

Inter-rater and Inter-institutional Reliability of Implant Wear Analysis Techniques

Stephanie M. McCarthy, Douglas W. Van Citters, Audrey J. Martin, Joshua J. Jacobs, Brian J. McGrory, Robin Pourzal
Rush University Medical Center, Chicago, IL, USA, Dartmouth College, Hanover, NH, USA, Tufts University School of Medicine, Boston, MA, USA
Stephanie_McCarthy@rush.edu

Disclosures: SM McCarthy (N), D Van Citters (5-DePuy Synthes, Medacta, RevBio, TJO), AJ Martin (N), JJ Jacobs (3B-Smith & Nephew, Zimmer-Biomet, 4-Hyalex, 5-Medtronic Sofamor Danek, Nuvasive, Zimmer-Biomet, 7B-JBJS, Am., 8-JBJS, Am., 9-ABOS Inc., OREF), BJ McGrory (3B-Smith & Nephew), R Pourzal (5-Stryker, Zimmer Biomet; 8-Biotribology)

INTRODUCTION: Fretting wear and corrosion within modular taper junctions of total hip arthroplasty (THA) components result in the release of metal particles and ions into the surrounding periprosthetic tissues. Adverse local tissue reactions (ALTR), painful pseudotumor-like reactions surrounding the implant, have been linked to these metallic debris, specifically debris from the taper of cobalt-chromium-molybdenum (CoCrMo) alloy femoral heads, and often require revision surgery [1]. Understanding the process and extent of material loss from the modular taper is imperative to reduce ALTR development and improve THA patient outcomes. The proper assessment of material loss from femoral head and stem tapers is essential to understanding the downstream results of local and systemic inflammation attributed to debris release. Coordinate measuring machines (CMM) and out-of-roundness machines are commonly used to measure volumetric material loss. Standards for the measurement and estimation of modular taper material loss from CMM have been defined by the ASTM for both tactile and optical techniques [2]. Even with defined standards, material loss assessment between raters using the same technique in the same lab or between labs using different techniques are not commonly compared. The purpose of this study was to evaluate 1) how volumetric material loss analysis compares between two observers from the same institution from the same data set, 2) how volumetric material loss analysis compares between two observers from different institutions using separate CMM data sets of material loss of the same femoral heads, and 3) how the resulting volumetric material loss measurements show trends with patient outcome data, specifically time *in situ* and cobalt serum metal ion concentrations.

METHODS: Analysis was conducted on 19 CoCrMo modular femoral heads retrieved from an institution independent of both participating laboratories (Lab A and B). The median (range) time *in situ* was 75.3 months (22.3-109.3). Serum cobalt ion levels were available for 6 patients. Metal ion data was provided by a third lab, independent of the two participating CMM labs. Lab A scored heads using a modified version of the Goldberg visual-based scoring system (1 minimal, 2 mild, 3 moderate, 4 severe), and presence of damage modes including 'imprinting' and 'column damage' were noted [3]. Head taper molds were created using a high-resolution replicating compound (Microset 202) and measured using an optical CMM (OrthoLux, RedLux) with a white light confocal sensor and commercially available software (RedLux Studio) [4]. A senior investigator (SI) and trainee (TI) used the same CMM data to compute volumetric material loss by fitting an ideal cone geometry to undamaged regions visible on the taper mold using a light intensity map. According to ASTM, head tapers were classified as Type 1- displaying undamaged regions on both distal and proximal regions, Type 2- undamaged regions on either the proximal or distal surface, or Type 3- displaying no areas of as-manufactured surface, indicating full contact with the stem taper along the entirety of the head taper area [2]. Lab B measured the same femoral heads directly with a tactile CMM (Zeiss Contura G2). Heads were secured at four points around the equator. A 3mm ruby stylus was controlled using a custom measurement script in which at least 72 axial scans were taken along the complete length of the bore with measurements every 0.1 mm along the scan. A supervised custom fitting algorithm digitally reconstructed the original taper surface from point clouds and computed volumetric material loss. Inter-rater and inter-laboratory reliability of modular junction damage assessment was tested using the intraclass correlation coefficient (ICC). SPSS 22 software was used to obtain ICC estimates and 95% confidence intervals (CI) based on a single measurement, consistency, two-way mixed effects model. ICC results were defined as ≤ 0.5 equating to poor agreement, 0.5-0.75 moderate agreement, 0.75-0.9 good agreement, and ≥ 0.9 indicating excellent agreement [5]. Kruskal-Wallis and linear regression were also run to examine material loss values and correlations with blood metal ion findings. Significance was set at $p < 0.05$.

RESULTS: Damage assessment of the head tapers found 7 with chemically-induced column damage and 14 with stem topography imprinting. The distribution of damage scores was: minimal damage ($n=1$), mild damage ($n=2$), moderate damage ($n=0$), and severe damage ($n=16$). At Lab A, 16 tapers were classified as either Type 1 or 2 tapers, and 3 were classified as Type 3. Material loss from Type 3 tapers could not be qualitatively assessed in two cases due to the lack of sufficient reference area. Among the measurable cases, the median (range) material loss was 4.78mm^3 (0.0-11.73) and 4.38mm^3 (0.0-15.45) for the SI and TI, respectively. At Lab B, the median volumetric material loss was 2.25mm^3 (-.44-17.3). A Kruskal-Wallis test showed no significant difference in volumetric material loss values between observers ($p=0.462$). The ICC results (95% CI) for volumetric material loss between all three reviewers and two techniques was $\text{ICC}=0.93$ (0.84-0.97), demonstrating good to excellent agreement between reviewers and institutions. Direct correlation of material loss measured for all cases at Lab A by SI and TI resulted in good agreement ($R^2=0.87$, $p<0.001$), as did cases measured by Lab A SI and Lab B ($R^2=0.73$, $p<0.001$). Considering patient outcomes, Lab A SI and TI and Lab B demonstrated a significant positive correlation between material loss and time *in situ* ($R^2=0.61$, $p=0.001$, $R^2=0.54$, $p=0.002$, $R^2=0.29$, $p=0.027$, respectively). Among the 6 cases with available Co ions, a trend towards a positive correlation with Co ion data was observed: Lab A SI: $R^2=0.61$, $p=0.066$, Lab A TI: $R^2=0.75$, $p=0.026$, Lab B: $R^2=0.31$, $p=0.25$.

DISCUSSION: We have demonstrated excellent agreement in the reliability of CMM data wear analysis methods between three different raters at two institutions using two different CMM techniques. These findings demonstrate that material loss can be accurately assessed using both tactile and optical CMMs. Furthermore, the results of Lab A demonstrate that a less experienced trainee will come to the same conclusions as an experienced investigator. Both techniques have advantages and disadvantages. Tactile CMMs allow for the direct measurement on the femoral head surface, while the optical CMM relies on the use of a mold. However, the light intensity map generated as a byproduct of the optical CMM scans aid the rater to better determine undamaged areas and makes the assessment of reference areas in severely damaged tapers easier. Regardless, the excellent ICC result indicates that both techniques are valid, which is also reflected in the common conclusion of the positive relationship between material loss and time *in situ*. We also observed a trend towards a relationship between Co ion levels and volumetric loss from the head taper in MoP THA. This finding was significant for Lab A TI, but not the others. However, the non-significance appears to be driven by a single outlier thus showcasing the largest limitation of this study, which is cohort size.

SIGNIFICANCE: While the (modified) Goldberg Score has long been established in the field as the primary assessment tool for taper damage, it has many limitations depending on the study objective. There remains a need in the field to establish reliable volumetric material loss data in order to determine if ALTRs caused by wear debris and corrosion products from modular junctions are dose dependent. Considering multiple confounding factors impacting material loss, a single laboratory will not be able to establish sufficiently large cohorts. It is therefore imperative to ensure that published data is comparable. This study marks a first step in this direction. Additionally, this study is the first to suggest a potential correlation between material loss at the taper and blood Co ions levels, which needs to be further explored.

REFERENCES: [1] Cooper et al, JBJS-A, 2013, 95: 865-72, [2] ASTM Standard F3129-16, 2016, [3] Hall et al, JBJS-B, 2018, 106: 1672-85, [4] McCarthy et al, CORR, 2021, 479: 2083-96, [5] Koo & Li, J. Chiropr. Med., 2016, 15:155-63.

ACKNOWLEDGEMENTS: This study was funded by NIH grant R56 AR070181-06A1.