependent on the time *in situ*. The average time *in situ* for the three groups is specified

Variable	Survival (days)	n	Mean	STD	Minimum	Maximum
Time *in situ*	<100	16	45.00	26.133	14	98
(days)	100–200	9	139.33	38.607	102	196
	>200	8	370.50	145.680	211	602
Max wear	<100	15	5.67	2.92	0	13.00
(μm)	100–200	8	10.50	6.63	4.00	21.00
	>200	8	13.50	22.50	3.00	69.00
Wear area	<100	11	37.00	44.54	0	102.56
(mm^2)	100–200	8	149.74	174.08	0	401.65
	>200	7	122.17	206.76	0	572.05
Wear volume	<100	11	0.17	0.21	0	0.47
(mm^3)	100–200	8	1.02	1.32	0	3.61
	>200	7	2.76	6.71	0	17.96

clear indication of the use of larger head sizes ($r = 0.73$, $p < 0.001$, $n_{tot} = 19$; $r = 0.57$, $p = 0.007$, $n_{tot} = 21$). The maximum wear depth, wear area, and wear volume increased significantly with time *in situ* ($r = 0.533$, $p = 0.005$; $r = 0.395$, $p = 0.046$; $r = 0.548$, $p = 0.004$; $n_{tot} = 26$; Table 1). There was a significant correlation between time *in situ* and wear volume with a mean rate of 0.012 mm^3/day ($r^2 = 0.30$, $p = 0.004$).

All but one cup showed no major wear ($n = 5$, depth $= 7 \pm 5$ μm, volume $= 0.25 \pm 0.38$ mm^3). The major wear in the one cup was due to rim loading (Fig. 2). Scratches and minor wear marks were found on most of the heads; those cannot be linked to the 'normal' wear process but might be as a result of the revision process or the failure event.

3.3 Morphology

The cement thickness was highest around the pole and lowest on the sides, with large variations between specimens (Fig. 8). Investigation of the cement underneath the head yielded surprisingly poor results ($n_{tot} = 31$): 31 per cent of the heads indicated cementing that had been performed according to suggested techniques (Fig. 9(e)), 6 per cent exhibited cement mantles at the pole exceeding 5 mm (Fig. 9(a)), 37 per cent showed cement penetrations exceeding 5 mm (Fig. 9(d)) and 20 per cent showed excessive penetration and mantle thickness (Fig. 9(c)). Only 6 per cent showed too little cement. Information regarding cementing procedure showed very different cementing techniques ($n_{tot} = 29$): jet lavage was used in 90 per cent, suction at the minor trochanter in 59 per cent and vacuum mixing in 21 per cent of the cases. Ten different cement types (high-viscosity and low-viscosity) with and without antibiotics and/or contrast medium were reported.

Cement curing and application procedures were of an even wider variety.

The five fracture types are illustrated in Fig. 9. Interestingly, significant differences and tendencies in the survival time between the different fracture-line locations were found (Table 2). The fractures completely outside the head occurred the earliest, the revisions without an acute fracture the latest ($p = 0.014$). Large within-group variations were found. There was also a tendency for differences in wear with the different fracture-line locations (Table 3). Those were, however, not significant. There was no correlation between fracture-line location and cement status ($p = 0.352$). From X-ray, three

Table 2 Time to failure for the different fracture line morphologies in ascending order

	Time to failure (days)				
Fracture site	n	Mean	STD	Minimum	Maximum
Outside	4	53	49	21	126
Edge to edge	3	114	106	15	225
Edge to out	10	117	97	17	332
Inside head	10	137	128	14	405
No fracture	5	355	226	98	602
Total	32	152	153	14	602

Table 3 Wear volume for the different fracture line morphologies in ascending order

	Wear volume (mm^3)				
Fracture site	n	Mean	STD	Minimum	Maximum
Inside head	9	0.40	0.70	0	2.10
Outside	3	0.67	0.84	0	1.62
Edge to edge	5	1.52	3.34	0	7.49
Edge to out	12	2.59	7.14	0	25.05
No fracture	11	5.27	7.81	0.13	22.50
Total	40	2.56	5.90	0	25.05
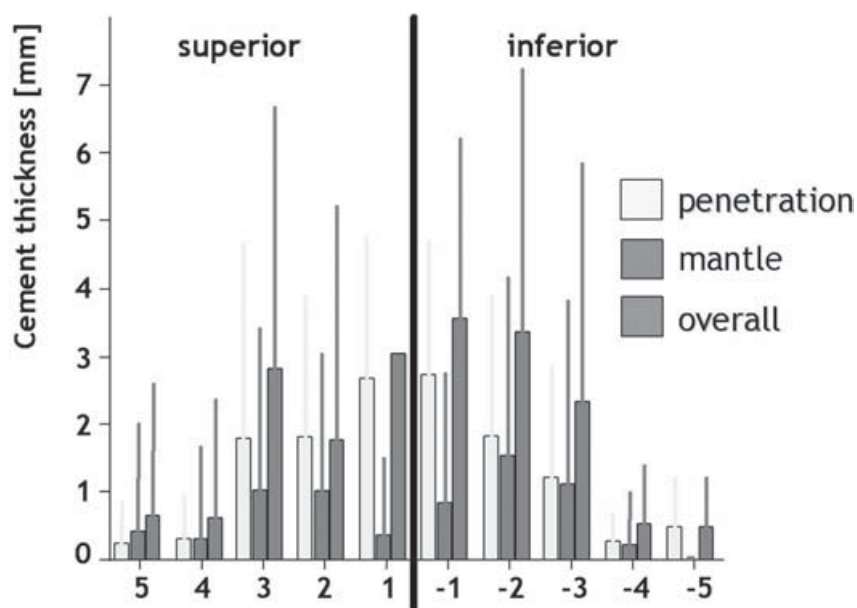
Fig. 8 Cement mantle thickness, cement penetration, and overall cement thickness underneath the head ($n = 21$). The sections are defined in Fig. 7

notchings of the superior cortex as well as three clear malpositions (two in varus, one in valgus) were identified from the 19 sets of X-rays available.

3.4 Histology

The histological results have to be treated carefully as, in many cases, insufficient information regarding patient history, failure history, and storage retreatment of the retrieved head after failure was available. The analysis of 28 cases was completed. In 32 per cent of the cases, no major abnormalities were found, in 29 per cent there were signs of avascular necrosis and in 39 per cent, signs of a 'two-instant' fracture pattern. In this fracture pattern, osteoblastic activity within the head was found in damage zones as well as signs of pseudoarthrosis (Figs 10, 11). Two cases showed signs of osteomyelitis.

4 DISCUSSION

This study presents a novel approach to the analysis of clinical failures. In contrast to controlled clinical studies, not all relevant information with regard to the failure phenomenon is available. However, it is anticipated that the analysis of a higher number of failures in future will yield a more realistic cross-section of the present situation in hip resurfacing and give a better estimate of the situation for an arbitrary patient.

The fractures reported in the literature are frequently seen to be related to a traumatic episode [16]. This is not surprising since the failure of the femoral neck owing to its biomechanical loading situation and the high forces and moments acting at the hip will, in most cases, be a traumatic single overload event rather than a fatigue failure. Furthermore, even if the fracture may have been initiated at an earlier point in time, ultimate failure will always be related to such a defined episode.

The patient population of the failures investigated seems to be in the appropriate range of indications and demographics. Revisions occur earlier in older women, which is not unexpected and might be explained by their decreasing bone quality. The number of prior surgeries performed increased the time to revision.

The algorithm used in this study for the determination of wear accounts for errors which may be introduced by least-square fitting of a sphere, which is the most common method cited [12, 13, 15]. Wear of a deformed implant was addressed by use of a non-spherical surface fit to a proportion of the measured points by removing points with the largest deviation (areas of wear) from the fitted ellipsoid. The use of 80 per cent points or less gave stable wear measurements and a good fit to the unworn surface.

The observed wear magnitudes (Table 1) were minor, which should not be a surprise as the time *in situ* for all the retrievals was short. It is noteworthy that a clear increase in wear can be observed with increasing time *in situ*, even though it is at a very low level (Table 2). The value of 0.012 mm³/day determined in this study is very close to values in
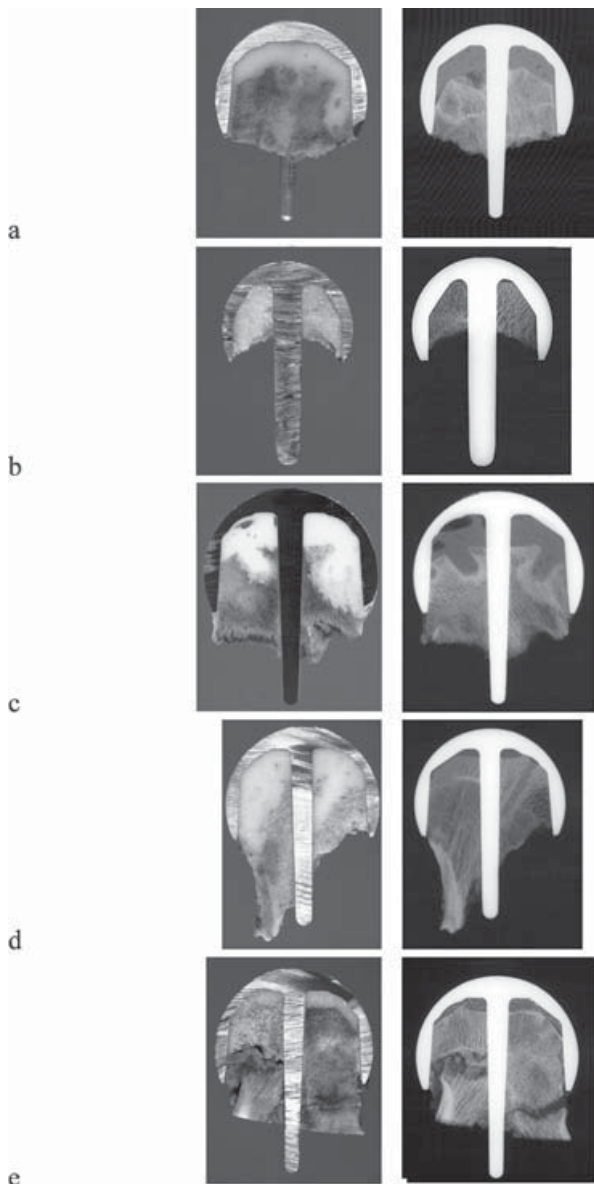
Fig. 9 The five different failure patterns observed: (a) fracture line between both edges of the implant ('edge to edge'); (b) fracture line completely underneath implant ('inside head'); (c) fracture completely outside the implant ('outside'); (d) fracture line from the edge of the implant to the outside ('edge to out'); (e) femoral neck cut for revision ('no fracture')

the literature of successful metal-on-metal McKee–Farrar, Müller, Huggler, and Ring endoprostheses (0.016 mm³/day [17], 0.014 mm³/day [12]). Linear wear rates in this study (10.5 μm for 100–200 days) are also similar to the mean values measured for the first year in a recent retrieval study of second-generation metal–metal heads (27.8 μm for 0–365 days) [18]. As the value determined in this study is based on short-term retrievals during their bedding-in phase, it can be speculated that the wear rate of

these new generation implants will decrease in the long term, as observed for retrieved second-generation bearings [18]. In early failure cases of hip surface replacements, wear cannot be involved in the causal relation with failure but, rather, can serve as information for validating the results of simulator studies. Major wear was only found in one situation with a rim-loaded head and cup, which is in accordance with the literature reporting more metal wear with high cup inclinations [19].

It is interesting to note that wear was smallest for the fractures inside the head, even though they did not fail earliest (Table 3). This suggests that those patients had pain from the beginning and did not put much load on the joint.

The cement situation found underneath the femoral component of the resurfacing implant is unsatisfactory. Of the analysed heads, 63 per cent showed major deviations from the desired situation. This can only be explained by the fact that the surgeon has no way of controlling the achieved result after the head has been put on. The result also suggests that the use of high viscosity cement (as in knee arthroplasty) may be favourable as, with viscous cement, the technique is much easier to control. The cement can be finger-packed onto the reamed head and pressed manually into the trabecular bone. Excessive cement can be removed before putting on the implant. This should eliminate situations shown in Figs 9(a) and (c) (excessive cement at the pole). Most of the surgeons seem to favour the use of jet lavage and suction at the trochanter minor, whereas vacuum mixing of cement does not seem to be widely used. It is not possible yet to speculate on the influence of these methodological aspects, especially as their exact mode of use cannot be determined (for example, duration of jet lavage, suction pressure). From an engineering point of view, quality control of the cementing procedure in any way is very difficult as so many different cement brands, viscosities, and cementing procedures are used.

As the polymerization temperature of the cement increases with its volume, excessive usage of cement should be avoided because the danger of necrosis can be expected to rise with the polymerization temperature. Necrosis should not, however, be a major factor in early hip resurfacing failures as mainly observed in this study. It could be speculated that this may play an important role later on.

One of the most interesting findings of this study is the difference in fracture-line patterns and the difference in the time to failure between these patterns. Weakened areas in the femoral neck distal to the implant (such as uncovered reamed bone or
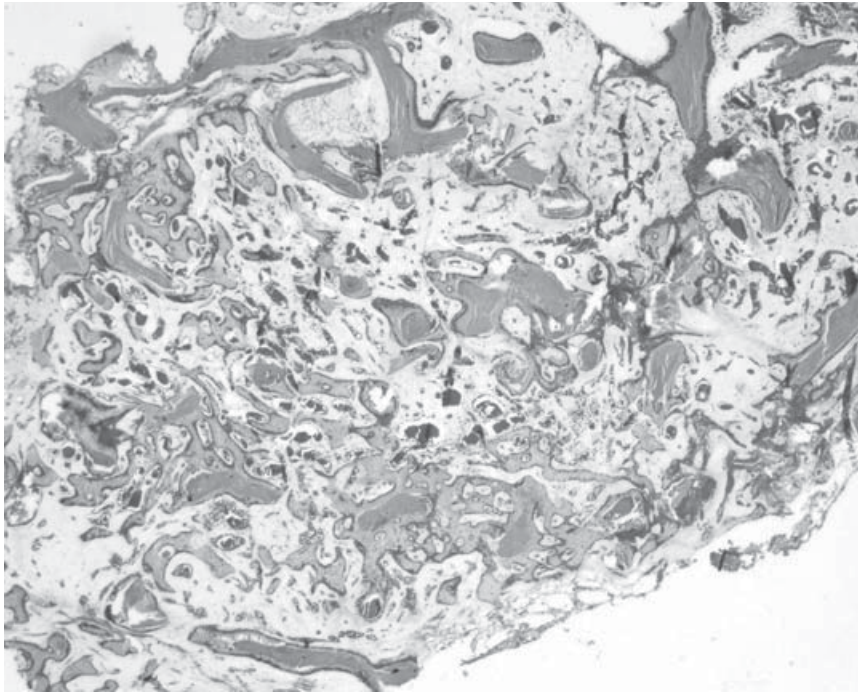
Fig. 10 Newly formed fibrous bone around micro fractures inside the head proximal to the failure line
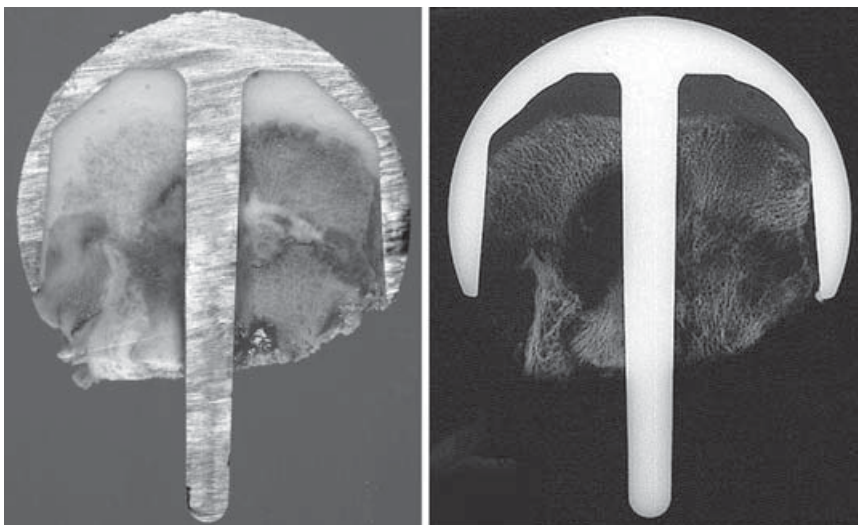
Fig. 11 Signs of pseudarthrosis underneath the head

microfractures caused by implantation trauma) exhibited the shortest survival times. It can be speculated that these fractures happen as soon as a loading episode with forces and moments large enough to exceed the remaining bone strength occur. The fractures involving the implant itself all occurred after some 4 months. It is interesting to note that those fractures which were completely inside the head occurred the latest (not significant, just a slight trend). This may be explained by a mechanical stabilization of the fracture by the implant itself. The high rate of pseudarthrosis or 'two time' fracture patterns indicate that the fractures beneath the head develop from an earlier event. This may also be true for the fractures involving the rim of the implant. Overall, it can be speculated that many of the failures could be attributed to a high implantation trauma. It is well known that a proud implant or uncovered reamed bone are high risk factors for failures [**10**]. If during surgery such a situation is anticipated, it is likely that the surgeon will use excessive force to seat the implant.

Most of the specimens received can be called 'typical learning curve failures' with either rather obvious insufficient cementing, alignment, or implantation technique. In 20 per cent of the failures analysed, those failures occurred during the first 10 surgeries of the respective surgeon. It has, for example, been shown that optimizing the femoral stem-shaft angle towards valgus orientation reduces the failure rate [20]. However, in approximately 30 per cent of the failures analysed in this study, no direct explanation for the failure was found. Those failures may be due to anything from poor patient selection to major trauma. These cases will be the interesting ones to be followed in future failure analysis studies. Learning curve failures are not really acceptable from a patient perspective and it will be a challenging task to develop training programmes, better instrumentation, and additional aids (navigation may be one) in the future.

The inherent limitation of an uncontrolled study is the unknown overall patient population and the incomplete information for many of the retrievals. X-rays were only received in fewer than half the cases, and similar for cementing information. Furthermore, only sections of the head were analysed, providing a sample but no overall information regarding the situation beneath the implant. Anatomical orientation of the specimens is difficult without landmarks provided by the surgeons. It is also unknown, how long the specimens were kept in Formalin and how long it had taken after the failure to immerse the specimens. Owing to the nature of the retrievals (bone is mostly covered by the implant), Formalin penetration is only possible once the head is cut, and varies greatly between specimens. Finally, only those surgeons who are interested in the failure mechanisms and in improvement of their technique provide the specimens for analysis; whether this is a representative sample or not can only be speculated.

Cemented resurfacing acetabular components show high revision rates with 21 per cent at 7 years [4]. The new generation of hybrid bearings has the potential to show much lower rates. These rates could be further improved if early failure could be eliminated. This study has shown that poor cementing technique and high trauma during implantation may be responsible for a large percentage of femoral neck fractures.

REFERENCES

1 **Howie, D. W., Campbell, D., McGee, M.,** and **Cornish, B. L.** Wagner resurfacing hip arthroplasty. The results of one hundred consecutive arthroplasties after eight to ten years. *J. Bone Jt Surg. Am.*, 1990, **72**(5), 708–714.

2 **Bell, R. S., Schatzker, J., Fornasier, V. L.,** and **Goodman, S. B.** A study of implant failure in the Wagner resurfacing arthroplasty. *J. Bone Jt Surg. Am.*, 1985, **67**(8), 1165–1175.

3 **Wiadrowski, T. P, McGee, M., Cornish, B. L.,** and **Howie, D. W.** Peripheral wear of Wagner resurfacing hip arthroplasty acetabular components. *J. Arthroplasty*, 1991, **6**(2), 103–107.

4 **Beaule, P. E., Le Duff, M., Campbell, P., Dorey, F. J., Park, S. H.,** and **Amstutz, H. C.** Metal-on-metal surface arthroplasty with a cemented femoral component: a 7–10 year follow-up study. *J. Arthroplasty*, 2004, **19**(8), Suppl. 3, 17–22.

5 **Daniel, J., Pynsent, P. B.,** and **McMinn, D. J.** Metal-on-metal resurfacing of the hip in patients under the age of 55 years with osteoarthritis. *J. Bone Jt Surg. Br.*, 2004, **86**(2), 177–184.

6 **Shimmin, A. J., Bare, J.,** and **Back, D. L.** Complications associated with hip resurfacing arthroplasty. *Orthop. Clin. North Am.*, 2005, **36**(2), 187–193.

7 **Amstutz, H. C., Beaule, P. E., Dorey, F. J., Le Duff, M. J., Campbell, P. A.,** and **Gruen, T. A.** Metal-on-metal hybrid surface arthroplasty: two to six-year follow-up study. *J. Bone Jt Surg. Am.*, 2004, **86A**(1), 28–39.

8 **Shimmin, A. J.** and **Back, D.** Femoral neck fractures following Birmingham hip resurfacing. *J. Bone Jt Surg. Br.*, 2005, **87**(4), 463–464.

9 **Little, C. P., Ruiz, A. L., Harding, I. J., McLardy-Smith, P, Gundle, R., Murray, D. W.,** and **Athanasou, N. A.** Osteonecrosis in retrieved femoral heads after failed resurfacing arthroplasty of the hip. *J. Bone Jt Surg. Br.*, 2005, **87**(3), 320–323.

10 **Goldberg, V.** Results of multi-center trial: strategies to prevent failure and minimize the 'learning curve'. In *Instructional Course Lecture Handout: Surface Arthroplasty of the Hip* (Eds D. W. Murray, P. E. Beaule, P. A. Campbell, V. Goldberg, and H. C. Amstutz), 2005 (AAOS, Washington).

11 **Herberts, P.** and **Malchau, H.** Annual Report 2003, The Swedish National Hip Arthroplasty Register, 2004, Goeteborg, Sweden.

12 **Willert, H. G., Buchhorn, G. H., Gobel, D., Koster, G., Schaffner, S., Schenk, R.,** and **Semlitsch, M.** Wear behavior and histopathology of classic cemented metal on metal hip endoprostheses. *Clin. Orthop. Rel. Res.*, 1996, **329**, Suppl., S160–S186.

13 **Kothari, M., Bartel, D. L.,** and **Booker, J. F.** Surface geometry of retrieved McKee–Farrar total hip

replacements. *Clin. Orthop. Rel. Res.*, **329**, Suppl., S141–S147.

14 Hahn, M., Vogel, M., and **Delling, G.** Undecalcified preparation of bone tissue: report of technical experience and development of new methods. In *Virchows Archiv A: Pathological Anatomy and Histopathology*, 1991, Vol. **418**, pp. 1–7.

15 Howie, D. W., Cornish, B. L., and **Vernon-Roberts, B.** The viability of the femoral head after resurfacing hip arthroplasty in humans. *Clin. Orthop. Rel. Res.*, 1993, **291**, 171–184.

16 Amstutz, H. C., Campbell, P. A., and **Le Duff, M. J.** Fracture of the neck of the femur after surface arthroplasty of the hip. *J. Bone Jt Surg. Am.*, 2004, **86A**(9), 1874–1877.

17 McKellop, H., Park, S. H., Chiesa, R., Doorn, P., Lu, B., Normand, P., Grigoris, P., and **Amstutz, H.** *In vivo* wear of three types of metal on metal hip prostheses during two decades of use. *Clin. Orthop. Rel. Res.*, 1996, **329**, Suppl., S128–S140.

18 Rieker, C., Schön, R., and **Köttig, P.** Development and validation of a second-generation metal-on-metal bearing. *J. Arthroplasty*, 2004, **19**(8), Suppl. 3, 5–11.

19 Brodner, W., Grubl, A., Jankovsky, R., Meisinger, V., Lehr, S., and **Gottsauner-Wolf, F.** Cup inclination and serum concentration of cobalt and chromium after metal-on-metal total hip arthroplasty. *J. Arthroplasty*, 2004, **19**(8), Suppl. 3, 66–70.

20 Beaule, P. E., Lee, J. L., Le Duff, M. J., Amstutz, H. C., and **Ebramzadeh, E.** Orientation of the femoral component in surface arthroplasty of the hip. A biomechanical and clinical analysis. *J. Bone Jt Surg. Am.*, 2004, **86A**(9), 2015–2021.