Introduction
In recent years, a number of competing bearing surface materials have been introduced for total hip replacements. While these materials show great promise based on laboratory assessment, only long-term patient follow-up studies can determine their true success or failure and distinguish among them. Unfortunately, clinical results may take many years to show differences between bearing surfaces, and even longer to determine why failure occurred. Therefore, it is imperative to develop and identify techniques capable of quantifying small amounts of wear at an early stage.

The goal in this study was first to independently assess the accuracy of available methods for measuring polyethylene wear from radiographs. We then compared maximum wear depth, measured from retrieval liners using CMM, to wear measured from clinical follow-up radiographs of the same patients using three manual and two computerized edge-detection techniques.

Methods
An apparatus was used to create and take AP and lateral radiographs of 116 random values of 3D wear in the laboratory using an HG-I shell and a 28mm ball, combining 20°, 45°, 55° of abduction and 0°, 10°, 20° of anteversion. Methods included were Devane’s PolyWare™ and Martell’s Hip32™ software packages, and Charnley’s duo-radiographic and Livermore’s methods. In addition to the above methods, a triangle method established by the authors was evaluated.

Using follow-up x-Rays of seventeen THR patients, the same radiographic methods were again used. These measurements were then compared to measurements taken directly from the acetabular liner retrieved from the same patients during revision surgery. Error (discrepancy) was calculated by subtracting the radiographic measurements from corresponding CMM measurements. Mean and maximum errors and associated confidence intervals and standard deviations were calculated for each method.

Results
In the laboratory setting, for 2D analysis, Livermore’s method was the most accurate; for 3D, computerized methods were equivalent (fig 1). The error was not significantly affected by the amount of wear (although the percentage error increased with smaller wear), or by the cup position.

Discussion
Although computerized methods outperformed manual methods in intra-observer and inter-observer reproducibility, both packages were difficult to learn. Cost of the software, dedicated computer, scanner and operator labor must be considered. For example, the edge-detection routines did not always work, and manual intervention was needed. Experience with the software was necessary to correct the edge-detection in these cases. For 3D measurements, good quality cross-lateral x-rays are necessary. Thus, the Livermore method, although limited to 2D projection of a 3D vector, can nevertheless be still considered the gold standard for routine clinical practice or research.

Previous studies that have assessed the accuracy of computerized techniques have done so using phantoms and ideal laboratory settings, including flawless lateral x-rays. Nominal accuracy of 0.15 is often reported. This is in agreement with our laboratory studies of 2D and 3D manual and computerized techniques using phantoms, which assessed 0.14-0.17mm accuracy.

Despite increasing popularity of computerized edge-detection methods for measuring THR polyethylene wear, few studies have independently validated these new methods in a clinical setting. Ohlin (J Arthroplasty, 8-4, 1993) found 1.1mm average and 1.9mm maximum error comparing coordinate measurement machine (CMM) measurements of retrieved cups to a computerized radiographic measurement method without automatic edge-detection.

The current study indicated that, in a clinical setting, achieving better than 0.3mm accuracy in measuring polyethylene wear may not be realistic. More importantly, manual methods, in particular, Dorr’s method, which is easy to use and can be performed on poor quality x-rays or obscure femoral ball margins, performed nearly as well as computerized methods.