Published in final edited form as:

Proc Inst Mech Eng H. 2010 ; 224(4): 585–597.

Using “Sub-cement” to simulate the long-term fatigue response of cemented femoral stems in a cadaver model: could a novel pre-clinical screening test have caught the Exeter matte problem?

Amos Race, Ph.D. [Associate Professor of Orthopedic Surgery]

SUNY Upstate Medical University

Mark A. Miller, M.S [Research Engineer]

SUNY Upstate Medical University

Kenneth A. Mann, Ph.D. [Professor of Orthopedic Surgery]

SUNY Upstate Medical University

Abstract

Previously, we formulated cement with degraded fatigue properties (sub-cement) to simulate long-term fatigue in short-term cadaver tests. The present study determined the efficacy of sub-cement in a ‘pre-clinical’ test of a design change with known clinical consequences: the “polished” to “matte” transition of the *Exeter* stem (revision rates were twice as high for matte stems). Contemporary stems were bead-blasted to give $R_a=1\mu\text{m}$ (matte finish). Matte and polished stems were compared in cadaver pairs under stair-climbing loads (3 pairs size-1, 3 pairs size-3). Stem micromotion was monitored during loading. Post-test, transverse sections were examined for cement damage.

Cyclic retroversion decreased for polished stems but increased for matte stems ($p<0.0001$). Implant size had a substantial effect: retroversion of (larger) size-3 stems was half that of size-1 stems and polished size-3 stems subsided $2\frac{1}{2}$ times more than the others. Cement damage measures were similar and open through-cracks occurred around both stems of two pairs. Stem retroversion within the mantle resulted in stem-cement gaps of 50–150 microns.

Combining information on cyclic motion, cracks, and gaps, we concluded that this test ‘predicted’ higher revision rates for matte stems (it also implied that polished size-3 stems might be superior to size-1).

Keywords

hip; pre-clinical; cadaver; cement; sub-cement; femoral stem; surface finish; micro motion; fatigue; damage

1 Introduction

Despite the widespread and ongoing use of PMMA cement for femoral stem fixation in hip arthroplasty, the behaviour of the cement mantle in a stem/cement/femur composite is poorly understood. In the decades since Sir John Charnley’s somewhat serendipitous development of successful total hip arthroplasty in the early 1960’s, many new designs of cemented femoral stem have been introduced, but with mixed clinical results [1]. Some seemingly trivial design

changes have resulted in markedly poorer clinical outcomes, others have not. The apparent unpredictability of clinical outcomes led to a more cautious attitude toward femoral stem innovation [2,3] and an interest in developing better pre-clinical screening methods.

Meeting the need for a better predictor of clinical success will require improvements to both physical and computational models, fed by data from in-vivo monitoring and post-mortem analyses. Physical models using cadaver femora provide the simplest method to generate clinically realistic cement mantles. However, stress concentrations at the external loading points cause unrealistic cortical fatigue-fractures during long-term cyclic loading [4]. In addition, there are concerns regarding bone degradation and excessive cost [5]. To get around these limitations, we have formulated bone cement with degraded fatigue fracture properties (Sub-cement) designed to have a time-scaling effect such that long-term fatigue can be simulated successfully in short-term cadaver tests [6].

Having developed sub-cement as a pre-clinical testing tool, we sought to test its efficacy by 'predicting' the outcome of a design change with known consequences. The original design of *Exeter* femoral stem had excellent results with respect to aseptic loosening [7] but had problems with stem fracture in heavier/active patients [8]. This led to slight modifications of stem geometry and improvements in the strength and toughness of the base material. The original stems had been polished, but the newer materials were left with a matte finish. These changes substantially reduced the incidence of fracture, but the long-term clinical outcomes were significantly worse, measured by rates of revision and osteolysis [7,9,10]. The polished surface finish was reinstated after ten years, along with some further modifications to the proximal geometry, which restored the earlier level of clinical success [11]. The success/failure/success of the *Exeter* stem was attributed to the changes in surface finish after retrieval studies showed clear evidence of abrasive wear at the stem cement interface of matte stems [12,13]. This clinical history led to the concept of "taper-lock", by which polished, tapered stems were thought to become more stable as they subsided into the cement mantle [14].

We hypothesised that using sub-cement in our previously described cadaveric fatigue loading model [15] would enable us to discriminate between *Exeter* stems with polished and matte finishes, based on analyses of stem micro-motions and cement fatigue damage. More specifically, we hypothesised that: (a) matte stems would display increasing cyclic motion (which would be indicative of loosening, as suggested by Maher et al [16]) and polished stems would not; (b) matte stems would result in more cement mantle damage than polished stems; (c) Cyclic motion would correlate with cement damage; and (d) Polished stems, but not matte ones, would show decreased cyclic motion with greater stem subsidence (although it would not have been made *a priori*, this last hypothesis was included because, to the authors' knowledge, the oft-quoted "taper-lock" phenomenon has never been experimentally validated in a stem/cement/femur construct).

2 Materials and Methods

2.1 Sub-cement manufacture

Sub-cement was made by adding a chain-transfer agent (1-dodecanethiol, Acros Organics USA, Morris Plains, NJ) to the liquid phase of standard PMMA based powder/liquid cement (Simplex-P, Stryker Orthopaedics Inc., Mawah, New Jersey). The addition of the chain-transfer agent reduced the molecular weight of the inter-bead matrix without changing reaction-rate or handling characteristics. The static properties of Sub-cement are approximately equivalent to normal cement, but fatigue failure is accelerated by a factor of between 30 & 100 [6]. Thus, Sub-cement enabled us to conduct time-scaled fatigue testing of cemented stems in cadaver femora

2.2 Specimen preparation

Six pairs of fresh-frozen cadaver femora were obtained from our institution's willied body program. Median donor age was 75 years (59–83, 3 male 3 female). Paired femora were implanted with pairs of contemporary *Exeter* stems, one of which was treated to mimic the finish of the historical matte stem.

Two, original, matte stems in our possession were examined microscopically and measured for surface roughness. We found that the same surface appearance and roughness ($R_a \sim 1\mu\text{m}$) could be produced by bead-blasting with 0.2mm diameter glass beads. Two stem sizes were used, three pairs of femora took a size-3 and three took a size-1. The stems were the same length and had the same sagittal plane taper but the size-3 stems were broader in the frontal plane, having a wider taper angle and a proximo-lateral curved shoulder. In order to fit the stems with a micro-motion target, they were drilled and tapped to take a USF10–32 screw. To prevent void formation by entrapped air during cementation, the tapped holes were temporarily filled with an appropriately shaped nylon screw.

The femora were stripped of all soft tissue, trimmed distally and potted in acrylic blocks (stem version was defined by the neck anatomy, so the distal condyles were not required for reference). Other than the substitution of Sub-cement, stem implantation followed contemporary protocol: femora were cut and broached following standard procedures for the *Exeter* stem, cleaned with vigorous brush lavage, fitted with a distal plug, and subjected to pressurisation with a proximal seal. These methods were intended to result in a bone free zone of at least 2mm around the stem and cement interdigitation to the cortex (“whiteout” on a radiograph). Reproducible implantation under simulated physiological conditions was carried out using previously described techniques [15]. To better simulate the in vivo thermal and fluid environment, cementation was conducted with the femora held in a vat of 37°C blood analogue. To ensure reproducible insertion, the stems were implanted (to the middle of the three indicator marks) using a materials testing machine. The femora were rigidly fixed to the base of the machine with the axis of the broached cavity vertical and coincident with the axis of the stem, which was rigidly fixed to the cross-head. Reproducibility was further ensured by triggering cement and stem insertion by cement viscosity measures. In a preliminary study, we recorded the viscosity of cement (using a Brookfield DV-E Viscometer with small sample adapter at a shear rate = 1/sec) at the application and stem insertion points defined by an experienced orthopaedic surgeon. Based on those data, and the manufacturer's recommended working phase times, we chose to apply cement at $\eta = 1000\text{Pa}\cdot\text{s}$ and to insert the stem at $\eta = 2000\text{Pa}\cdot\text{s}$. After stem insertion, the constructs were allowed to cure in a blood analogue solution for four days prior to testing. During mechanical testing, hydration was preserved by wrapping the femora in wax film.

2.3 Mechanical testing

The stem/cement/femur constructs were then subject to stair-climbing loading (Figure 1) for up to $10^{5.5}$ (320,000) cycles (equivalent to at least 10^7 cycles ~ 5 years), using a previously described custom jig and protocol [15]. In brief, sinusoidal loading at 3Hz was applied to the femoral head, via a cup, and to the greater trochanter, via a strap. The temperature of the stem/cement/femur construct was maintained at 37°C by wrapping the femur with a heated water circuit, as has been described previously [4]. The femoral head force, in terms of a unit vector in the anatomical co-ordinate system, was directed: 0.83 distally; 0.40 posteriorly; 0.39 laterally. The force on the greater trochanter, was directed: 0.72 proximally; 0.47 anteriorly; 0.51 medially. The ratio of head load to abductor load was 1.5:1. The applied loads were normalised by stem-cement-femur construct stiffness, as measured during the first 100 cycles of loading (the initial applied load was estimated using data from previous studies and cortical dimensions obtained from AP x-rays). This was done in an effort to prevent premature bone

failure. Actual body weights were not used since weight at death is not necessarily a good indicator of a donor's normal weight because of fluid retention or muscle wasting. In addition, in an overweight and under-exercised donor, our aggressive loading protocol would probably have been unrealistic: had they engaged in the simulated activity level then their weight would have been lower and their bone strength higher. We chose our normalisation method to allow maximal loading, making performance differences more apparent, with a minimal chance of premature failure.

Stem micro-motion was measured using a previously described six degrees-of-freedom system [17,18] consisting of a three-ball target attached to the stem (via a 10mm diameter portal through the cement and bone), which was monitored by an array of six LVDTs fixed to the cortex. Stem position, relative to the cortex in the same plane as the micro-motion target, was recorded at the peaks and valleys of cyclic loading.

2.4 Cement damage assessment

Post-test, cement damage was assessed using a staining/imaging technique that has been described previously [15]. The femora were transversely sectioned, using a high speed water irrigated abrasive disk, at 4 locations (20, 40, 80 & 100mm distal to the calcar), stained with a fluorescent dye-penetrant to reveal cement cracks, and electronically imaged via an epifluorescence microscope. These sectioning and staining techniques did not generate microcracks in isolated cement specimens. The imaging system had a resolution of 6 microns/pixel but could detect narrower microcracks because of the fluorescent penetrant. Crack lengths and the areas of the broached and interdigitated parts of the cement mantles were measured using ImagePro.

2.5 Outcome measures

Micromotion data was reduced to two migrations and two corresponding cyclic motions. Stem migrations (subsidence and retroversion) were defined by the valley position of the stem, corresponding cyclic motions by the differences between peak and valley positions.

Cement mantle damage was divided into 3 exclusive categories: *Stem-cement*, encompassing all cracks that touched the stem; *cement-bone*, cracks within the broached area that did not reach the stem; and *interdigitated*, cracks within the interdigitated area. To account for differences in mantle area, normalised damage measures were defined by summing crack lengths for the 4 transverse sections of each bone and dividing by the sum of the relevant area (broached or interdigitated) of the cement mantles. Correlation between cement damage and stem motion was tested using the stem-cement damage category, based on the assumption that this was the most relevant measure.

2.6 Statistical analyses

For hypotheses (a) & (b), donor pairing was accounted for by using generalised linear models with repeated measures. Data from left and right femora were treated as repeated measures on each donor. This was accomplished by using stem finish {*matte* | *polished*} as a within-donors effect in the model. Change in micromotion over the course of loading was analysed by treating $\text{Log}_{10}(\text{Cycles})$ as another within-donors effect, defining additional repeated measures at $\text{Cycles} = 10^2, 10^{2.5}, 10^3, 10^{3.5}, 10^4, 10^{4.5}, 10^5$ & $10^{5.5}$ (the statistical tests were run using SPSS for Mac v16, which can run generalised linear models with multiple within-subjects effects). Stem size {#1 | #3} was included as a between-donors effect.

Hypotheses (c) & (d) dealt with correlations and did not include pairing. The correlation between cyclic motion and stem-cement damage was tested by a simple linear regression. The

taper-lock hypothesis was tested by one-sample t-tests (hypothesised mean = 0) of the gradients of linear fits of subsidence versus cyclic retroversion for the stems of each finish.

3 Results

Two pairs of constructs completed $10^{5.5}$ cycles. Of the other 4 pairs, one or other construct experienced earlier failure at the abductor attachment (the proximal greater trochanter was crushed and/or pulled off). In three of those pairs the first construct tested failed early (at 120, 200 & 230 kcycles), so the contralateral test was halted at the same number of cycles to preserve pairing. In one case, the second construct failed early, but since it reached 81% of completion (260 kcycles), it was retained in the analyses. Because of the early failures, analyses of micromotion data were limited to 10^5 cycles, which allowed all the constructs to be included. For the correlation between micromotion and cement damage, the last micromotion data from each test was used.

Micromotion data, and statistical analyses, are presented in Figure 2 and Table 1. The polished stems subsided and retroverted immediately upon loading, migrating further as loading continued. The matte stems did not immediately migrate, behaving as if the stem-cement interface were bonded (henceforth, “bond” refers to the micro-mechanical interlock between cement and stem). However, all matte stems showed evidence of stem-cement disbonding beyond 10^3 cycles, migrating further under subsequent loading. Disbonding was inferred from the transition from zero-migration to measurable migration.

A striking and unexpected finding was that implant size had a substantial effect: retroversion for size-3 stems was about half that for size-1 stems and polished size-3 stems subsided 2½ times more than the others (Figure 2).

Cement damage data and statistical analyses are presented in Table 2. Damage measures were largely similar across stems and through-cracks were generated in both stem types and sizes (Figure 3).

3.1 Tests of hypotheses

Hypothesis (a), that, as loading progressed, cyclic motion would increase for matte stems but not for polished stems, was accepted for cyclic retroversion and rejected for cyclic subsidence. Cyclic subsidence was approximately two orders of magnitude lower than subsidence, which was itself small - ranging from 5 to 25 microns at 10^5 cycles. Cyclic subsidence did not change with loading, except in the case of polished size-1 stems (2 of the 3 showed decreasing cyclic motion). However, cyclic retroversions, which were a greater fraction of migration, increased with loading for matte stems and decreased for polished.

Hypothesis (b), that matte stems would result in more cement damage, was rejected (Table 2). However, as with micromotion, stem size had an unexpected effect – there was a significant interaction between size and finish for stem-cement cracks. Polished size 1 stems had an order of magnitude more damage at the stem/cement interface than did polished size 3 stems.

Hypothesis (c), that cyclic motion would correlate with cement damage, was accepted for retroversion - cement cracks emanating from the stem were a significant predictor of cyclic retroversion ($p = 0.007$, Figure 4) but rejected for cyclic subsidence ($p = 0.27$, Figure 4).

Hypothesis (d), that taper-lock would occur for polished but not matte stems, was accepted (Figure 5). The mean gradients, of the linear fits for each stem of subsidence versus cyclic retroversion, were 0.34 ± 0.21 (polished, $p = 0.011$) and -0.42 ± 0.66 (matte, $p = 0.18$).

3.2 Additional findings

In general, our micromotion measures could not differentiate between position changes due to elastic deformation, interface slip, or creep. However, because the matte stems were still bonded at 10^2 cycles, their cyclic motion at that point was probably a good estimate of the elastic component of subsequent cyclic motions. This unanticipated finding allowed us to estimate the cyclic sliding motion at the stem/cement interface by subtracting out the elastic deformation included in the cyclic motion measures.

Using the cyclic retroversion at 10^2 cycles for each matte stem to estimate elastic deformation for each donor, we estimated that, at 10^5 cycles, slipping at the stem/cement interface accounted for $0.051 \pm 0.010^\circ$ (size-1) and $0.041 \pm 0.006^\circ$ (size-3) of cyclic rotation for the matte stems, and about $0.068^\circ \pm 0.009$ (size-1) and $0.012 \pm 0.025^\circ$ (size-3) of cyclic rotation for the polished stems. The longest chords passing through the long axes of the stems in the transverse plane of measurement were 10.5mm (size-1) and 13.5mm (size-3). So, by cycle 10^5 , the maximum cyclic sliding distances at the stem/cement interface were approximately 5 microns for the matte stems of both sizes, 7 microns for polished size-1 and 1 micron for polished size-3. Our cyclic subsidence data suggested that, at 10^5 cycles, there was minimal (1 micron or less) proximal/distal sliding.

Originally, we had suspected that the matte/polished difference might have been due to differences in stem-cement apposition, having noted increased stem-cement gaps around rough stems [19] and a correlation between retroversion and stem-cement gap fraction [4]. However, unlike previously studied stems [15,19], it was obvious from the transverse sections that the *Exeter* stems, both matte and polished, had locked into a retroverted position such that all stems had gaps between the stem and cement of between 50 and 150 microns (Figure 3). This meant that we could not properly quantify stem-cement gaps. We did note that, qualitatively, there was not an obvious difference in stem-cement gap fraction between the two stem finishes.

4 Discussion

4.1 Experimental limitations

There are, of course, limitations to any cadaveric model, most obviously the absence of biology. However, despite the lack of bone remodelling, the results of such tests may still shed light on clinical outcomes by careful synthesis of experimental and clinical data, as described below.

A further limitation of this study was that only one type of loading was applied, and that without interruption. Previous workers have suggested, not unreasonably, that hip testing should include alternating normal gait and stair-climbing with dwell periods and occasional 'stumbling' overloads interspersed [20]. It could be argued that constant stair-climbing was the most conservative test, as it was representative of the highest torques on a femoral stem [21] and generated more damage in a previous computational model of a one stem [22]. However, if one were to investigate a new stem design, we believe that it would be prudent to use the alternating loading suggested by Bergmann et al. (2001) since such changes may have unforeseen consequences in a novel system. A related issue is the manner in which hip loads were approximated. Our stair-climbing jig used a single strap to represent the abductor complex and simulated no further muscle forces. This seemed reasonable in the light of the work of Stolk et al [22], who published data for cement stresses in response to various combinations of muscle loads in a computational model.

The sub-cement that we used to time-scale fatigue loading was limited in that it did not accelerate creep. It has oft been asserted that cement creep, leading to subsidence, is crucial to the functional benefit of polished tapered stems in that it facilitates taper-lock. However, to the authors' knowledge, this has never been demonstrated. Computational models of polished

tapers have shown that cement creep contributed minimally to stem subsidence [23] and that subsidence is not required to maintain taper-lock [24]. Rather, taper-lock is a function of the taper angle and the coefficient of friction between stem and cement. So, whilst we acknowledge that it would be preferable to have sub-cement with time-scaled creep, we believe its absence is not a fatal flaw.

Several femora failed prior to completion of the planned cycle limit, which may have biased the cement damage data. Given that 10^5 cycles (which represent between 3 & 10 years of use) were sufficient to demonstrate significant effects in micromotions, we recommend limiting future studies to that number.

Demonstrating that our sectioning and staining techniques did not introduce significant artefact in a stem-cement-femur construct was complicated by the fact that, under a construct's inherent constraints, there are curing induced microcracks in the cement mantle [18,25–27]. However, we were confident that significant artefacts were not introduced by overheating, vibration, or chemical action for the following reasons: a) The material removal rate was slow - each cut took about an hour - so the flood coolant could easily prevent heat build-up; b) Examination of the cut surfaces often revealed extremely tenuous trabecular structures and barely connected or supported bits of cement mantle, which would surely have broken off or been displaced under vibration sufficient to cause cracking; c) In a previous study, we found that microcrack-density increased with applied head-load [18]; d) In the present study, stem related microcrack-density correlated with micromotion.

Although it may have appeared that our cement damage data was at odds with previous reports of post-mortem retrieval studies, which had found cracks initiating at the stem and predominately in thin mantle regions [28,29], this was not in fact the case. Those previous studies did not use a penetrant, so could not have observed the vast majority of microcracks – without a penetrant, the section shown in Figure 3 would only have revealed a stem-initiated through crack in a thin mantle region.

4.2 Statistical limitations

Specimen numbers were limited by available funding and by competing demands for donated tissue. Since we did not anticipate the effect of stem size, the study was, in some respects, underpowered, particularly with regard to the damage measures.

For a pre-clinical test to be practical, it would help if it were more sensitive than clinical outcomes. This may be possible via careful inference from in vitro data, depending on what factors or, more likely, combination of factors, cause the ultimate clinical failure. Some factors may be inherent to the stem design and others may be random. Very small samples may suffice to reveal certain inherent factors (e.g. debris generation, which could be either directly measured or deduced from micromotion data and surface finish) that, in combination with known random possibilities (e.g. mantle defects to transport debris from the stem-cement interface to active bone), would predictably lead to relatively high revision risk. However, if failure depended on some unanticipated random factor, (e.g. some previously undescribed anatomical variation) then it would only be uncovered by a sufficiently large sample to include examples of such unknown-unknowns (say, $n > 100$). The results of the present study suggest that it would be prudent to test implants of all sizes, again requiring greater sample numbers.

4.3 Relation to previous pre-clinical testing studies

Since the outcome measures of pre-clinical tests cannot include “revision”, which is the ultimate measure of clinical failure, what measure can define failure in-vitro? Revision risk must be inferred from surrogate measures of micromotion and cement damage. Inferences from

micromotion are rooted in Radiostereometric Analysis (RSA) studies that have shown that early stem migration is highly correlated with “late” loosening [30]. Inferences from cement damage are based on the many clinical/autopsy/animal studies that have implicated cement fracture, wear particles and pressure pulses as causes of focal osteolysis. However, there is not a well-quantified causal link between clinical failure and these measures: Jasty et al. demonstrated cement fractures in asymptomatic post-mortem retrievals [28], so cement damage cannot be used in isolation; and, although early stem subsidence correlates with late loosening for many stems, the correlation does not hold for polished stems. Thus, it is far from clear what limits on these measures are appropriate to define a poor implant.

In the last decade the challenge of pre-clinical testing has inspired a range of studies, including both physical [16,31–33] and computational modelling [34], with both micromotion and cement damage as outcomes. Rather than come up with an absolute measure to define a poor implant, these recently proposed pre-clinical tests have used micromotion and cement damage measures to rank stems as better/worse. We agree that this seems to be the only viable strategy. This being the case, it would be of great benefit to generate a database of previous implants and implantation techniques. Generating this large body of data would be greatly facilitated if a standardised procedure were to be adopted.

To the authors' knowledge, all other proposed physical pre-clinical tests have been based on synthetic femora. Unfortunately, at present, synthetic femora do not adequately model cancellous bone: they are filled with closed-cell foam that does not have the correct structure or small-scale mechanical properties. This means that there can be no depth of interdigitation and that the constraints on the curing cement mantle are different from that in a human bone.

This difference in constraint, and lack of interdigitation, is manifest by the lack of curing induced microcracks in cement mantles formed in synthetic femora [32]. Also, mantles formed in synthetic femora may be unrealistically thin, since a clinical Grade A mantle (i.e. one that shows ‘whiteout’ on a radiograph [35]) is impossible with a closed-cell foam. Further, interface apposition is likely to be different in synthetic and cadaver femora, due both to the difference in constraint and the difference in thermal properties. Current synthetic femora do have the advantage that more than two stem types can be easily compared. Ideally, a range of synthetic femora will be developed that include a reasonable model of trabecular bone and encompass the range of anatomical variation.

It has been argued that cadaver femora are too variable to allow testing of conveniently small numbers. However, we note (a) that we have previously shown very low variability in a cadaver model of a well-functioning stem/cement/femur system, and (b) that variability is, in itself, a potential indicator of likely clinical problems [15]. RSA studies have shown that stems that were ultimately revised were ‘outliers’ from an early stage [36]. This has implications for sample numbers, statistical analyses and experimental methods – an ideal pre-clinical test protocol would need to both generate and reliably notice such outliers in a clinically relevant way.

We have argued that pre-clinical screening tests based on physical models should use cadaver femora, but that presents practical problems, of cost and availability, in running larger data sets. Obviously, these problems cannot be solved solely by using improved synthetic femora as the required anatomical variability (assuming it could be manufactured cost effectively) would still demand large data sets. So, either the implant development process must include the cost of such large in vitro tests or future stem developments (like those of the past) must be tested in clinical practice. In the latter case, we argue that there is a moral obligation to maintain sufficiently detailed and timely records (such as exist in the Scandinavian Joint Registers) to pull poorly functioning implants from the market at the earliest moment.

4.4 Interpretation and relation to previous clinical data

At first glance, the fact that cyclic retroversion decreased for polished stems and increased for matte appeared to give a clear-cut indication of the inferiority of the latter stem (Figure 2 & Table 1). However, further examination suggested a more careful interpretation. The micromotion data for polished stems showed smooth transitions, whereas the matte stems produced discontinuities. The increase in cyclic retroversion of matte stems occurred during the first $10^{3.5}$ cycles, after which it appeared to stabilise within the range of the polished stems.

In isolation, the cyclic motion of the matte finish could not be considered worse. However, the prospect of increased debris generation at the stem/cement interface invites consideration of the relative risk of osteolysis. Although the aetiology of osteolysis is not fully understood, conduits to living bone, small wear particles and cyclic fluid flow have been strongly implicated [37,38]. These three factors are considered below.

In the present study, all of the stems had gaps between the stem and cement (due to rotational migration within the mantle). These gaps would have presented a conduit for fluid and debris along the length of the stem — this finding was compatible with previous reports of fibrous tissue, up to 100 microns thick, around explanted stems of both polished and matte finish [14,39]. In addition, we found through cracks in some mantles of each stem group, presenting conduits to what would, in vivo, have been living bone. Previous workers have suggested that polished double taper stems showed less osteolysis because the subsiding stem cut off the flow of fluid along the stem /cement interface [13,40]. The current study, and the previous retrieval studies of others [14,39] indicated that this was probably not the case.

With regard to wear particles, the estimated 5 microns of cyclic sliding at the matte stem/cement interface was predictive of small debris generation — this was compatible with previously published work showing burnished regions around retrieved *Exeter* matte stems along with 0.5 – 1 micron metal and cement debris [13]. In contrast, the polished stems would not be expected to produce debris in great quantity.

We found gaps at all stem/cement interfaces. Cyclic sliding at the stem/cement interface would have changed the dimensions of those gaps resulting in cyclic fluid flows. Our estimates of stem/cement sliding implied that, in the case of polished size-3 stems, the stem/cement conduits became almost static. In contrast, the much larger sliding motions of the matte and the size-1 polished stems would have caused fluid to be cyclically pressurised and transported. Cyclic fluid flow induced by cyclic stem motion was previously proposed to explain the nutrient supply to the tissue found between stem and cement [39] and has been demonstrated using an in vitro model of an *Exeter* size-1 stem [41].

By considering our data in combination with the known biological effects of fluid flow and debris we conclude that our proposed cadaver based pre-clinical test, using Sub-cement to accelerate fatigue damage, could indeed have “caught the *Exeter Matte* problem”.

Acknowledgments

This publication was made possible by Grant Numbers AR 50553 and AR 42017 from NIHNIAMS. The authors would like to thank Dr Robert Ploutz-Snyder for assistance with statistical methods.

References

1. Barrack RL. Early failure of modern cemented stems. *Journal of Arthroplasty* 2000;15:1036–50. [PubMed: 11112201]
2. Huiskes R. Failed innovation in total hip replacement. Diagnosis and proposals for a cure. *Acta Orthopaedica Scandinavica* 1993;64:699–716. [PubMed: 8291421]

3. Walker PS. Innovation in total hip replacement--when is new better? *Clinical Orthopaedics & Related Research* 2000;9–25. [PubMed: 11127675]
4. Race A, Miller MA, Clarke MT, Mann KA, Higham PA. The effect of low-viscosity cement on mantle morphology and femoral stem micromotion: a cadaver model with simulated blood flow. *Acta Orthop* 2006;77:607–616. [PubMed: 16929438]
5. Gheduzzi S, Miles AW. A review of pre-clinical testing of femoral stem subsidence and comparison with clinical data. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine* 2007;221:39–46.
6. Race A, Miller MA, Mann KA. A modified PMMA cement (Sub-cement) for accelerated fatigue testing of cemented implant constructs using cadaveric bone. *Journal of Biomechanics* 2008;41:3017–3023. [PubMed: 18774136]
7. Malchau H, Herberts P, Soderman P, Oden A. Prognosis of Total Hip Replacement: Update and validation of results from the Swedish National Hip Arthroplasty Register 2000:1979–1998.
8. Gie GA, Ling RSM. Loosening and migration of Exeter THR. *JBJS-Br* 1994;76:506–507. [PubMed: 8175874]
9. Rockborn P, Olsson SS. Loosening and bone resorption in exeter hip arthroplasties. Review at a minimum of five years [see comments]. *Journal of Bone & Joint Surgery - British Volume* 1993;75:865–8. [PubMed: 8245072]
10. Howie DW, Middleton RG, Costi K. Loosening of matt and polished cemented femoral stems. *Journal of Bone & Joint Surgery - British Volume* 1998;80:573–6. [PubMed: 9699814]
11. Middleton RG, Howie DW, Costi K, Sharpe P. Effects of design changes on cemented tapered femoral stem fixation. *Clinical Orthopaedics & Related Research* 1998;47–56. [PubMed: 9917590]
12. Ling RS. The History and Development of the Exeter Hip. 2004
13. Anthony PP, Gie GA, Howie CR, Ling RSM. Localised endosteal bone lysis in relation to the femoral components of cemented total hip arthroplasties. *Journal of Bone and Joint Surgery* 1990;72:971–979. Series B
14. Fowler JL, Gie GA, Lee AJ, Ling RS. Experience with the Exeter total hip replacement since 1970. *Orthop Clin North Am* 1988;19:477–89. published erratum appears in *Orthop Clin North Am* 1989 Oct;20(4):preceding 519. [PubMed: 3269211]
15. Race A, Miller MA, Clarke MT, Mann KA. Cement/implant interface gaps explain the poor results of CMW3 for femoral stem fixation: a cadaver study of migration, fatigue and mantle morphology. *Acta Orthopaedica* 2005;5:679–87. [PubMed: 16263615]
16. Maher SA, Prendergast PJ. Discriminating the loosening behaviour of cemented hip prostheses using measurements of migration and inducible displacement. *Journal of Biomechanics* 2002;35:257–65. [PubMed: 11784544]
17. Berzins A, Sumner DR, Andriacchi TP, Galante JO. Stem curvature and load angle influence the initial relative bone-implant motion of cementless femoral stems. *Journal of Orthopaedic Research* 1993;11:758–69. erratum appears in *J Orthop Res* 1995 Jan;13(1):151. [PubMed: 8410476]
18. Race A, Miller MA, Ayers DC, Mann KA. Early cement damage around a femoral stem is concentrated at the cement/bone interface. *Journal of Biomechanics* 2003;36:489–496. [PubMed: 12600339]
19. Race A, Miller MA, Ayers DC, Cleary RJ, Mann KA. The influence of surface roughness on stem-cement gaps. *Journal of Bone & Joint Surgery - British Volume* 2002;84:1199–204. [PubMed: 12463671]
20. Bergmann G, Deuretzbacher G, Heller M, Graichen F, Rohlmann A, Strauss J, Duda GN. Hip contact forces and gait patterns from routine activities. *Journal of Biomechanics* 2001;34:859–71. [PubMed: 11410170]
21. Bergmann G, Graichen F, Rohlmann A. Is staircase walking a risk for the fixation of hip implants? *J Biomech* 1995;28:535–53. [PubMed: 7775490]
22. Stolk J, Verdonschot N, Huiskes R. Hip-joint and abductor-muscle forces adequately represent in vivo loading of a cemented total hip reconstruction. *Journal of Biomechanics* 2001;34:917–26. [PubMed: 11410175]
23. Verdonschot N, Huiskes R. Acrylic cement creeps but does not allow much subsidence of femoral stems. *J Bone Joint Surg Br* 1997;79:665–9. [PubMed: 9250762]

24. Norman TL, Thyagarajan G, Saligrama VC, Gruen TA, Blaha JD. Stem surface roughness alters creep induced subsidence and 'taper-lock' in a cemented femoral hip prosthesis. *Journal of Biomechanics* 2001;34:1325–1333. [PubMed: 11522312]
25. Roques A, Browne M, Taylor A, New A, Baker D. Quantitative measurement of the stresses induced during polymerisation of bone cement. *Biomaterials* 2004;25:4415–4424. [PubMed: 15046932]
26. Orr JF, Dunne NJ, Quinn JC. Shrinkage stresses in bone cement. *Biomaterials* 2003;24:2933–2940. [PubMed: 12742733]
27. Lennon AB, Prendergast PJ. Residual stress due to curing can initiate damage in porous bone cement: experimental and theoretical evidence. *Journal of Biomechanics* 2002;35:311–21. [PubMed: 11858806]
28. Jasty M, Maloney WJ, Bragdon CR, O'Connor DO, Haire T, Harris WH. The initiation of failure in cemented femoral components of hip arthroplasties. *J Bone Joint Surg Br* 1991;73:551–8. [PubMed: 2071634]
29. Kawate K, Maloney WJ, Bragdon CR, Biggs SA, Jasty M, Harris WH. Importance of a thin cement mantle. Autopsy studies of eight hips. *Clin Orthop* 1998;70–6. [PubMed: 9917592]
30. Karrholm J, Herberts P, Hultmark P, Malchau H, Nivbrant B, Thanner J. Radiostereometry of hip prostheses. Review of methodology and clinical results. *Clinical Orthopaedics & Related Research* 1997;94–110. [PubMed: 9372762]
31. Cristofolini L, Teutonico AS, Savigni P, Erani P, Viceconti M. Preclinical assessment of the long-term endurance of cemented hip stems. Part 1: Effect of daily activities - A comparison of two load histories. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine* 2007;221:569–584.
32. Cristofolini L, Erani P, Savigni P, Bordini B, Vicenconti M. Preclinical assessment of the long-term endurance of cemented hip stems. Part 2: In-vitro and ex-vivo fatigue damage of the cement mantle. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine* 2007;221:585–599.
33. Britton JR, Prendergast PJ. Preclinical testing of femoral hip components: An experimental investigation with four prostheses. *Journal of Biomechanical Engineering* 2005;127:872–880. [PubMed: 16248319]
34. Stolk J, Maher SA, Verdonchot N, Prendergast PJ, Huiskes R. Can finite element models detect clinically inferior cemented hip implants? *Clinical Orthopaedics & Related Research* 2003;138–50. [PubMed: 12671496]
35. Barrack RL, D. MR 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;74:385–9. [PubMed: 1587883]
36. Karrholm J, Borssen B, Lowenhielm G, Snorrason F. Does early micromotion of femoral stem prostheses matter? 4–7-year stereoradiographic follow-up of 84 cemented prostheses. *Journal of Bone & Joint Surgery - British Volume* 1994;76:912–7. [PubMed: 7983118]
37. Sundfeldt M, Carlsson LV, Johansson CB, Thomsen P, Gretzer C. Aseptic loosening, not only a question of wear: A review of different theories. *Acta Orthopaedica* 2006;77:177–197. [PubMed: 16752278]
38. Aspenberg P, Van Der Vis H. Migration, particles, and fluid pressure: A discussion of causes of prosthetic loosening. *Clinical Orthopaedics and Related Research* 1998;75–80. [PubMed: 9678035]
39. V F, Cameron HU. The femoral stem/cement interface in total hip replacement. *Clin Orthop* 1976:248–52.
40. Crawford RW, Evans M, Ling RS, Murray DW. Fluid flow around model femoral components of differing surface finishes: In vitro investigations. *Acta Orthopaedica Scandinavica* 1999;70:589–595. [PubMed: 10665724]
41. Bartlett GE, Beard DJ, Murray DW, Gill HS. The femoral stem pump in cemented hip arthroplasty: An in vitro model. *Medical Engineering and Physics* 2008;30:1042–1048. [PubMed: 18280197]

Figure 1.
Loading jig and micro-motion measurement device attached to a right femur, which was wrapped with a heated water circuit.

Figure 2.

Charts showing changes in the position of the stem, relative the cortex at the level of the micro-motion measurement device, during the application of cyclic loading. “#m” = size # matte stems, “#p” = size # polished stems. Bars are 1s.d. Data are presented for subsidence (left) and retroversion (right), showing migration (top), the change in valley position, and cyclic motion (bottom), the difference between peak and valley positions. Statistical analyses of these data are shown in Table 1.

Figure 3.

Epifluorescence micrograph of a transverse section through a matte size 3 stem in a right femur. The section was taken 40mm distal to the calcar, at the level of the lesser trochanter. This stem, which completed $10^{5.5}$ cycles, migrated 1.3° into retroversion and subsided 0.11mm (0.7° & 0.06mm @ 10^5 cycles), illustrates several points of interest. In general, note the extensive interdigitation between cement and trabecular bone (which would be absent in a closed cell foam analogue). Also, note that, although this section would have shown a Grade A 'whiteout' on a radiograph and a mantle of over 2mm, trabecular bone almost touched the stem – this was common finding. (a) The stem had locked into a retroverted position, preventing an accurate assessment of the initial stem/cement contact fraction. (b) Typical microcracks in the interdigitated region (where they were most commonly found). Also, note the frequent gaps between cement and trabecular bone – although some of these could have been formed by loading, such gaps are frequently observed in unloaded specimens due to cement shrinkage. (c) An example of an open through crack. (d) A large void, with some associated cement damage. On its left, another large void had no cement damage. Such voids are not uncommon. (e) Oblique microcracks adjacent to the stem, suggesting torsional failure.

Figure 4.

Charts showing the relationship between stem related cement damage and cyclic subsidence (top) and cyclic retroversion (bottom) at the end of cyclic loading. Circles are polished, stars are matte, and pairs are matched by colour.

Cyclic retroversion (deg.)

Figure 5.

Chart showing the relationship between stem subsidence and cyclic retroversion over the course of 10^2 – 10^5 cycles of loading. Data points are at log-regular intervals of cyclic loading.

Table 1

Results of the repeated measures analysis of the micromotion outcome measures (Fig 2). Finish{*matte/polished*} and $\text{Log}_{10}(\text{Cycles} = 10^2, 10^{2.5}, 10^3, 10^{3.5}, 10^4, 10^{4.5} \text{ \& } 10^5)$ were treated as within-donors effects to define repeated measures for each donor. Size{#1|#3} was treated as a between-donors effect.

	p-values for micromotion outcomes (Fig 2)			
Effect	<i>Retroversion</i>	<i>Subsidence</i>	<i>Cyclic rotation</i>	<i>Cyclic translation</i>
Finish	0.060	0.005	0.20	0.21
Size	0.002	0.043	0.062	0.29
Finish*Size	0.11	0.026	0.059	0.14
Log10(Cycles)	<0.0001	<0.0001	0.16	0.011
Finish*Log10(Cycles)	<0.0001	0.94	<0.0001	0.74
Size*Log10(Cycles)	<0.0001	0.10	0.59	0.048
Finish*Size*Log10(Cycles)	0.81	0.39	0.39	0.038

Table 2

Results of the repeated measures analysis of the cement mantle damage outcome measures. Donor pairing was accounted for by treating Finish{*matte/polished*} as a within-donors effect to define repeated measures for each donor. Size{#1|#3} was treated as a between-subjects effect.

		Cement crack length-density (microns/mm ² ± s.d.)		
		Broached area		Interdigitated area
		<i>Stem</i>	<i>Bone</i>	
Matte	<i>size 1</i>	24 ± 13	10 ± 5	116 ± 62
	<i>size 3</i>	17 ± 16	15 ± 8	142 ± 21
Polished	<i>size 1</i>	49 ± 28	13 ± 12	117 ± 70
	<i>size 3</i>	5 ± 2	10 ± 4	113 ± 28
		p-values for damage measures		
Effect	Size	0.12	0.87	0.80
	Finish	0.35	0.62	0.33
	Finish*Size	0.042	0.28	0.29