
CLINICAL ARTICLE

Analysis of wear-related *failures* in long-term follow-up of crosslinked polyethylene hip sockets

CJ Grobbelaar MBChB, MMed Surg (Orth), MD Orth

Professor (extra-ord) University of Pretoria; Department of Orthopaedic Surgery, University of Pretoria
TA du Plessis MSc, DSc

Gammatron, Pretoria

FE Steffens MSc, DSc, THOD

Professor, Department of Statistics, University of Pretoria

MJ van der Linde PhD

Senior research consultant, Department of Statistics, University of Pretoria

Reprint requests:

Prof CJ Grobbelaar

PO Box 2083

Brooklyn Square

Pretoria 0075

Email: cjgrobbelaar@netactive.co.za

Abstract

The publication of our long-term hip replacement follow-up study computer analysis has clearly demonstrated the association between polyethylene (polyethylene/UHMWPE) wear on the one hand and pain, interface widening and osteolysis on the other¹. From a total of 97 cases, seven had true failure and needed ultimate revision. These seven cases formed the nucleus of the current study (Group 1). The course of events could then be analysed in great detail, and compared with the other 90 cases (Groups 2, 3, and 4).

The purpose of this paper was to confirm the role of polyethylene wear (the independent variable) in the limitation of longevity of the implant; furthermore to identify the cut-off point of wear that is considered acceptable. At the same time quantification of P, I, O (pain, interface widening and osteolysis) was studied (the dependent variables). Once again, computer analysis provided us with detailed values that would constitute the cut-off points for acceptance.

These findings then enabled us to categorise the 97 cases into four groups. Of special interest was Group 2: 'impending failure'. Even though none of the nine cases in Group 2 were revised, this study clearly demonstrated that they should be classified as failures and managed accordingly.

In this study some other controversial issues were addressed: the **degree** of wear proved important as opposed to the **tempo**, which was not! Equally unimpressive were the patients' age, activity and body mass, which according to digital analysis had little effect on implant longevity.

Introduction

The correlation between **wear** and P, I, O (**pain, interface status** and **osteolysis**) was already clearly established in a long-term study of metal/polyethylene hip replacement¹. Before the present study we realised that there emerged an opportunity to find answers to more detailed questions pertaining to longevity in total hip replacement.

Some answers came from 'simple' statistical studies, such as the influence on implant longevity of age, sex, body mass and activity level. Other important issues needed more sophisticated analysis:

- Tempo of wear versus total wear
- The cut-off point for acceptable total wear
- In-depth study of cut-off points for pain, interface status and osteolysis (collective as well as individual data).

In these aspects the Department of Statistics at the University of Pretoria provided invaluable sophisticated computer assistance.

Group 1 of our study series consisted of seven cases in which everyone had been revised due to polyethylene wear failure. In keeping with numerous literature studies^{3,4} we had to consider the Meier-Kaplan statistical analysis program, which might have been an option if we had been dealing purely with a survival study. However, this study needed much more than a survival analysis in order to accommodate the other three groups, consisting of 90 cases or 93% of our study cohort. The Meier-Kaplan analysis method thus had to be rejected in favour of a choice of the following statistical methods:

- Generalised Linear Model (GLM)
- Least Square Means comparisons (LSM)
- Correlation Analysis (Regression)
- Chi Square test
- Fisher's Exact test
- Frequency Procedure

Presentations were done through scatter-plots, regression graphs, bar charts and cross tables.

We attempted in this study to emphasise the importance that Group 2 cases should be seen as **impending failures** which is really only one small step away from **true failures**, as our materials and methods will clearly show. These impending failures should thus be included in our failures cohort, and consequently, the criteria to determine failure should **not** be limited to 'revision as end point', but should rather include pain, function, and radiographic issues.

An important goal of this study was therefore to get clarity on the acceptable values of these criteria, and so to redefine the concept of failure, as per *Figure 1*. At this stage values of independent variables were not yet known and only became available towards the end of this study. The expected (if incomplete) outline as per *Figure 1* was already formulated at the very start of this study, even if completion could only be done later, in response to study results. It did however, form the anticipated nucleus of the study.

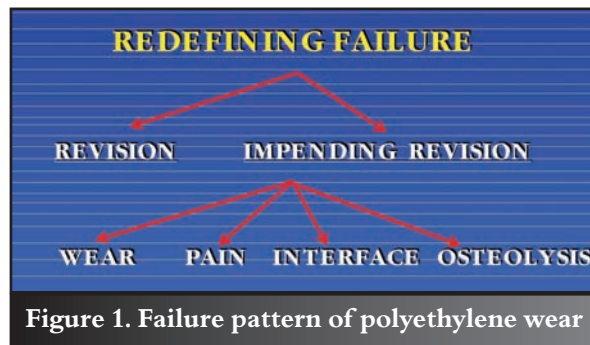


Figure 1. Failure pattern of polyethylene wear

Materials, methods and results

Materials were provided by 97 hips utilising metal/UHMWPE (ultra high molecular weight polyethylene). Gamma crosslinking of polyethylene took place on pre-manufactured machined cups to a depth of 300 $\mu^{5,6}$. *Table I* confirms the statistics as per previous publication by the first author². Wear measurement was according to the DMM method.¹²

On the negative side in *Table I*, there were seven revisions out of the 97 cases. Six of these revisions had wear in excess of 1 mm. On the positive side, it should be noted that revisions were only performed after a mean period of 20.12 years post-op which makes this a true long-term follow-up study.

In the current study we acknowledged the four groups of hips in terms of outcomes^{1,2}. Special attention was now given to the failures leading to revisions in Group I, consisting of seven cases. The differences were then demonstrated in detail as the other three groups were studied individually (*Table II*).

Group 1

All seven hips in this group had to be revised for wear-related failure. The course of every patient is displayed in *Table II*. The mean average wear of 0.079 mm/year was more than five times that of the 97 cases series average of 0.015 mm/year. Since the outcome of every one of the seven cases in this group was revision surgery, we could pin down implant longevity to 19.00 years (*see Table VIII*). In contrast, the mean follow-up values of groups 2, 3 and 4 were only 'way-side' values that were still going to rise with time, since in these groups the failure-driven end points have not been reached by any means. Furthermore in Group 1, the average total wear was 1.59 mm, with six out of the seven presenting with wear of more than 1 mm. The dependent variables (P, I, O) were likewise excessively raised, and it is this association that our computer programs have already clearly showed¹. Perhaps the most important independent variable indicative of failure was osteolysis – invariably a grave sign indicating a poor prognosis!¹⁷ The average Group 1 osteolysis score was degree (med) 3 as opposed to degree 0 which is normal.

Table I: Previous analogue study – follow-up polyethylene wear after 10-33 years (mean 20.12)

Wear	Number of cases	Degrees 1-3 osteolysis	Revisions - actual	Revisions - impending	Interface improvement
Nil	68	0	0	0	15
<1 mm	15	6	1	0	
≥1 mm	14	14	6	9	1
Total	97	20	7	9	16

Table II: Group 1 (seven cases): Cases with wear and absolute indication for revision – true revisions

Patient gamma number	Annual individual wear (mm)	Total individual wear (mm)		Follow-up 2 - pain (Merle d' Aubigne)	Follow-up 2 - acetabular interface (mm)	Follow-up 2 - femoral interface (mm)	Follow-up 2 - degree of osteolysis ≥2	Outcome scores	True revisions
		<1mm	≥1mm						
154	0.102		2.00	4 (score 2)	2.00	3.00	3	12.10	1
297	0.088		1.25	5 (score 1)	5.00	0.00	3	10.34	1
354	0.082		2.00	4 (score 2)	8.50	0.00	3	15.58	1
369	0.068		1.10	4 (score 2)	2.00	0.00	2	7.17	1
550	0.199		3.00	5 (score 1)	8.00	0.50	3	15.70	1
621	0.081		1.00	4 (score 2)	5.00	0.00	1	9.08	1
686	0.038	0.75		4 (score 2)	1.00	2.00	3	8.79	1
Total	0.658	11.10		12.00	31.50	5.50	18.00	78.76	7
Average	0.094	1.59		1.71	4.50	0.79	2.57	11.25	all

Table III: Pain values in relation to pain scores

M d' Aubigne	Pain score
6 = no pain	0
5 = slight pain	1
4 = moderate pain	2
3 = severe pain	3

Next, a scoring system (outcome score) was introduced, which proved invaluable to determine management of the patient: the sum of the wear, pain, interface width and the degree of osteolysis (P, I, O) determined the management of the patient who presented with polyethylene cup wear. Real values were added except for pain which had to be inverted since the lower the score, the better the result achieved (*Table III*). The row totals in column 9 (*Table II*) depicted the outcome score.

Group 2

This was a most interesting and important group. Although wear of 1 mm or more was measured in all nine cases (total average individual wear averaging 1.33 mm) the dependent variables (P, I, O) were still acceptable in terms of the association (corr.) programme as well as the regression graphics. The caveat proved to be the combination of wear more than 1 mm combined with osteolysis more than degree 2, when careful and regular follow-up was indicated. An average outcome score of 5.89, (column 9, *Table IV*) in Group 2 clearly excludes it from being classified as being a successful outcome¹. For these reasons Group 2 should be seen as **impending revisions/failures**.

Group 3

This group represented the other 13 cases with measurable wear (*Table V*). The wear, however was limited to 0.019 mm average/year, which only marginally surpassed the study average of 0.015 mm/year. Once again, the independent variables (P, I, O) reflected the excellent prognosis through exceptional P-values in the computer study.

Group 2

Table IV: Group 2 (nine cases): Cases with wear ≥ 1 mm but no indications for revision								
Patient gamma number	Annual individual wear (mm)	Total individual wear (mm)		Follow-up 2 - pain (Merle d' Aubigne)	Follow-up 2 - acetabular interface (mm)	Follow-up 2 - femoral interface (mm)	Follow-up 2 - degree of osteolysis ≥ 2	Outcomes score
		<1 mm	≥ 1 mm					
287	0.138		2.00	4 (score 2)	1.00	0.00	1	6.14
590	0.102		1.75	4 (score 2)	4.00	1.00	3	11.85
647	0.066		1.20	5 (score 1)	1.00	5.00	3	11.27
168	0.034		1.00	6 (score 0)	-0.25	0.00	0	0.78
288	0.088		1.25	4 (score 2)	1.50	2.00	0	6.84
295	0.035		1.05	6 (score 0)	0.00	0.00	1	2.09
319	0.043		1.25	5 (score 1)	0.50	0.00	0	2.79
715	0.057		1.55	6 (score 0)	2.00	0.00	3	6.61
815	0.063		1.10	5 (score 1)	0.50	0.00	2	4.67
Total	0.626	12.15		9.00	10.25	8.00	13	53.04
Average	0.070	1.35		1.00	1.14	0.89	1.44	5.89

Group 3

Table V: Group 3 (13 cases): Cases with wear <1 mm, but no impending revision – insufficient associated danger signs (raw data P6/7/8)								
Patient gamma number	Annual individual wear (mm)	Total individual wear (mm)		Follow-up 2 - pain (Merle d' Aubigne)	Follow-up 2 - acetabular interface (mm)	Follow-up 2 - femoral interface (mm)	Follow-up 2 - degree of osteolysis ≥ 2	Outcomes score
		<1 mm	≥ 1 mm					
168	0.029	0.900		6 (score 0)	0.00	0.00	1	1.93
201	0.026	0.500		6 (score 0)	-0.25	0.50	0	0.78
212	0.031	0.500		6 (score 0)	0.00	0.00	0	0.53
362	0.003	0.092		6 (score 0)	0.35	0.00	0	0.45
469	0.015	0.400		5 (score 1)	0.23	0.00	0	1.65
472	0.007	0.187		6 (score 0)	0.00	0.00	0	0.19
479	0.030	0.814		6 (score 0)	0.50	0.00	0	1.34
538	0.018	0.500		5 (score 1)	0.00	0.00	0	1.52
553	0.033	0.500		6 (score 0)	1.00	0.00	0	1.53
660	0.030	0.810		6 (score 0)	-1.00	0.00	0	-0.16
673	0.010	0.260		6 (score 0)	1.50	0.00	0	1.77
870	0.010	0.250		6 (score 0)	-0.80	0.00	0	-0.54
1076	0.015	0.345		6 (score 0)	-0.25	0.00	0	0.11
Total	0.257	6.058		2.00	1.28	0.50	1.00	11.09
Average	0.0198	0.047		0.15	0.098	0.039	0.077	0.85

Table VI: Group 4 (68 hips): Neither wear nor danger signs

	Annual individual wear (mm)	Total individual wear (mm)	Follow-up 2 - pain (Merle d' Aubigne)	Follow-up 2 - acetabular interface (mm)	Follow-up 2 - femoral interface (mm)	Follow-up 2 - degree of osteolysis ≥ 2	Danger signs (revision)
Average	0	0	Mode 1	<1	N/A	0	0

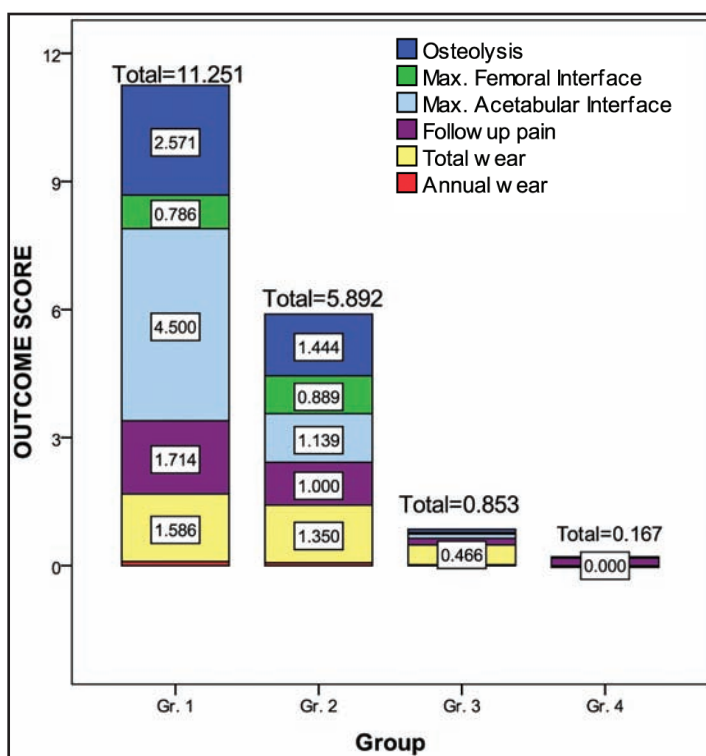


Figure 2. Combination of values of different variables to classify patients into groups 1-4.

See below for bar 1 statistics:

Indiv. annual wear mm (refer to bar 1 above)	Group 1	Group 2	Group 3	Group 4	Series average
	0.094	0.07	0.0198	0	0.015

Of particular interest, but not unexpected, was the virtual absence of **osteolysis**. Danger signs in this group were widening of the acetabular interface, which was uniformly present, albeit in very limited measurements.

Group 4

The 68 hips in this cohort (70.1%) presented with remarkable absence of any degree of measurable wear, combined with a strong association with independent variables

(P, I, O – pain, interface and osteolysis). A notable feature (made possible by our computerised programs) was the following: in individual as well as collective data, the average 18.53 years' follow-up time for Group 4 correlated well with the rest of the study groups (1, 2 and 3) – *Table VIII*. It cannot therefore be argued that these excellent results could be due to a shorter follow-up period.

The column graph in *Figure 2* emphasises the convincing value of the scoring system to categorise our patients into groups 1-4.

In view of the minute values of annual cup wear after gamma crosslinking, bar 1 in each column was so small that a separate box was created below the graph. The individual annual wear for groups 1, 2, 3 and 4 ranged from 0.094 mm down to 0.000 mm, which is in close agreement with the combined values of variables in this graph.

According to the model depicted in *Figure 2*, management can be facilitated by that information, as depicted in *Table VII*.

Next, we studied the importance of body mass, age at operation, and the affect of the tempo of wear versus total wear, on an individual as well as a collective basis (*Table VIII*).

Mean ages at operation, body mass at operation and mean follow-up times were in good agreement in all four groups. Their influence on longevity of the cups thus seemed insignificant. Likewise, the excellent results in Group 4 were clearly not due to **shorter** follow-up times. Total wear of more than 1 mm resulted in markedly higher failure rates clearly shown in Groups 1 and 2 (column 7, *Table VIII*).

Discussion

With the arrival and application of computerised statistical analysis our long-term wear study (10-33 years) has taken on a different meaning. Analogue results previously obtained (*Table I*) proved correct and valuable, but lacked true sensitivity and versatility. It was clear through these computer programs that the 1 mm cut-off point for wear was of great importance, and it also became clear that a new category should be added, namely **impending failure** (Group 2). There were nine cases in this group versus seven cases of **true failures** (Group 1).

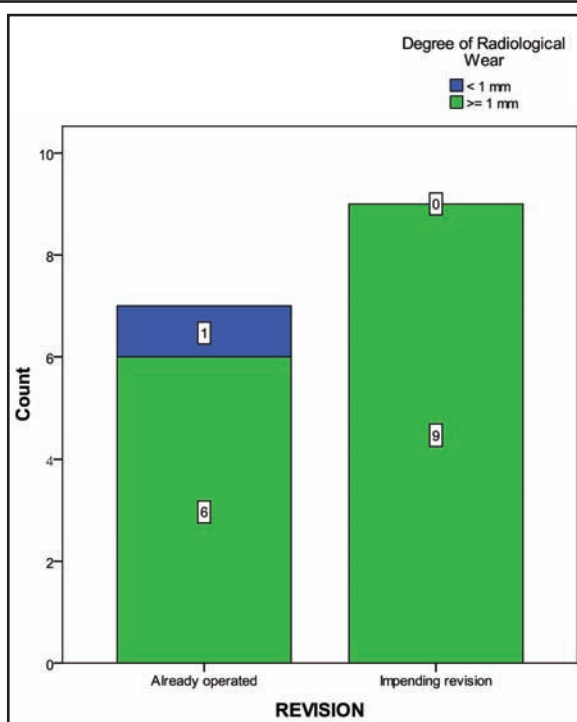
The column graphs in *Figures 3 and 4* display conclusive evidence of the practical implications of the 1 mm cut-off point in acetabular cup wear.

Table VII: Management according to outcome score

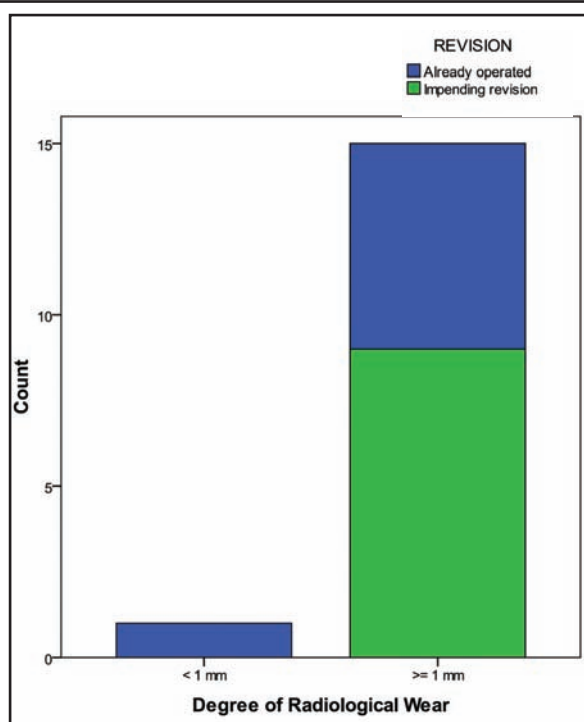
Group	Average score	Management	Score range
1	11.25	Revision	>10
2	5.89	Impending revision/ close observation	5 to 10
3	0.85	Routine observation	< 5
4	0.16	No treatment	± 0.00

Table VIII

Group	Mean age at operation (yrs)	Mean mass at operation (kg)	Mean follow-up (yrs)	Total wear in 21.12 yrs (mm)	Total wear/case in 20.12 yrs (cases)	Total wear/case/yr (indiv. ann. wear) (mm)
1	50.29	71.14	19	11.1	7	1.58
2	50.67	65.78	19.3	12.15	9	1.35
3	44.83	68.75	24.8	6.06	13	0.47
4	54.62	70.15	18.53	0.00	68	0.00
Series average	50.1	68.96	20.12			0.15
Total				29.308	97	

**Figure 3**

From a combined total of 16 cases of true and impending revision, 15 out of 16 (93%) had wear more than 1 mm. The importance of the 1 mm cut-off point is confirmed.

**Figure 4**

From a combined total of 16 cases of true and impending revision only one case had wear of less than 1 mm. Again, the 1 mm cut-off point for wear is confirmed.

Instead of the traditionally calculated (Meier-Kaplan) seven failures from 97 cases (7.2%), we now had a self-inflicted failure rate of 16 failures from the 97 cases (16.4%). The question immediately arose – is crosslinking really that good? The answer lies in the following comparisons with virgin polyethylene cups' wear performance: our crosslinked group failures took 12.6 years average from implantation to revision. Wear tempo was 0.079 mm average/year. According to world literature the average virgin cup will wear at 0.1 mm/year. The failure point of 1 mm wear was thus reported to occur at ± 10 years in uncrosslinked polyethylene cups⁸. Consequently the accepted wear performance (longevity) of the seven failed crosslinked cups in our study was still better than those of the average 'successful' virgin cups.

Summary of relevant information arising from this study

- The tempo of wear was not found to be important (Table VI). What did matter was the degree of total wear.
- International experience found wear in virgin polyethylene cups to be ± 1 mm average per year. The series average of our gamma crosslinked polyethylene cups was only 0.015, i.e. 6.66 times better.
- The cut-off point of 1 mm total wear was confirmed and was independent of the time it took to reach this value¹. Previously virgin polyethylene cups were found to fail after 10 years only if wear exceeded 1 mm.
- The provision was that other (dependent) variables also had to be considered and included in the equation. In particular these variables made an all-important contribution towards classification as either **true (established) failures or impending failures**.
- The 1-2 mm cut-off point, for instance, also applied to the acetabular bone cement interface.
- Patients' age, sex and body mass were remarkably similar in all four groups and therefore played no statistical role in wear variation and its consequences.

- Pain was not found to be the result of wear per se; however, when wear exceeded 1 mm, pain became common due to an increase in particular in polyethylene debris and its consequences (inflammation, granuloma, osteolysis and finally looseness).
- Literature on 30 years plus' results with polyethylene hip cups are uncommon. Those that did reach publication did not always echo our findings, in some important aspects. Wroblewski⁹ did not find evidence of osteolysis even in advanced worn Charnley cups. He ascribed looseness to gross mechanical features causing impingement of the neck against the socket. John Callaghan *et al.*¹⁰ reported on an ultra-long (35 years) follow-up of Charnley hips. There was only 4.8% survival with 50% of them already revised. Thus, actually only 2.4% survived without revision. However polyethylene wear was not even mentioned as a possible cause of failure (Meier-Kaplan study).
- The Meier-Kaplan Survivorship Analysis¹¹ was rejected by our Statistics Department, since it was unsuitable for ongoing patients as in our groups 2, 3 and 4, constituting 90% of our cases. Our analysis program selection had to be able to handle a different purpose and this was mentioned in Figure 1.

Conclusion

The fact that we had the privilege to study a most reliable, wear-resistant hip prosthesis enabled us to report on a reasonable number of survivors at 10-33 years. The expertise of the Department of Statistics at the University of Pretoria brought a new dimension to the issues pertaining to implant longevity.

Associations were established for the different variables¹, which yielded surprising results in some ways, and confirmed certain facts that were already suspected.

Two weaknesses of our study need mentioning:

- Collection of wear data on X-ray images remains our Achilles heel. We believe that the measuring method that we use is simple and reliable. However, although magnification of the X-rays by 4-5 times improves matters, the interface sometimes remains somewhat hazy. We hope that future radiographic research can provide us with even better quality images.

Table IX: Criteria for failure and treatment

Wear ≥ 1 mm			
Group 1	True revisions	Group 2	Impending revisions
Pain (Md'A)	3 or 4	Pain (Md'A)	5 or 6
Interface (mm)	>2	Interface (mm)	1 to 2
Osteolysis (degree)	2 or 3	Osteolysis (degree)	1
Wear + 2 other variables - revision surgery		Wear + 1 other variable - regular follow up	

This study enabled us to set criteria by which we could categorise our hips into four groups, where cut-off values for revisions could be set

- We did not find femoral looseness/interface data of much value in this study; however, this did not limit the value of this study in any significant way.

Finally, this study enabled us to set criteria by which we could categorise our hips into four groups, where cut-off values for revisions could be set, as demonstrated in *Table IX*.

Once we have classified our patients as Group 1 (failure) or Group 2 (impending failure), *Table IX* presents us with the criteria for revision versus regular follow-up. The value of a scoring system is clearly depicted, where Groups 1 and 2 can be differentiated between. Surgery versus regular follow up will depend on the frequency (> 2 or < 2) of the dependent variables.

The future

The high occurrence (16%) of improvement of the interface after total hip replacement, in our opinion, is of great importance and has not been generally acknowledged in world literature. It justifies in-depth statistical analysis and this study is presently under way.

The content of this article is the sole work of the authors. No benefits of any form have been received from a commercial party related directly or indirectly to the subject of this article. All subjects included in this study provided their written informed consent.

References

1. Grobbelaar CJ, *et al.* Statistical analysis of association between wear and pain, interface changes and osteolysis. *Article in Press, SA Orth J.* 2011.
2. Grobbelaar CJ, Weber FA, Du Plessis. Thirty-three years of clinical experience with polyethylene crosslinking in cemented total hip replacement. *Article in Press, SA Orth J.* 2011.
3. Zhongmin JIN, *et al.* Tribology of hip joints. *J of Biomech Science and Engin.* 2006; **Vol I**:1.
4. Ingham J, Fisher J. The role of macrophages. *J Biomat.* 2005; **26**:1271-1286.
5. Du Plessis TA. Van Kinetika tot kunstgewrigte. *SA tydskrif vir Natuurwet en Tegn.* 1984; **3**:16-19.
6. Du Plessis TA, Grobbelaar CJ, Marais F. The improvement of polyethylene prosthesis through radiation crosslinking. *Rad Phys Chem.* 1977; **9**:647-51.
7. Harris WH. The problem is osteolysis. *Clin Orthop Rel Res.* 1995; **46**:311-16.
8. Kurtz S, Medel FJ, Manley M. Wear in highly crosslinked polyethylenes. *Current Orthopedics.* 2008; **22**:392-99.
9. Wroblewski BM, Siney BA. Charnley low-frictional torque arthroplasty – follow up 30-40 years. *J Bone Joint Surg Br.* 2009; **91-B**: 447-50.
10. Callaghan JJ, Bracha P, Liu Steve S. *J Bone Joint Surg.* 2009; **91-A** 11: 25-8.
11. Kaplan EL, Meier P. Non parametric estimation from incomplete observations. *J Am Stat Assoc.* 1958; **53**: 457-81.
12. Grobbelaar CJ. A new method to measure polyethylene cup wear in total hip replacement – a preliminary report. *SA Orth J.* 2007; **6**:18-24.