Validation of a New Method for Determination of Cup Orientation in THA

Simon D. Steppacher, 1,2 Moritz Tannast, 2 Guoyan Zheng, 3 Xuan Zhang, 2 Jens Kowal, 1,3 Suzanne E. Anderson, 4 Klaus A. Siebenrock, 2 Stephen B. Murphy 1

1 Center for Computer Assisted and Reconstructive Surgery, New England Baptist Hospital, Tufts University School of Medicine, Boston, Massachusetts, 2 Department of Orthopaedic Surgery, Inselspital, University of Bern, Murtenstrasse, 3010 Bern, Switzerland, 3 ARTORG Center for Biomedical Engineering Research, ISTB—Institute for Surgical Technology and Biomechanics, University of Bern, Bern, Switzerland, 4 Department of Radiology, Inselspital, University of Bern, Bern, Switzerland

Received 6 December 2008; accepted 4 May 2009
Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/jor.20929

ABSTRACT: Our goal was to validate accuracy, consistency, and reproducibility/reliability of a new method for determining cup orientation in total hip arthroplasty (THA). This method allows matching the 3D-model from CT images or slices with the projected pelvis on an anteroposterior pelvic radiograph using a fully automated registration procedure. Cup orientation (inclination and anteversion) is calculated relative to the anterior pelvic plane, corrected for individual malposition of the pelvis during radiograph acquisition. Measurements on blinded and randomized radiographs of 80 cadaver and 327 patient hips were investigated. The method showed a mean accuracy of 0.7 ± 1.7° (−3.7° to 4.0°) for inclination and 1.2 ± 2.4° (−5.3° to 5.6°) for anteversion in the cadaver trials and 1.7 ± 1.7° (−4.6° to 5.5°) for inclination and 0.9 ± 2.8° (−5.2° to 5.7°) for anteversion in the clinical data when compared to CT-based measurements. No systematic errors in accuracy were detected with the Bland–Altman analysis. The software consistency and the reproducibility/reliability were very good. This software is an accurate, consistent, reliable, and reproducible method to measure cup orientation in THA using a sophisticated 2D/3D-matching technique. Its robust and accurate matching algorithm can be expanded to statistical models. © 2009 Orthopaedic Research Society. Published by Wiley Periodicals, Inc. J Orthop Res

Keywords: hip arthroplasty; cup orientation; 2D/3D-matching; inclination; anteversion

Both short- and long-term results after total hip arthroplasty (THA) are correlated with correct component positioning, particularly optimal orientation of the acetabular cup. Cup malorientation is associated with prosthetic impingement, dislocation, wear, osteolysis, loosening, and irritation of the iliopsoas tendon often resulting in early revision surgery. 1–5 The ability to measure the position of the acetabular component accurately is essential, both to correlate component positioning with outcome and to assess potentially improved methods of cup placement during surgery. 6–8 Unfortunately, measurement of component orientation on plain radiographs is highly inaccurate because of the individual pelvic orientation during radiograph acquisition. 9–11 Improved methods of measuring component position postoperatively are essential to evaluate surgical techniques and navigation methods, and to improve our ability to determine optimal component positioning in groups or in individual patients.

The current study assessed the accuracy of measuring acetabular component positioning on plain radiographs in patients upon whom a CT study was available. This method of 2D/3D matching allows for calculation of component position of any near-anteroposterior (AP) radiograph taken following surgery. We hypothesized that the 2D/3D-matching technique is an accurate, consistent, reliable, and reproducible tool for measurement of postoperative cup orientation in patients with CT-based computer-assisted THA.

MATERIALS AND METHODS
Software “HipMatch”
The mathematical background of the noncommercial software “HipMatch,” details of the implemented algorithm, and a preliminary evaluation were previously reported. 12–14 Briefly, the software uses a fully automated registration procedure that can match the 3D-model of a CT with the actual projected pelvis on a postoperative radiograph (Fig. 1). This is done by matching the contours of digitally reconstructed radiographs of the 3D pelvic model to the radiograph (Fig. 2). Required software input includes the preoperative CT stored in DICOM format, a digital radiograph, and the film-focus distance. No specific calibration of the x-ray or a CAD (computer-assisted design) model of the implant is required. The software then calculates cup radiographic inclination and radiographic version relative to the anterior pelvic plane. 15 Since the method for calculating cup version according to Ackland et al. 16 was used, the software is applicable to all bearing types (even those where the minor axis of the projected ellipse of the cup opening surface is hidden; Fig. 3). The only precondition of the implant is a circular opening surface. If any doubt exists about cup retroversion, the method of Seradge et al. 17 using a regular AP pelvic radiograph and a radiograph with the central x-ray beam centered on the cup can be used. By moving the central beam away from the cup, the apparent opening of the cup will decrease if the cup is anteverted. If the cup is retroverted, it will appear to open more.

Validation
The validation study consisted of external validation (1), internal validation (2), and reproducibility and reliability analysis (3). The validation study was based on cadaver trials and clinical data.

External validation was designed to measure software accuracy by direct comparison with the cup orientation...
measured on CT scans, which was considered the gold standard. For the cadaver trials, 10 dry human pelves (20 hips) with a bilaterally implanted prosthetic polyethylene liner were used. The liners were marked with a circular metal wire. The cup orientation measured on postoperative CTs (with the technique described later) ranged for the inclination from 34° to 57° (mean, 44.0 ± 6.0°) and for the anteversion from 1° to 24° (mean, 14.5 ± 6.8°). From each pelvis, one CT including the liners for ground truth measurement and one CT without the liners for the 2D/3D-matching process were performed. The CTs were acquired with a pixel spacing of 0.74 mm/pixel and an inter-slice distance of 2 mm. Additionally, four AP pelvic radiographs with unknown, different orientations of the pelvis during acquisition were performed, resulting in 80 measurements for accuracy. Pelvic orientation can drastically affect the apparent cup orientation on plain radiographs.9–11 Pelvic rotation along the sagittal axis can be controlled on standardized radiographs and therefore measurement of cup inclination on plain radiographs is more reliable.10 However, pelvic tilt (rotation along the horizontal axis) shows a large individual variation and therefore makes the measurement of cup version on plain radiographs highly inaccurate.11 The mean pelvic tilt measured on the radiographs (positive values for forward tilt) was 11.4 ± 15.7° (−28 to 42°) and the mean rotation (positive values for rotation to the right) was 20.2 ± 20.3° (−48 to 67°).

For the clinical measurements, all patients with a pre- and postoperative (24 patients, 25 hips) or with only a postoperative CT (44 patients, 48 hips) out of a series of CT-based computer-assisted THAs were used. The preoperative CTs were performed for operative planning, and postoperative CTs were available when a THA without CT-based planning on the contralateral side was performed previously or a new CT for the contralateral CT-based THA was acquired. The postoperative CT was used as the ground truth, and if available the preoperative CT was used for the 2D/3D-matching process. If only a postoperative CT was available, the postoperative CT after manual excision of the cup from every CT slice with Osirix (Version 3.2, Geneva, Switzerland)18 was used for the 2D/3D-matching process. The pre- and postoperative CTs had a variable pixel spacing ranging from 0.55 to 0.78 mm/pixel with a constant inter-slice distance of 2 mm. The cup orientation measured on postoperative CTs ranged for the inclination from 28° to 61° (mean, 42.7 ± 6.7°) and for the anteversion from 1° to 35° (mean, 22.6 ± 8.0°). The mean pelvic tilt of the anterior pelvic plane (APP) on the postoperative CTs was 3.8 ± 5.0° (−7° to 18°) and the mean rotation was −1.5 ± 2.4° (−10° to 7°). One to six postoperative radiographs were available from each patient, resulting in a 327 measurements. AP pelvic radiographs were performed supine. Radiographs with gonadal shielding (47 radiographs), deep centering (88 radiographs), or a combination of both (12 radiographs) were assessed separately from the radiographs without shielding or deep centering (192 radiographs) to detect a possible influence on the 2D/3D matching accuracy. Deep centering was defined as missing anterior superior iliac spines and the cranial part of the pelvis on the radiograph.
Ground truth measurements of cup orientation based on CT relative to the APP were done using commercially available Osirix software. The four anatomical landmarks (both anterior superior iliac spines and pubic tubercles) of the APP were defined on a rendered 3D-CT pelvis model by one observer (S. D. S.). This was very reliable and reproducible. Then, six points on the cup-opening surface were used to define a best fit circle. The corresponding normal vector and the real radiographic anteversion and inclination were calculated.

Differences in inclination and anteversion compared to the ground truth after 2D/3D-matching were calculated. The graphical Bland–Altman analysis was used to detect systematical errors in the software.

Internal validation was used to evaluate the consistency of the software. The 40 radiographs of the 10 cadaver pelvises (20 hips) were used. Each of the hips had four AP pelvic radiographs with the pelvis in different, unknown orientations during acquisition. From each radiograph, the cup position was calculated and compared with the three corresponding results of the same hip. Despite the different pelvic orientation during radiograph acquisition, the results should not vary widely indicating a high consistency of the software (Fig. 4). The results for inclination and anteversion were graded using the intraobserver correlation coefficient (ICC).

Reproducibility and reliability analysis was used to evaluate the variability between two measurements or two observers and was based on the cadaver trials. All the radiographs of the 80 cadaver hips were blinded, randomized, and analyzed by two observers at two occasions, at least 1 month apart, resulting in 320 measurements. The results for inclination and anteversion were graded using the ICC.

Statistical Analysis
Normal distribution was determined using the Kolmogorov–Smirnov test. Because not all parameters showed normal distribution, non-parametric tests were used. Comparison of accuracy of inclination and anteversion between cadaver and clinical tests or between radiographs with or without gonadal shielding, with or without deep centering, or a combination of both was assessed using the Mann–Whitney U-test. To detect systematical errors in the software, the Bland–Altman analysis was used for both parameters by plotting the difference between the two measurement techniques (2D/3D-matching and CT measurement) against their averages. The ICC was used for calculation of agreement of software consistency, reproducibility, and reliability, and was graded as: ICC < 0.20 for slight agreement; 0.21–0.40 for fair agreement; 0.41–0.60 for moderate agreement; 0.61–0.80 for substantial agreement; and >0.80 for almost perfect agreement. The level of significance was set as 0.05.

RESULTS
External Validation
The mean accuracy in the cadaver trials was 0.7 ± 1.7° (−3.7° to 4.0°) for the inclination and 1.2 ± 2.4° (−5.3° to 5.6°) for anteversion. In the clinical series, the mean accuracy for radiographs without gonadal shielding or deep centering was 1.7 ± 1.7° (−4.6° to 5.5°) for the inclination and 0.9 ± 2.8° (−5.2° to 5.7°) for anteversion. No significant difference was found for accuracy of anteversion between the cadaver and clinical series (p = 0.57), whereas the accuracy for inclination differed significantly (p < 0.001). Compared to radiographs without gonadal shielding or deep centering, no significant difference was found in accuracy for inclination in radiographs with gonadal shielding (1.4 ± 1.1°, range −0.8° to 3.3°, p = 0.94), with deep centering (1.9 ± 1.9°, range −6.3 to 6.2°, p = 0.42), or with a combination of both (2.0 ± 1.1°, range −0.8 to 3.3°, p = 0.54). Calculation of anteversion with gonadal shielding was less accurate compared to radiographs with gonadal shielding (1.9 ± 3.4°, range −8.1 to 8.5°, p = 0.05), radiographs with deep centering (3.3 ± 3.7°, range −7.1 to 9.2°, p < 0.001), or a combination (2.8 ± 3.1°, range −2.7 to 8.5°, p = 0.05).

Bland–Altman analysis showed that the mean of the measurement pairs was spread evenly and randomly for inclination and anteversion, in both the cadaver and clinical data for radiographs without gonadal shielding or deep centering, indicating no systematical error (Fig. 5).

Internal Validation
The software showed a good consistency for both parameters with an ICC for inclination of 0.96 (95% CI: 0.93–0.98) and for anteversion of 0.95 (95% CI: 0.91–0.98).

Reproducibility and Reliability Analysis
A good reproducibility and reliability was found for both parameters (Table 1).

DISCUSSION
Measurement of cup orientation out of an AP pelvic radiograph is highly inaccurate because of individual pelvic orientation during radiograph acquisition. In particular, pelvic tilt shows a large variation predominantly affecting the measurement of cup anteversion. Optimal cup position critically affects the long-term result after THA, but our inability to measure cup position accurately limits our ability to
determine optimal cup orientation, to assess new methods of improving cup positioning in surgery, and to assess the effect of acetabular component positioning on outcome. Our goal was to perform a thorough validation assessment of a new and fully automated 2D/3D-matching technique. We showed HipMatch to be an accurate, consistent, reliable, and reproducible tool, matching a 3D-model based on CT with the contours of the projected pelvis on a postoperative radiograph, to measure cup orientation relative to anatomical landmarks, corrected for malposition of the pelvis on the radiograph.

We confirmed prior studies of 2D/3D matching,9,25,26 and further demonstrated that CAD models of specific implants are not necessary for accurate measurement of cup orientation, so long as the opening face of the acetabular component is planar and round. An additional matching process of the CAD model to the projected cup (as proposed by other authors9,25,26) is unnecessary. The additional step to match the CAD model to the projected cup was either performed manually,9,26 or a manual coarse alignment was initially required.25 This implicates, according to the authors, a certain degree of subjective judgment. Nevertheless, neither an intra/interobserver study nor an analysis of software consistency were performed for the previously presented methods.9,25,26 The only published validation results on accuracy are comparable to our study; however, their results are based on relative small numbers.

Blendea et al.9 reported a mean error for inclination of 0.9°/C6.13°8 and 0.0°/C6.20°8 for version with maximal errors of 4.7°8 and 5.0°8 for inclination and version based on 19 measurements using a matching technique performed manually.9,26 This implicates, according to the authors, a certain degree of subjective judgment. Nevertheless, neither an intra/interobserver study nor an analysis of software consistency were performed for the previously presented methods.9,25,26 The only published validation results on accuracy are comparable to our study; however, their results are based on relative small numbers.

Blendea et al.9 reported a mean error for inclination of 0.9° ± 1.3° and 0.0° ± 2.0° for version with maximal errors of 4.7° and 5.0° for inclination and version based on 19 measurements using a matching technique performed manually. Jaramaz and Eckman25 published an update of this software using automatic algorithms for matching, and found in the experimental setup, using an acrylic pelvis with fiducial markers, an accuracy for inclination of 0.1° ± 0.2° (−0.5° to 0.2°) and −0.4° ± 0.9° (−3.1° to 1.7°) for version. Clinically, the error for

Table 1. Results of Reproducibility and Reliability

<table>
<thead>
<tr>
<th>Parameters</th>
<th>ICC Intraobserver 1&lt;sup&gt;a&lt;/sup&gt;</th>
<th>ICC Intraobserver 2&lt;sup&gt;a&lt;/sup&gt;</th>
<th>ICC Interobserver&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inclination</td>
<td>0.98 (0.97–0.99)</td>
<td>0.99 (0.98–0.99)</td>
<td>0.98 (0.98–0.99)</td>
</tr>
<tr>
<td>Antversion</td>
<td>0.96 (0.94–0.98)</td>
<td>0.97 (0.95–0.98)</td>
<td>0.95 (0.93–0.96)</td>
</tr>
</tbody>
</table>

ICC, intraclass correlation coefficient.
<sup>a</sup>Mean (95% confidence interval).
inclination was $0.4 \pm 0.8^\circ$ and $0.6 \pm 0.8^\circ$ for version with maximum errors of $2.2^\circ$ and $2.0^\circ$ for inclination and version. Penney et al.\textsuperscript{26} found an accuracy for inclination of $0.4 \pm 1.1^\circ$ and $1.1 \pm 2.5^\circ$ for version based on 10 measurements.

While accurate, our method is currently limited to those hips that have had a CT study at some point during treatment. Still, statistical modeling of large populations of human pelvises may allow the technique to be applied to patients who have not had a CT study. Further studies will show whether this is possible with a clinically sufficient accuracy. Also, since acetabular components that are retroverted look identical on an AP radiograph to components that are identically anteverted, retroverted components cannot be accurately measured unless these cases are identified as being retroverted by the user, based on images other than the AP radiograph.\textsuperscript{17}

The best results with the smallest error were achieved using radiographs that included the anterior superior iliac spines and did not have gonadal shielding. These anatomical landmarks should be included on the radiograph whenever HipMatch is applied. Also, the measurements of consistency, reproducibility, and reliability were based on cadaver data under optimal conditions, including the maintenance of a nearly perfect centering of the x-ray beam and no disturbing soft tissue, femora, or bony spinal structures. Analogous studies performed under less optimal clinical circumstances might yield different results.

Several possible sources for errors in accuracy exist. First, any difference between the definition of the APP on the preoperative CT using HipMatch and on the post-operative CT using Osirix\textsuperscript{18} would result in differences in the calculations. However, the definition of the APP is reproducible.\textsuperscript{7} Further, metal artifacts on the postoperative CTs could affect the ground truth measurements. Also, some of the postoperative radiographs were taken years after surgery, and although no case had radiographic signs of osteolysis or loosening, small changes in position of the acetabular component within the pelvis could have affected the measurements. Finally and perhaps most importantly, the accuracy is dependent on the manual implementation of the Ackland et al.\textsuperscript{16} method with exact definition of the projected ellipse and the corresponding axis.

The current study demonstrated that the orientation of the acetabular component can be calculated accurately from plain radiographs using 2D/3D matching. This method may have significant implications for future assessment of THA. This technique may be used to assess the utility of any proposed method of aligning acetabular components during surgery, whether the method uses mechanical alignment guides, local landmarks, or any type of surgical navigation. Similarly, assessment of larger groups of THAs that did and did not have a postoperative dislocation may allow for greater refinement of the “safe zone,”\textsuperscript{82} which may in fact be different for different surgical techniques, implants, and patients. Also, it may be possible to employ these methods intraoperatively to measure cup position correctly on an intraoperative radiograph.\textsuperscript{27,28} Since the orientation of the pelvis itself is always calculated using this technique, orientation of the pelvis can also be calculated both before and after surgery in various positions such as standing or lying.\textsuperscript{29,30} Optimal cup orientation will likely be determined for individual patients based on preoperative assessment of pelvic position and motion using preoperative radiographs and 2D/3D matching in the future.

ACKNOWLEDGMENTS

One or more of the authors (S. D. S., M. T., G. Z., K. A. S.) has received funding from a grant from the National Center of Competence in Research “Computer Aided and Image Guided Medical Interventions (Co-Me)” of the Swiss National Science Foundation (SNSF), a personal grant for prospective researchers (S. D. S.) or advance researchers (M. T.) by SNSF, and from the Association for Orthopaedic Research AFOR (M. T.). No author has a commercial association that might pose a conflict of interest in connection with the submitted article. Each author certifies that his institution has approved the human protocol for this investigation, that all investigations were conducted in conformity with ethical principles of research, and that informed consent was obtained.

REFERENCES


