Hothi, H, Matthies, Ashley K, Berber, R, Whittaker, R, Bills, Paul J., Racasan, Radu, Blunt, Liam, Blunn, Gordon, Skinner, John and Hart, A. J.

The reproducibility of a semi-quantitative scoring method for taper corrosion and fretting, and its usefulness for predicting the volume of material loss

Original Citation

Hothi, H, Matthies, Ashley K, Berber, R, Whittaker, R, Bills, Paul J., Racasan, Radu, Blunt, Liam, Blunn, Gordon, Skinner, John and Hart, A. J. (2014) The reproducibility of a semi-quantitative scoring method for taper corrosion and fretting, and its usefulness for predicting the volume of material loss. In: Orthopaedic Research Society - 60th Annual Meeting, March 15-18 2014, New Orleans, LA, USA. (Unpublished)

Title: The reproducibility of a semi-quantitative scoring method for taper corrosion and fretting, and its usefulness for predicting the volume of material loss

Authors: H Hothi, AK Matthies, R Berber, R Whittaker, P Bills, L Blunt, R Racasan, G Blunn, J Skinner, A Hart

Introduction

It has been suggested that the release of metal debris due to mechanical wear, fretting and corrosion at the head-stem taper junction may contribute to the high revision rates of MOM-THAs [1, 2].

A peer reviewed, semi-quantitative corrosion and fretting scoring system of the femoral stem and head tapers using visual assessment, was developed by Goldberg et al. [3]. This method involves assigning scores on a scale of 1 (no visible signs of corrosion or fretting) to 4 (severe corrosion or fretting) based on the prevalence of black debris, pits or etch marks (corrosion) and fretting scars.

However, this method has not been validated and the reproducibility of such a measure is unknown. Furthermore, whilst scoring a component allows the examiner to quickly quantify the appearance of taper surface damage in the form of corrosion or fretting, it is unclear as to how this relates to the actual volume of material loss at the taper.

The aims of this study were to: (1) determine the inter-observer variability of the visual scoring method for the assessment of corrosion and fretting of MOM-THA taper junctions; (2) determine the prevalence and severity of corrosion and fretting at the taper junction; (3) determine the strength of correlation between corrosion and fretting scores and the actual volume of material lost at the taper junction.

Methods

Macroscopic and stereomicroscopic examinations of the head taper surface of 150 retrieved MOM-THA implants were performed by two experienced independent observers using methods previously defined [3] to assess corrosion and fretting. Both observers were blinded to all other clinical and component data. A Leica M50 microscope with x40 magnification was used to examine each surface. The visual appearance of fretting was considered as regions of the taper surface that were damaged with small scars running perpendicular to the circumferential machine lines of the taper screw thread. The visual appearance of corrosion was defined as regions of discolouration or dullness on the taper surface or the presence of black debris, pits or etch marks. The prevalence of corrosion and fretting was quantified using a scoring scale of 1 (none) to 4 (severe), as described previously [3]. Corrosion and fretting scores were assigned to the distal and proximal regions of the taper surface, Figure 1. Overall scores for corrosion and fretting were then assigned to each taper following assessment of each surface as a whole.

The volume of material loss at each of the taper surfaces was measured using a Talyrond 365 (Hobson, Leicester, UK) roundness measuring machine, using a published method [4]. The corrosion score relating to the overall surface of each taper was plotted against the measured volume of material loss.

All statistical analysis was performed using Stata/IC version 12.1 [StataCorp, College Station, TC, USA] and throughout a p value < 0.05 was considered statistically significant.

Cohen's weighted Kappa statistic (κ) was used to measure the inter-observer agreement of the Goldberg scores [5]. Kappa values were assessed using the criteria described by Landis and Koch where $\kappa \leq 0$ = poor, 0.01 to 0.20 = slight, 0.21 to 0.40 = fair, 0.41 to 0.60 = moderate, 0.61 to 0.80 = substantial, 0.81 to 1 = almost perfect [6].

Neither the volumetric measurement data nor corrosion scores were normally distributed. Therefore, to determine the strength of correlation between the two, we used the Spearman Rank test.

Results

Table 1 summarises the inter-observer agreement data for the Goldberg scores of the taper surfaces. The observed agreement for the overall corrosion and fretting scores were 95% and 84% respectively. The reliability of the proximal and distal taper corrosion scores was moderate to substantial (κ=0.52 to 0.70), whilst reliability of the overall head taper corrosion score was substantial (κ=0.64). The reliability of the proximal, distal and overall taper fretting scores was poor (κ=0.14, 0.13 and 0.18 respectively).

Figures 2a and 2b plot the distribution of the overall corrosion and fretting scores by both examiners for the tapers examined. Examiners A and B reported that 99% and 94% of tapers respectively had visual evidence of corrosion whilst 54% and 59% of tapers respectively showed evidence of fretting.

Figure 3 plots the severity of corrosion assessed using the visual scoring method against the material loss at the head taper. Taper corrosion score was significantly and moderately correlated with the volume of material loss measured (Spearman's $r = 0.59$; 95%CI = 0.47 - 0.68; $p < 0.001$). Similarly, fretting score was significantly correlated with the volume of material loss, but the correlation was weak (Spearman's $r = 0.24$; 95% CI = 0.08-0.39; $p = 0.003$).

Discussion

The results of this study found that the observed agreement was higher for corrosion assessments of the head taper surface than for fretting. The inter-observer reliability for corrosion scoring was moderate whilst the reliability of fretting scores was slight to fair.

Both examiners reported that it was considerably easier to identify regions of discolouration and the presence of black debris, both key indicators of corrosion, than it was to identify clear fretting scars. It was difficult to distinguish between surface damage due to fretting and damage that may have occurred during implantation or retrieval of the components.

In all instances of fretting, there was also evidence of corrosion however there was not always evidence of fretting when corrosion was observed, particularly when corrosion scores were severe. This supports the claim that lower fretting scores may be due to the fact that the black deposits, due to corrosion, mask the presence of fretting scars. Another explanation may be that the occurrence of crevice corrosion over time may remove some of the fretting scars, therefore underestimating the degree of fretting that has occurred.

The presence of black corrosion deposits may also explain the large variability between the two examiners when reporting on fretting. There were a number of examples of when deposited material partially obscured suspected fretting scars, adding to the existing uncertainty due to difficulties in distinguishing between fretting and other mechanical damage.

Our study is the first to report on the relationship between the visual assessment of the taper and quantified material loss. Significant positive correlations were found to exist between head taper corrosion and fretting scores and the actual measured volume of material loss. The correlation was stronger for corrosion scores than fretting scores. Whilst these results certainly support a mechanism involving fretting corrosion, the strengths of these correlations do not support the use of semi-quantitative scoring methods to predict the actual measured volume of material loss.

The results of our study suggest that corrosion plays a significant role in the loss of material: (1) there was evidence of corrosion in almost every taper examined; (2) a positive correlation was observed between the severity of corrosion and the absolute volume of material loss; (3) imprinting of the male taper surface was observed on all female tapers, and suggests a mechanism involving galvanic corrosion; (4) observations of less fretting may in part have been due to crevice corrosion destroying fretting scars over time.

Significance

Our study has contributed to the work on implant retrievals, showing that detailed visual examination of taper surfaces can produce reliable data, which may be able to predict the severity of material loss but is not a substitute for complex metrology methods. The results of our study suggest that corrosion may be the main mechanism of material loss at the taper junction.

References

[1] Meyer et al. [2012], [2] Matthies et al. [2012], [3] Goldberg et al. [2002], [4] Matthies et al. [2013], [5] Cohen [1968], [6] Landis and Koch [1977]

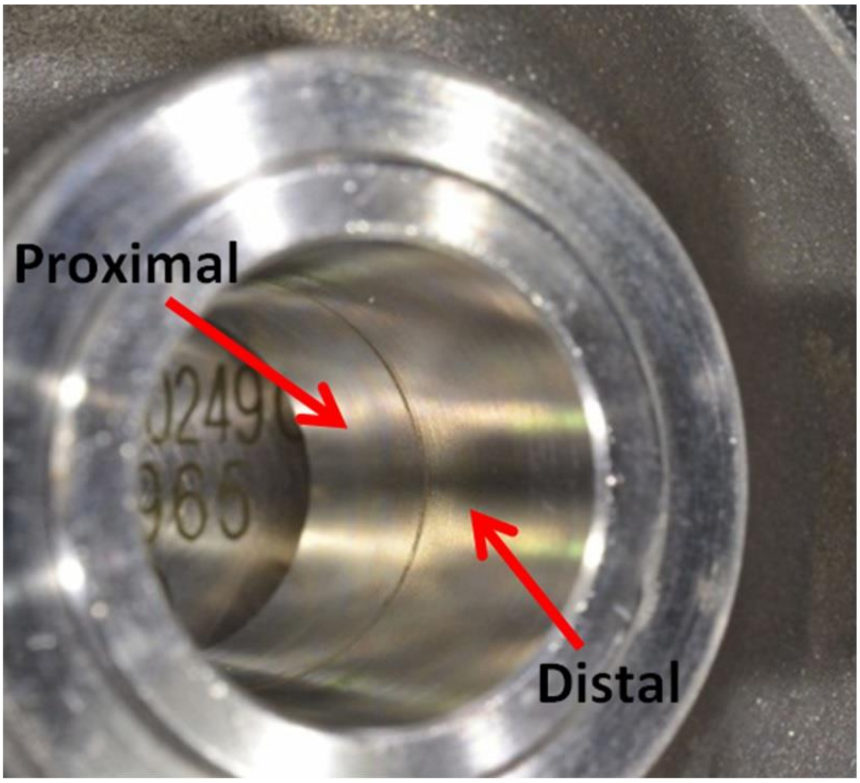

Figure 1: *Corrosion and fretting scores assigned to proximal and distal regions of taper*

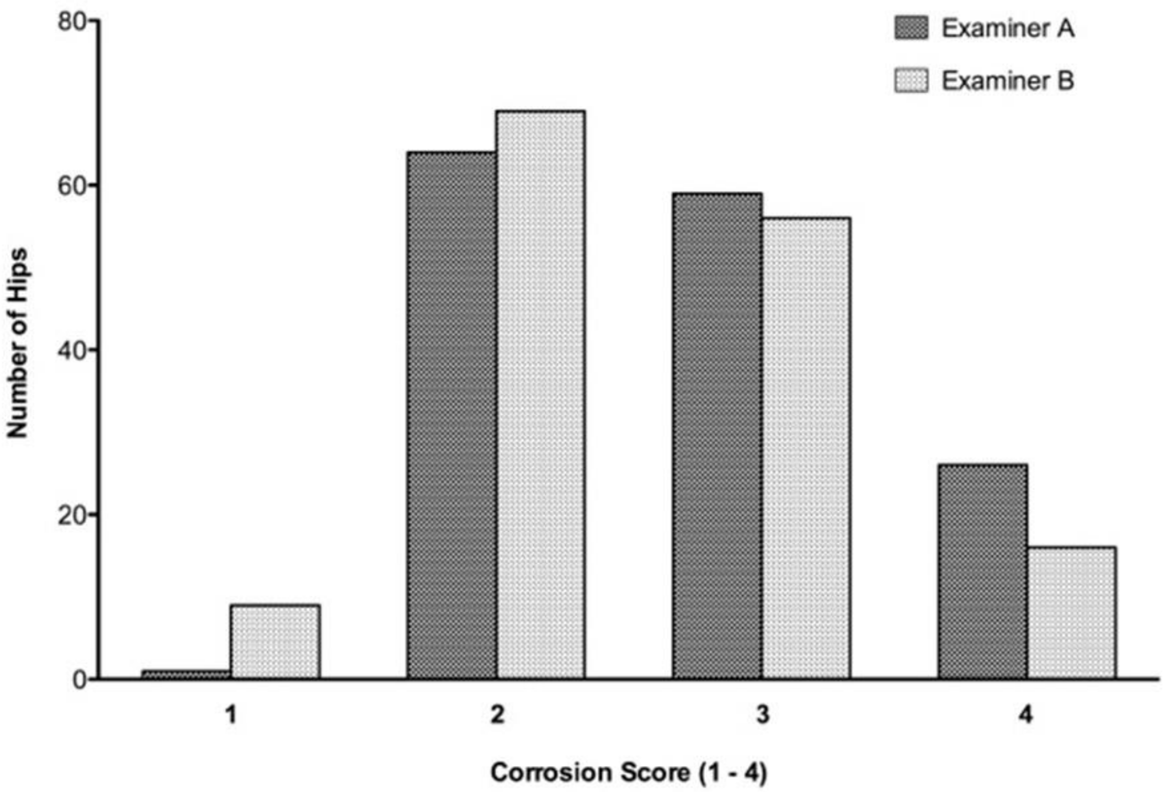

Figure 2a: *Distribution of overall corrosion scores for both examiners*

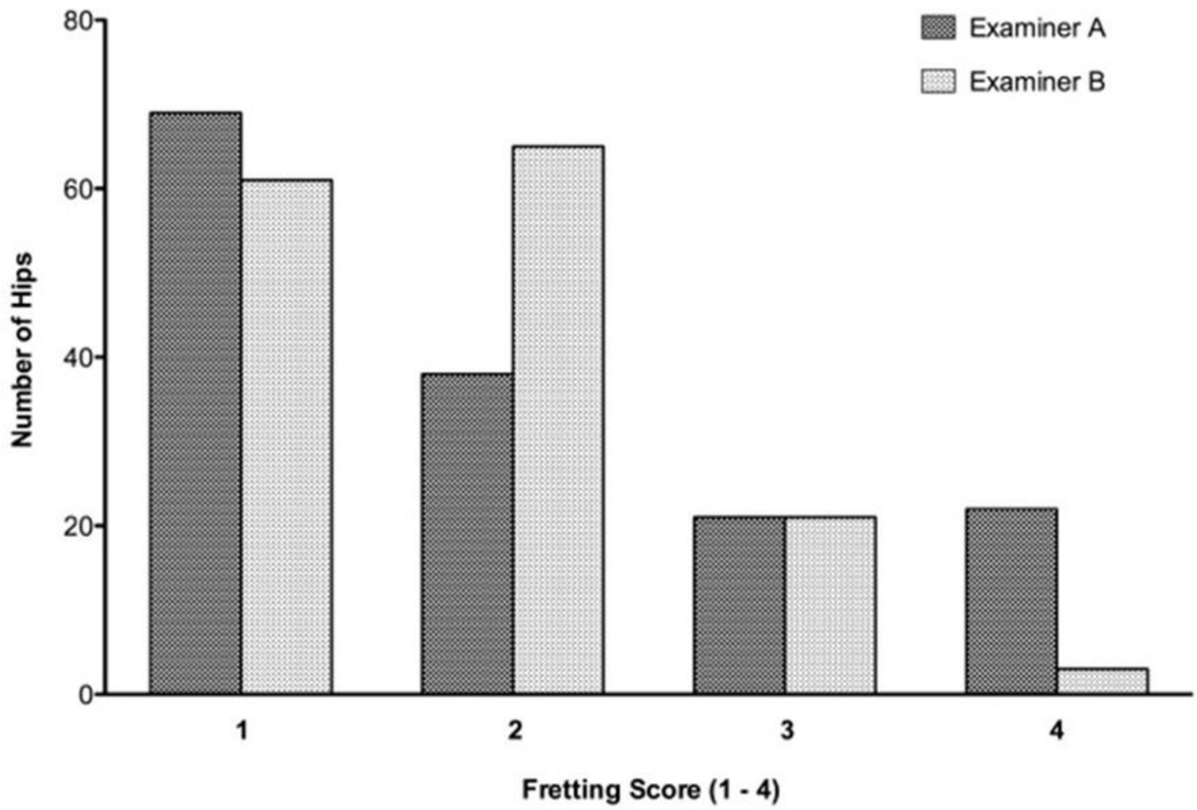

Figure 2b: *Distribution of overall fretting scores for both examiners*

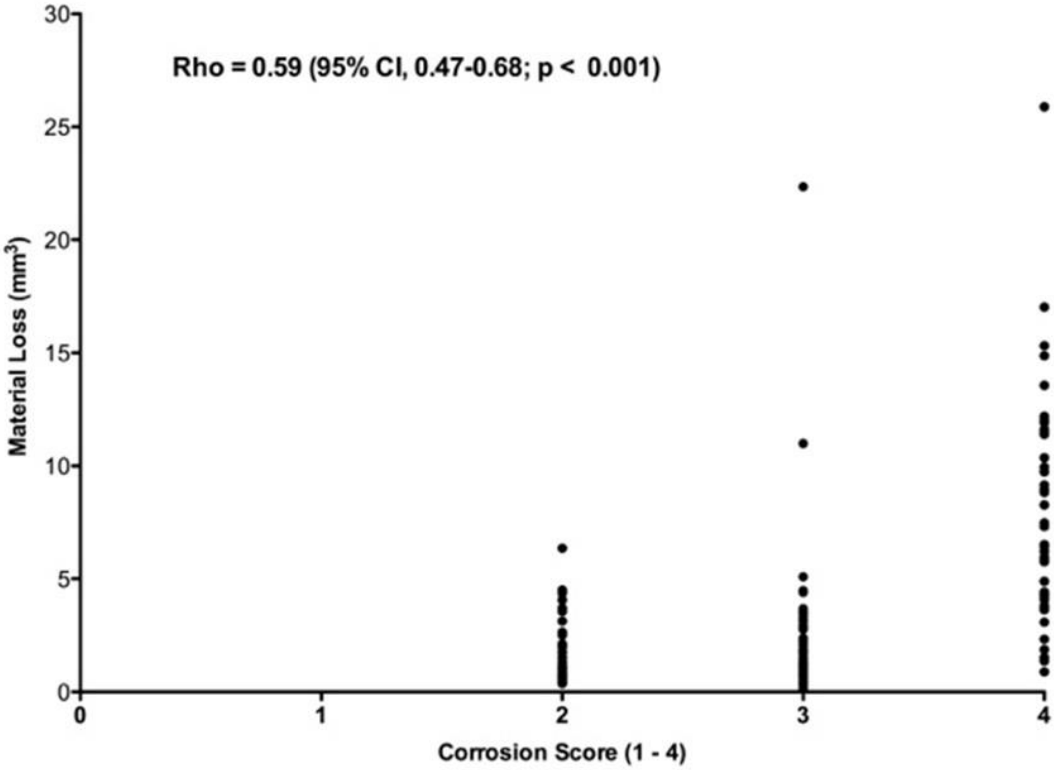

Figure 3: *Corrosion scores for examiner A plotted against measured material loss*