Reliability of Predictors for Screw Cutout in Intertrochanteric Hip Fractures

Kirstin De Bruijn, MD, Dennis den Hartog, MD, PhD, Wim Tuinebreijer, MD, PhD, MSc, MA, and Gert Roukema, MD

Investigation performed at Maasstad Hospital, Rotterdam, The Netherlands, and Erasmus MC, University Medical Center, Rotterdam, The Netherlands

Background: Following internal fixation of intertrochanteric hip fractures, tip apex distance, fracture classification, position of the screw in the femoral head, and fracture reduction are known predictors for screw cutout, but the reliability of these measurements is unknown. We investigated the reliability of the tip apex distance measurement, the Cleveland femoral head dividing system, the three-grade classification system of Baumgaertner for fracture reduction, and the AO classification system as predictors for screw cutout.

Methods: All patients with an intertrochanteric hip fracture who were managed with either a dynamic hip screw or a gamma nail between January 2007 and June 2010 were evaluated from our hip trauma database.

Results: The tip apex distance measurement was reliable and patients with device cutout had a significantly higher tip apex distance. The agreement between observers with regard to screw position and fracture reduction was moderately reliable. After adjustment for tip apex distance and screw position, A3 fractures were at more risk of cutout compared with A1 fractures. Poor fracture reduction was significantly related with a higher incidence of cutout in univariate analysis, but not in multivariate analysis. Central-inferior and anterior-inferior positions, after adjustment for tip apex distance and screw position, were significantly protective against cutout.

Conclusion: To decrease probable risks of cutout, the tip apex distance needs to stay small or the screw needs to be placed central-inferiorly or anterior-inferiorly.

Level of Evidence: Therapeutic Level III. See Instructions for Authors for a complete description of levels of evidence.

Failed osteosynthesis has a major impact on mortality and morbidity among the elderly with intertrochanteric hip fractures.

Osteosynthesis with a fixed angle device in combination with a dynamic sliding lag screw is the favored treatment for intertrochanteric hip fractures. The sliding lag screw can be used in combination with an intramedullary nail (for example, a gamma nail) or an extramedullary side plate (for example, a dynamic hip screw).

For unstable intertrochanteric hip fractures (AO classification A2 and A3), our clinic uses the gamma nail because of the possible mechanical advantage over plate fixation in unstable intertrochanteric hip fractures. For stable intertrochanteric hip fractures (AO classification A1), our clinic uses a two-hole dynamic hip screw.

The most common mechanism of fixation failure due to cutout of the screw from the femoral head is the collapse of the neck-shaft angle into varus. Earlier research has shown that several variables, such as fracture type, fracture reduction, tip apex distance, and position of the screw in the femoral head, can increase the risk of cutout, after internal fixation of intertrochanteric hip fractures.

The tip apex distance is the sum of the distance in millimeters from the tip of the screw to the apex of the femoral head on anteroposterior and lateral radiographs (Fig. 1). It has been demonstrated that the tip apex distance is a predictor for cutout. A tip apex distance <25 mm is protective of the screw cutting out of the femoral head. However, the reliability of the tip apex distance measurement has only been partially investigated. Baumgaertner et al. found that it was a reproducible...
measurement, but the reliability of the tip apex distance measurement has never been investigated. Cleveland et al. introduced a system in which the femoral head was divided into nine zones. Screw cutout in relation to the position of the screw in the femoral head and the reliability of the Cleveland femoral head dividing system were reported.

There is no general classification system for classifying hip fracture reduction. Baumgaertner et al. used a well-defined, three-grade classification system for fracture reduction. Because reliability has not been established for this classification, we investigated this issue. Additionally, we studied the relationship between the three-grade fracture classification system and screw cutout from the femoral head.

To classify fracture type, we used the reliable AO classification system with only the groups A1, A2, and A3. No further subgroups were used.

Materials and Methods

All patients from the hip trauma database who were managed for an intertrochanteric hip fracture between January 2007 and June 2010 at a level-2 trauma hospital in the Netherlands were evaluated. Patients were managed with either a dynamic hip screw or a gamma nail, as noted above. Patients who did not undergo a complete follow-up at three months postoperatively were excluded. Patients who died within three months after surgery were also excluded. If the quality of the postoperative radiographs was too poor to allow visualization of the femoral head, the patient was excluded.

Clinical data were collected from the hip trauma database. Available data included sex, age at surgery, operating time, American Society of Anesthesiologists (ASA) classification, type of implant, and date of death (if death occurred). The fracture classification, quality of fracture reduction, position of the screw, and tip apex distance were determined with use of preoperative and postoperative anteroposterior and lateral digital radiographs. All specifications were measured in the picture archiving and communication system (PACS). An earlier study showed that the PACS is suitable for measuring the tip apex distance.

The follow-up period varied from three months to three years after the operation. The mean follow-up period (and standard deviation) was 2.0 ± 0.99 years. We expected screw cutout to occur within three months after the operation, as seen in previous studies, which correlates with the fact that fracture healing can be expected at three months postoperatively. We ensured that every patient was followed for at least three months, thus minimizing the chances of missing a patient with screw cutout.

The fractures were classified with use of the AO classification system, without the use of subgroups.

The postoperative quality of fracture reduction was described as good, acceptable, or poor, according to the definitions of the three-grade classification system proposed by Baumgaertner et al. A good reduction was classified as normal or slight valgus alignment on the anteroposterior radiograph, <20° of angulation on the lateral radiograph, and ≤4 mm of displacement of any fragment. An acceptable reduction had to meet the criteria for a good reduction, but for either the alignment or the displacement, not for both. A poor reduction met none of the criteria.

The measurement of the tip apex distance, as developed by Baumgaertner et al., was also performed. The tip apex distance is the sum of the distance, in millimeters, from the tip of the screw to the apex of the femoral head, on the anteroposterior and lateral radiographs (Fig. 1). The first postoperative radiographs were used to measure the tip apex distance.

With regard to position, the femoral head was divided into nine areas, as reported by Cleveland et al. On the anteroposterior radiograph, the locations are superior, central, and inferior. On the lateral radiograph, the locations are anterior, central, and posterior. With use of this information, the exact position of the screw was determined.

Two independent observers performed tip apex distance measurements and determined screw position and fracture reduction. One observer was a consultant trauma surgeon and the other observer was a surgical resident. The observers were shortly briefed on the tip apex distance measurement, the Cleveland zone system, the Baumgaertner fracture reduction system, and the AO classification system.

All observations were done on the immediate postoperative radiograph.

Statistical Analysis

Statistical analysis was performed with use of SPSS version 16.01 (SPSS, Chicago, Illinois) and MedCalc version 9.2 (MedCalc Software, Mariakerke, Belgium) software on a personal computer. The chi-square test and Student t test were used as applicable for univariate analysis. The independent variables from the univariate analysis with a p value of <0.10 were entered in a multivariate logistic regression, with the occurrence of cutout of the hip screw as the dependent variable. The likelihood ratio backward test was conducted to find the best-fit model by selecting variables, one by one. The probability for entry was set at p < 0.05, and the probability for removal was set at p = 0.10. In this model for the tip apex distance, the mean of the tip apex distance of the two observers was used.

Univariate logistic regression was used to calculate the predicted probability of a cutout for each observed tip apex distance value. The intraclass correlation coefficient with its 95% confidence interval (95% CI) for interval data was calculated to assess the interobserver reliability of one observer (single-measure intraclass correlation coefficient) and of two observers (average-measure intraclass correlation coefficient). The two-way random-effects model with absolute agreement was selected. To assess the interobserver reliability of nominal and ordinal data, kappa coefficients

\[
TAD = (X_{ap} \times \frac{D_{true}}{D_{ap}}) + (X_{lat} \times \frac{D_{true}}{D_{lat}})
\]

Fig. 1
A line drawing showing the measurement of the tip apex distance (TAD). Use of the true diameter of the screw (Dtrue) controlled for magnification. Dtrue is the known diameter of the lag screw. Dap is the measured diameter of the lag screw on an anteroposterior radiograph. Dalat is the measured diameter of the lag screw on a lateral radiograph. Xap is the distance from the tip of the screw to the center of the femoral head on an anteroposterior radiograph. Xlat is the distance from the tip of the screw to the center of the femoral head on a lateral radiograph. (Reproduced from: Baumgaertner MR, Curtin SL, Lindskog DM, Keggi JM. The value of the tip-apex distance in predicting failure of fixation of peritrochanteric fractures of the hip. J Bone Joint Surg Am. 1995;77:1058-64.)
were calculated with use of 95% CI. The standard error of measurement of the tip apex distance (SE_{measurement} = \text{the square root of the mean square residual}) was used to calculate the amount of errors between the two tip apex distance measurements. To account for the relationship between the mean and the standard error, the coefficient of variation (CV) was calculated with use of the following formula: CV (%) = SE_{measurement}/\text{mean} \times 100. The coefficient of variation expresses the standard error of measurement as a proportion of the mean and is unitless. A low coefficient of variation represents a better measurement than a high coefficient of variation.

The receiving operating characteristic (ROC) curve for the mean tip apex distance was used to calculate the area under the curve (C-statistics), and the best threshold value of the tip apex distance was used to predict screw cutout.

The differences in the mean tip apex distance in the groups A1, A2, and A3 of the AO classification system were tested with use of one-way analysis of variance (ANOVA).

### Source of Funding
There was no external funding source used for this study.

### Results

#### Data on the Patients

Two hundred and ninety patients who had had intertrochanteric hip fractures were identified. Twenty-four patients who did not have a complete duration of follow-up at three months postoperatively were excluded. Thirty-two patients died within three months after surgery and were excluded. The quality of nineteen of the postoperative radiographs was too poor to see the femoral head properly, so these patients were excluded. Therefore, 215 patients who had had fractures were suitable for our study.

There were 168 women (78%) and forty-seven men (22%) in our study. The mean age (and standard deviation) was 78.0 ± 14.9 years. A gamma nail was used to obtain fracture fixation in 175 patients (81%), and a dynamic hip screw was used in forty patients (19%).

Twenty-nine people died more than three months after surgery, a mortality rate of 13.5%. There is no accurate time interval for this mortality rate because the overall duration of follow-up was not the same for every patient.

In total, sixteen patients (7%) had screws in which the fixation failed. Of these patients, three had screw cutout after the third postoperative month, and thirteen had screw cutout within three months. Additionally, six patients required reoperation that was not due to screw cutout. Of these patients, one had severe infection, another had an unsuccessful initial operation because of malalignment of the dynamic hip screw plate on the postoperative radiograph, and, after the third postoperative month, four reported pain and had a second operation in which the fixation device was removed after consolidation of the fracture.

### Reliability of the Measurements

To assess the interobserver reliability of tip apex distance measurement, the intraclass correlation coefficient was calculated. The single-measure intraclass correlation coefficient (the tip apex distance measured by a single observer) was 0.81 (95% CI, 0.71 to 0.87) and the average-measure intraclass correlation coefficient (the mean measure of the two observers) was 0.89 (95% CI, 0.83 to 0.93), which both correspond with good reliability. The corresponding standard error of measurement and coefficient of variation were 2.78 and 13%.

The single-measure intraclass correlation coefficient was calculated. The single-measure intraclass correlation coefficient (the tip apex distance measured by a single observer) was 0.81 (95% CI, 0.71 to 0.87) and the average-measure intraclass correlation coefficient (the mean measure of the two observers) was 0.89 (95% CI, 0.83 to 0.93), which both correspond with good reliability. The corresponding standard error of measurement and coefficient of variation were 2.78 and 13%.

#### Univariate Analysis

In Table II and the Appendix (Student t test, chi-square test, logistic regression), different variables are presented, split for patients with or without cutout of the screw. Table II presents and compares the numbers and percentages as well as the means and the standard deviations of patient characteristics. A table in the Appendix compares the nonadjusted odds ratios of the categorical variables for patients with and without cutout of the screw. Measurements of both observers are shown.

We found that for the sixteen patients who had screw cutout, the average tip apex distance (and standard deviation) was 25.9 ± 6.6 mm compared with 21.3 ± 6.1 mm for the 199 patients who had screws that did not cut out. The difference between the two groups was significant (p = 0.005).

Among the patients who had screw cutout, univariate analysis revealed no significant results for the Student t test or chi-square test for the variables of age, sex, ASA classification system, device, operation time, length of hospital stay, or time to operation (Table II).

With regard to the AO classification system, the A3 classification had an odds ratio of 3.32 compared with the A1 classification; however, this result was not significant (p = 0.089) (see Appendix).
Poor fracture reduction was associated with a significantly higher risk of screw cutout compared with good reduction (odds ratio [OR], 5.19; p = 0.017) for a single observer.

Division of the femoral head into nine zones has been recommended by Cleveland et al.\textsuperscript{12}. Screws were found to have been placed in all possible locations by one observer (see Appendix). No screws cut out of the posterior-inferior position. Screw position proved to be significantly related to cutout of the screw for both observers (p = 0.009 and p = 0.014). The anterior-inferior and central-inferior positions had significantly favorable odds ratios (0.11 for both positions) compared with the anterior-superior position for the patients without cutout. The central-central position had a significant odds ratio (<1) for both observers (p = 0.025 and p = 0.028, respectively).

| TABLE II Overview of the Patients without or with Screw Cutout |
|-----------------------|-----------------|-----------------|-----------------|-----------------|
| Characteristic        | Without Cutout  | With Cutout     | CI Mean Difference | P Value |
| Age\* (yr)            | 77.8 ± 15.0     | 79.9 ± 13.1     | (−9.7; 5.6)       | 0.595 |
| Operation time\* (min)| 97.7 ± 32.0     | 121.4 ± 47.6    | (−49.3; 2.0)      | 0.069 |
| Length of hospital stay\* (d) | 10.8 ± 11.8 | 20.1 ± 20.8     | (−20.5; 1.9)      | 0.098 |
| Time to operation\* (d) | 1.7 ± 5.9       | 2.9 ± 6.6       | (−4.3; 1.9)       | 0.452 |
| Tip apex distance\* (mm) | 21.3 ± 6.1      | 25.9 ± 8.6      | (−7.9; −1.5)      | 0.005 |
| No. of patients†      | 199 (93%)       | 16 (7%)         |                 |      |
| Sex†                  |                 |                 |                 | 0.532 |
| Female                | 154 (77%)       | 14 (88%)        |                 |      |
| Male                  | 45 (23%)        | 2 (13%)         |                 |      |
| Device†               |                 |                 |                 | 1.0  |
| Dynamic hip screw     | 37 (19%)        | 3 (19%)         |                 |      |
| Gamma nail            | 162 (81%)       | 13 (81%)        |                 |      |
| Death†                |                 |                 |                 | 0.459 |
| Yes                   | 26 (13%)        | 3 (19%)         |                 |      |
| No                    | 173 (87%)       | 13 (81%)        |                 |      |
| ASA classification†‡  |                 |                 |                 | 0.299 |
| I                     | 6 (4%)          | 0 (0%)          |                 |      |
| II                    | 98 (65%)        | 4 (50%)         |                 |      |
| III                   | 46 (30%)        | 4 (50%)         |                 |      |
| IV                    | 2 (1%)          | 0 (0%)          |                 |      |
| AO classification†     |                 |                 |                 | 0.112 |
| A1                    | 77 (39%)        | 4 (25%)         |                 |      |
| A2                    | 93 (47%)        | 7 (44%)         |                 |      |
| A3                    | 29 (15%)        | 5 (31%)         |                 |      |
| Fracture reduction†‡  |                 |                 |                 | 0.228 |
| Observer 1            |                 |                 |                 |      |
| Good                  | 125 (63%)       | 8 (50%)         |                 |      |
| Moderate               | 66 (33%)        | 6 (38%)         |                 |      |
| Poor                   | 8 (4%)          | 2 (13%)         |                 |      |
| Observer 2            |                 |                 |                 | 0.087 |
| Good                  | 83 (42%)        | 5 (31%)         |                 |      |
| Moderate               | 100 (50%)       | 6 (38%)         |                 |      |
| Poor                   | 16 (8%)         | 5 (31%)         |                 |      |
| Screw position (see Appendix) |                |                 |                 | 0.009, 0.014 |
| Reoperation†          |                 |                 |                 | 0.000 |
| Yes                   | 2 (1%)          | 16 (100%)       |                 |      |
| No                    | 197 (99%)       | 0 (0%)          |                 |      |

*The values are given as the mean and the standard deviation. †The values are given as the number of patients, with the percentage in parentheses. ‡Data were unavailable for sixty-seven patients.
Multivariate Analysis and Adjusted Odds Ratios
Table III shows the results of the multivariate logistic regression with the mean tip apex distance of the two observers and the categorical variables of the more experienced observer as independent variables. The odds ratio of the tip apex distance adjusted for the AO classification system and screw position showed significant results (OR, 1.11 [95% CI, 1.02 to 1.21]; p = 0.022). The odds ratio for the A3 category of the AO classification system compared with that of the A1 category was 14.24 (95% CI, 2.29 to 88.72; p = 0.004). The odds ratio for the anterior-inferior screw position was 0.07 (95% CI, 0.01 to 0.62; p = 0.027); and the odds ratio for the central-inferior screw position was 0.08 (95% CI, 0.01 to 0.57; p = 0.016).

To estimate a threshold value that can predict cutout of the screw, we used an ROC curve. The criterion for cutting out was at the 19.9-mm level. Sensitivity was 87.5% and specificity was 44.7%. The area under the curve (AUC) was 0.67 (Fig. 2). At the 25-mm level, the sensitivity and specificity were 50% and 71.4%, respectively. The sensitivities of both thresholds were not significantly different (p = 0.0565). In contrast, the specificity of the 25-mm cutoff was significantly higher than the 19.9-mm cutoff value (p < 0.0001).

Figure 3 shows the predicted probability of cutout with ascending tip apex distance. The odds ratio of the tip apex distance adjusted for the AO classification system and screw position per 5-mm tip apex distance increase was 1.69 (95% CI, 1.08 to 2.64); for a 10-mm increase, this value was 2.84 (95% CI, 1.16 to 6.95).

The one-way ANOVA test showed that there was no significant difference (p = 0.42) among the mean tip apex distance values of the three AO fracture classifications. The mean tip apex distance was 22.1 mm for A1 fractures, 21.0 mm for A2 fractures, and 22.1 mm for A3 fractures.
Discussion

Published reports have associated variables such as intertrochanteric hip fracture reduction, tip apex distance, and position of the screw with risk of screw cutout from the femoral head.[5-11] Our main goal was to investigate the reliability of the tip apex distance measurement, the Cleveland femoral head dividing system, and the three-grade classification system of Baumgaertner for fracture reduction. We also studied the more general variables associated with a patient with an intertrochanteric hip fracture to see whether they were of any interest.

Our results showed that of all measurements, the only reliable one was measuring tip apex distance in the PACS. Baumgaertner et al.[6] studied the agreement of the tip apex distance measurement for two observers, but only for 28% of the fractures in their study. We calculated the intraclass correlation coefficient for all our fractures and obtained a good reliability score. An advantage of the tip apex distance measurement is that it is not influenced by the experience of the observer. Its measurement is based not on skillful interpretation but on simple, well-defined rules. Like other studies,[6,9-11], the present study demonstrated a correlation between tip apex distance and cutout of the screw from the femoral head. The multivariate model shows that a higher tip apex distance is associated with a 1.1-times higher risk of cutout per millimeter increase of the tip apex distance. Although earlier studies[5,7] showed that patients who have a tip apex distance >25 mm are at risk of screw cutout, we found a threshold value for tip apex distance at the 19.9-mm level. The 19.9-mm threshold thus proves to be a better predictor (although not significantly) than the 25-mm threshold. Greater numbers of patients should be evaluated to clarify this observation.

The 25-mm threshold, as stated in the study of Baumgaertner et al.[3], is significantly more reliable in predicting that a screw will not cut out. Although there is evidence for a tip apex distance threshold with a maximum of 25 mm, we argue that a 19.9-mm threshold would be effective in reducing the risk of cutout.

The fractures classified as A3 fractures were shown to be at fourteen times more risk of cutout, which could be explained by the fact that A3 fractures are unstable and more difficult to reduce. This fact corresponds with our results that poor fracture reduction also contributes to higher risks (up to five times) of screw cutout. Still, the definite risk of cutout remains relatively small.

The measurement of fracture reduction proved to be only moderately reliable. This is likely due to the fact that classifying fracture reduction is based largely on interpretation rather than on well-defined measurements.

There is no relation between fracture classification and tip apex distance. Our assumption was not confirmed that A3 fractures would have a larger tip apex distance because of the expected difficulties during operation that were due to the more complicated fracture.

Determining the exact position of the screw in the femoral head is also of importance and needs refining. The central-central, anterior-inferior, and central-inferior positions contribute significantly to protect the screw from cutting out of the femoral head. This finding is in agreement with previous studies,[6,9-11], in which a central and inferior position is preferred. Despite not experiencing any screw cutout, the posteroinferior zone did not have a significant correlation with cutout and was not significantly protective against screw cutout. The moderate reliability of the Cleveland zone system that we found in our study could be due to the fact that nine zones are too many to yield good reliability. It is possible that other circumstances and other observers may result in a different reliability outcome for this system. Sex had no influence on the screw cutting out of the femoral head. Of the sixteen patients whose fixation had failed, fourteen (88%) were women. This result was to be expected because the sex of 78% of the total population that had been studied was female and a previous study had showed no significant association between these variables[6].

The operation time, hospital stay, and time to operation were longer in the cutout group, but the differences were not significant. The small sample size of the cutout patient group (n = 16) reduced the power to detect significant differences. An explanation could be that there were higher percentages of patients in ASA class 3 and A3 fractures within the group of patients with screw cutout, although these values were not significant. Larger study groups are needed for determining whether the variables mentioned above significantly contribute to a longer operation time, hospital stay, and time to operation.

The population of interest is a fragile one, with an overall mortality rate of almost 14% supporting our concerns about the need for the first operation to be successful. It is obvious that the success rate of the operation partly depends on factors that the surgeon cannot influence. Surgeons should therefore be aware of the factors that they can manipulate with a positive outcome. For the tip apex distance, for example, Baumgaertner and Solberg showed that surgeon awareness is associated with a lower rate of screws cutting out of the femoral head[6].

Because this was an observational and retrospective study, it had limitations. However, by calculating the adjusted
odds ratios and therefore adjusting for confounders, we partly eliminated the confounding bias. We lost a substantial amount of patients because of incomplete follow-up.

Because only sixteen patients had screw cutout, the confidence intervals of our estimates were wide, indicating a low level of precision. Therefore, research with larger study groups is needed to obtain more reliable results.

We can conclude that the tip apex distance measurement in the PACS is a reliable method for reducing the risk of screw cutout. By keeping the tip apex distance at a <25-mm threshold (without penetrating the joint), placing the screw centrally or inferiorly, and paying even more attention to A3 fractures, it can be expected that cutout of the screw of the femoral head will be minimized.

Appendix

A table showing the nonadjusted odds ratio for effects of different categorical variables on the cutout of hip screws with use of univariate logistic regression and a figure showing the observer differences for determining screw position, by zone, for the 215 patients are available with the online version of this article as a data supplement at jbjs.org.

References