Introduction: Multicenter studies in which different examiners will be measuring polyethylene wear from radiographs to assess conventional polyethylene wear and particularly to evaluate the new crosslinked polyethylenes will want to consider using the most accurate technique to reduce measurement errors. The most precise in vivo method of quantifying wear thus far reported is radiostereometry (RSA) [1], the gold standard at this time. However, the widespread use of this technique is limited in that it requires expensive specialized equipment and a complicated analysis. Over two decades ago, investigators introduced simple 2D techniques for measuring femoral head penetration from a single radiograph [2], with the technique of Livermore [3] being popularized. Computerized edge-detection schemes were then developed to improve upon the repeatability of the 2D caliper measurements by introducing either a digitizing tablet or film scanner [4]. More recently, 3D analyses by Devane et al. [5] and Martell et al. [6] incorporating both AP and lateral films, in addition to more complete volume wear formulas, have been introduced to provide an accurate assessment of the linear penetration and volumetric wear rate.

The purpose of this study was to measure the accuracy and repeatability of the contemporary 3D computerized techniques of Devane and of Martell and compares those with measurements obtained with the more familiar 2D technique of Livermore. Observer variation for a range of operator experience was investigated with a controlled in vitro wear phantom profile, and a prospective clinical study on the Harris-Galante prosthesis (Zimmer, Warsaw, IN) acetabular reconstruction with an average duration of 10.3 years (range=5.2-13.2 years) was completed to evaluate clinical performance.

Materials and Methods: Wear phantoms were created by machine milling 6 Trilogy (Zimmer Corp., Warsaw, IN) acetabular cups (diameter=62 mm) with a spherical mill to simulate linear femoral head penetration (diameter=32 mm) to wear depths of 0.5, 1.0, 1.5, 2.0, and 4.0 mm. The wear path was oriented along a vector directed towards a common area that is subjected to maximum wear in THA, in the superior and lateral aspect of the cup. AP and true lateral plain radiographs were taken with the prosthetic components positioned at 45° anteversion and 20° tilt in a custom-built jig which incorporated industrial elastic bands to seat the femoral head into the cup. Each cup was probed at 1013 specific points on its articulating surface using a coordinate measuring machine (CMM) accurate to 80μm (MicroVal PFx, Browne & Sharpe, RI) before and after milling to report the wear depth and volume achieved with machining. Radiographs were digitized at 300 dpi using a film scanner (VXR-12 Plus™, Vidar Systems Corp, Herndon, VA) at 200 dpi with sixteen byte encoding for computerized analysis. One observer evaluated the wear depth and volume for each phantom using the technique of Livermore with a Vernier caliper to measure the dimensions of the head evaluated the wear depth and volume for each phantom using the technique of Livermore (-18±228% [-51, 534], and Reader 3=-14±43% [-83, 494] (ranked in increasing years experience). Readers tended to underpredict the actual wear (i.e. negative error). Mean accuracy for repeat readings was 12%.

Discussion: There was a considerable amount of intra- and interobserver variability in the degree of accuracy, and repeatability for each reader in contrast to the reported performance reported. This was primarily due to how different readers interpreted the location of the cup apex, and points on the ellipse. It is estimated that film quality did not significantly impact readings, nor did technique programmalitics. However the ability to clearly visualize the cup opening for different geometries was essential. Each step was methodologically correct to the best of the readers’ knowledge as confirmed by interactive training sessions with the authors. Towards this regard, error of this magnitude may be significant for the interpretation of data on a case by case basis; however, interpretation of previously published clinical series depends upon the averaging factor. Use of multiple readers helps to minimize the effect of this variability.