Displaced femoral neck fractures: towards better practice
Heetveld, M.J.

Citation for published version (APA):
Chapter 6

Expert prediction of internal fixation failure in displaced femoral neck fractures

M.J. Heetveld, E.L.F.B. Raaymakers, J.S.K. Luitse, D.J. Gouma

Submitted
Abstract

Background: The influence of precise surgical technique on clinical outcome of internal fixation (IF) for displaced femoral neck (DFN) fractures is underreported in the literature. Inadequate fracture reduction and/or implant position, as determined by expert rating early after surgery, potentially has clinical consequences. The aim of this study was to determine the correlation between expert rating of IF technique and clinical outcome.

Methods: A prospective multicenter 2-year follow-up study involving 102 patients selected for IF was performed. Operations were performed or supervised by experienced surgeons. Technical recommendations for fracture reduction and implant positioning were described in the study protocol according to recent literature. All pre-, intra- and 1-2 day postoperative radiographs were digitalised and rated by two independent experts, who were blinded to each other, for the hospital and for clinical and radiological outcome. If one or both raters assigned an inadequate rating, the case was considered technically inadequate in the analysis. Ratings were divided into 3 groups: overall technique, fracture reduction and implant positioning. The clinical endpoints were failure of IF necessitating revision, postoperative complications and mortality. The association between the ratings and the clinical endpoints were calculated with correlation coefficients and regression analysis.

Results: The highest correlation with 2-year clinical failure of IF was found for fracture reduction technique rating \( r = 0.294, p < 0.01 \); agreement between raters: kappa = 0.39). For overall technique rating a significant correlation with IF failure was also found \( r = 0.265, p < 0.01 \); agreement between raters: kappa = 0.65). A likely correlation existed between implant positioning rating and IF failure \( r = 0.181, p = 0.07 \); agreement between raters: kappa = 0.16). Logistic regression demonstrated that inadequately rated reduction (varus) on the AP view was independently associated with clinical failure (Odds Ratio: 17.5, \( p < 0.01 \)). Inadequate reduction on the lateral view, overall technique and implant positioning were not associated with failure. No association between technical rating and postoperative complications was calculated.

Conclusion: Early postoperative inadequate expert rating of IF technique correlated with clinical failure within 2 years. This result opens the discussion about expert rating of surgical performance preferably intraoperatively or early postoperatively. Although two experts may disagree, a single inadequate expert rating should prompt at least consideration of a pre-emptive change in the choice of therapy.
Introduction

The importance of precise surgical technique for reduction and internal fixation (IF) of displaced femoral neck fractures is underreported. Surgical technique or experience level of surgeons can influence clinical outcome.\(^1\) Fracture reduction criteria were developed more than 30 years ago and are generally accepted.\(^5\) Implant positioning criteria have also been well described,\(^2,6\) but have not been subjected to a technical analysis in relation to prospective clinical outcome. The principles described in the literature for fracture reduction and implant positioning are not always optimally adopted and it is unclear what constitutes an adequate or inadequate surgical performance. Few retrospective studies have correlated technical performance with outcome.\(^2,3\) No study has reported independent expert technical rating and a possible correlation with clinical outcome in displaced femoral neck fracture patients. In a recently performed study of internal fixation in elderly patients with a high physiologic demand, an inadequate technical performance was observed in 13% of patients.\(^7\)

Overall, clinical failure resulting in revision occurs in about one-third of patients with a displaced femoral neck fracture treated with IF.\(^8-11\) In these patients, an early expert rating, performed intra- or postoperatively, may detect inadequate reduction or fixation technique. This could hypothetically lead to a pre-emptive change in clinical strategy. Early (re-) intervention before clinical failure becomes evident would potentially reduce patient discomfort, improve outcome and save costs. The aim of this study was therefore to determine whether an inadequate technical rating by independent experts was correlated with clinical failure, defined as revision to arthroplasty, development of postoperative complications and 2-year mortality in displaced femoral neck fracture patients.

Patients and Methods

Design and inclusion  The selected IF patients represented one arm of a prospective, multi-center study, which aimed to verify a modified Physiologic Status Score (PSS) treatment protocol for patients with a displaced femoral neck fracture, which has been reported previously.\(^7,12\) Four university and 6 non-university teaching hospitals participated in that study, in which patients were selected for either IF or hemiarthroplasty based on their quantified physiologic status. All patients aged 60-90 years with a displaced femoral neck fracture were eligible for inclusion. The in- and exclusion criteria are listed in table 1. Inclusion was commenced after the protocol had received Institutional Review Board consent.
Table 1: In- and exclusion criteria

Inclusion criteria:
- Aged 60 years and above
- Displaced intracapsular femoral neck fracture, no trochanteric fractures
- Physiologic Status Score \( \geq 20 \) points, defined as:
  - Mobile with no or one walking aid outside the home
  - Living in own home, partial dependence on social services/relatives allowed
  - A good cognitive function
  - American Society of Anaesthesiologist's (ASA) class 1 or 2 medical condition

Exclusion criteria:
- Undisplaced femoral neck fractures
- Fractures of the femoral neck older than 48 hours
- Patients with advanced arthritis of the ipsilateral hip joint who would normally be treated by total hip replacement. The diagnostic signs on plain X-ray of this condition were: subchondral sclerosis, narrowing of the articular crevice, formation of osteophytes and subchondral cyst formation
- Patients with established rheumatoid arthritis or showing signs of this disease on plain X-ray.
- Patients with a pathological fracture or who sustained a spontaneous fracture after a comparatively minor trauma with signs of osteolysis at the fracture sight on plain X-ray.

Operation

The previously reported PSS study protocol \(^7,^{12}\) required operation or supervision by staff surgeons with experience in IF for displaced femoral neck fractures (i.e. operator or supervisor of at least 10 IF per year) in all participating hospitals. In the Anterior-Posterior (AP) fluoroscopic image a reduction of the femoral head in 0-20 degrees of valgus, corresponding with a Garden Index of 160-180 degrees, was required.\(^5\) In the lateral view the line between the centre of the femoral head and long axis of the femoral shaft was required be as close as possible to 180 degrees.\(^6\) Up to 10 degrees of femoral head retroversion was considered to be acceptable. Cannulated cancellous bone screw insertion according to the three-point fixation \(^2,^6\) principle or compression hip screws (CHS) were recommended as implants.\(^13\) For steep Pauwels type 3 fractures a fixed angle CHS was the preferred option.\(^14-16\)

Positioning of the CHS was required to be in the central to caudal half of the femoral head in the AP view. In the lateral view CHS positioning in the central-dorsal part of the femoral head was advised. Postoperatively, after consultation with physiotherapy, all patients were mobilised with full weight-bearing as much as discomfort would allow.

Data collection and follow-up

The pre-, intra-, and 1-2 day postoperative radiographs (AP and lateral views) of all internal fixation patients were digitally photographed and branded onto two CD-ROM's. All references to patient name and institution were removed digitally before branding.
Case report forms (CRF) were completed on admission, at operation and upon discharge. Follow-up was scheduled at 1 and 2 years. Plain radiographs were also taken at each follow-up interval. The primary end-point of this study was revision of IF to arthroplasty. Secondary end points were postoperative complications and mortality.

**Technical rating** As no guidelines exist for the rating of technique on radiographs, two independent expert professors with an orthopaedic and a general surgery-traumatology background each with over 20 years of clinical and scientific experience, were asked to blindly rate fracture reduction and positioning of the implants. Both experts were requested to rate technique according to the recommended methods in the study protocol above in combination with their expertise. The experts were blinded to each others’ findings, the hospital and to clinical and radiological outcome. They were asked to rate the combination of reduction and fixation technique (overall rating) and then to rate fracture reduction and implant positioning separately. An adequate or inadequate rating for technique, regardless of their opinion of likely clinical outcome, could be assigned. The rating results were divided into 3 groups: overall, reduction and implant position rating. For analysis purposes inadequate rating was defined as an inadequate rating by one or both experts. A sub-analysis was performed of cases where both experts agreed on inadequate technique.

**Regression analysis** A regression analysis with 2 year clinical failure, postoperative complications and mortality as the dependent outcomes and the following co-factors was performed: mobility, accommodation status, femoral neck bone density, cognitive function, ASA class, delay to surgery, fracture angle (Pauwels type), AP fracture reduction, lateral fracture reduction and type of implant.

**Statistical methods** The database of this study was created in an SPSS, version 11.5 (Statistical Package for Social Sciences Incorporated, Chicago IL, USA), spreadsheet format. The distribution of data was tested with a one-sample Kolmogorov-Smirnov test. For normally distributed data, p-values and 95% Confidence Intervals (CI) were calculated with an independent samples t-test. The two-sample Mann-Whitney test was used to calculate significance in abnormally distributed data. A p-value below < 0.05 was considered significant. For non-parametric correlation the Spearman rank co-efficient was calculated. Regression analysis was performed with a binary logistic model.
Results
Internal fixation after closed reduction was performed in 126 patients selected using the Physiologic Status Score protocol. Nine patients received internal fixation, but should have received hemiarthroplasty according to the PSS study protocol (protocol violation). In 9 other patients, in whom there were 6 clinical failures after 2 years, the intraoperative fluoroscopic images, which we deemed essential for proper technical assessment, had not been printed and added to the radiology files. Two patients were lost to follow-up and 4 patients died within 3 months before any clinical follow-up had taken place. For these reasons 24 patients were excluded from the analysis. The results of the remaining 102 patients (34 men) were analysed. For the 3 groups of ratings the cases with adequate and inadequate technical ratings are shown in table 2a-c. Age, sex, pre-injury mobility, accommodation status, bone density of the contralateral femoral neck, cognitive status, ASA class and fracture angle (Pauwels type) showed no differences. Delay to surgery in the overall technique rating group and type of implant in the fixation rating group was different, but regression analysis identified no association with the clinical endpoints. For overall technique an inadequate rating was assigned to 25 (25%) patients. Sixteen

Table 2a. Overall technique. Demographics and baseline characteristics of the adequately and inadequately rated cases. Data are expressed as means (n = 102).

<table>
<thead>
<tr>
<th></th>
<th>Adequate</th>
<th>Inadequate</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number (n)</td>
<td>77</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>77.4</td>
<td>75.4</td>
<td>0.08</td>
</tr>
<tr>
<td>Gender M/F</td>
<td>27 / 50</td>
<td>7 / 18</td>
<td>0.51</td>
</tr>
<tr>
<td>Mobility score</td>
<td>4.8</td>
<td>4.6</td>
<td>0.20</td>
</tr>
<tr>
<td>Accommodation score</td>
<td>4.9</td>
<td>5.0</td>
<td>0.21</td>
</tr>
<tr>
<td>T-score femoral neck</td>
<td>-2.2</td>
<td>2.0</td>
<td>0.28</td>
</tr>
<tr>
<td>Cognition score</td>
<td>4.7</td>
<td>4.8</td>
<td>0.78</td>
</tr>
<tr>
<td>Medical history score</td>
<td>4.3</td>
<td>4.2</td>
<td>0.85</td>
</tr>
<tr>
<td>Delay to Surgery (hours)</td>
<td>26.1</td>
<td>16.2</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Pauwels type</td>
<td>2.5</td>
<td>2.6</td>
<td>0.35</td>
</tr>
</tbody>
</table>

Implant:
Dynamic Hip Screw (DHS)  21       10
DHS with antirotational CS* 33       9     } 0.30
CS                    23       6

*3 cannulated screws
**Table 2b.** Fracture reduction technique. Demographics and baseline characteristics of adequately and inadequately rated cases. Data are expressed as means (n = 102).

<table>
<thead>
<tr>
<th></th>
<th>Adequate</th>
<th>Inadequate</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number (n)</td>
<td>76</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>77.2</td>
<td>77.1</td>
<td>0.87</td>
</tr>
<tr>
<td>Gender M/F</td>
<td>25 / 49</td>
<td>9 / 17</td>
<td>0.82</td>
</tr>
<tr>
<td>Mobility score</td>
<td>4.8</td>
<td>4.6</td>
<td>0.21</td>
</tr>
<tr>
<td>Accommodation score</td>
<td>4.9</td>
<td>4.9</td>
<td>0.56</td>
</tr>
<tr>
<td>T-score femoral neck</td>
<td>-2.2</td>
<td>-2.2</td>
<td>0.85</td>
</tr>
<tr>
<td>Cognition score</td>
<td>4.7</td>
<td>4.9</td>
<td>0.14</td>
</tr>
<tr>
<td>Medical history score</td>
<td>4.3</td>
<td>4.2</td>
<td>0.23</td>
</tr>
<tr>
<td>Delay to Surgery (hours)</td>
<td>24.3</td>
<td>21.6</td>
<td>0.32</td>
</tr>
<tr>
<td>Pauwels type</td>
<td>2.4</td>
<td>2.7</td>
<td>0.09</td>
</tr>
</tbody>
</table>

**Implant:**
- Dynamic Hip Screw (DHS) 21 10 0.31
- DHS with antirotational CS* 32 10
- CS 23 6

*3 cannulated screws

**Table 2c.** Implant positioning technique. Demographics and baseline characteristics of adequately and inadequately cases. Data are expressed as means (n = 102).

<table>
<thead>
<tr>
<th></th>
<th>Adequate</th>
<th>Inadequate</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number (n)</td>
<td>59</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>76.7</td>
<td>77.8</td>
<td>0.38</td>
</tr>
<tr>
<td>Gender M/F</td>
<td>22 / 37</td>
<td>12 / 31</td>
<td>0.32</td>
</tr>
<tr>
<td>Mobility score</td>
<td>4.9</td>
<td>4.7</td>
<td>0.12</td>
</tr>
<tr>
<td>Accommodation score</td>
<td>4.9</td>
<td>4.9</td>
<td>0.68</td>
</tr>
<tr>
<td>T-score femoral neck</td>
<td>-2.2</td>
<td>-2.1</td>
<td>0.39</td>
</tr>
<tr>
<td>Cognition score</td>
<td>4.8</td>
<td>4.7</td>
<td>0.43</td>
</tr>
<tr>
<td>Medical history score</td>
<td>4.2</td>
<td>4.4</td>
<td>0.09</td>
</tr>
<tr>
<td>Delay to Surgery (hours)</td>
<td>25.9</td>
<td>20.7</td>
<td>0.16</td>
</tr>
<tr>
<td>Pauwels type</td>
<td>2.4</td>
<td>2.6</td>
<td>0.06</td>
</tr>
</tbody>
</table>

**Implant:**
- Dynamic Hip Screw (DHS) 13 18 0.03
- DHS with antirotational CS* 26 16
- CS 20 9

*3 cannulated screws
Table 3. Technical rating results in the three groups in relation to fracture reduction criteria and the study endpoints (n = 102).

<table>
<thead>
<tr>
<th>Overall Technique</th>
<th>Adequate</th>
<th>Inadequate</th>
<th>p value</th>
<th>Adequate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>77</td>
<td>25</td>
<td></td>
<td>76</td>
</tr>
<tr>
<td>Garden Index (mean)</td>
<td>166</td>
<td>160</td>
<td>&lt;0.01</td>
<td>165</td>
</tr>
<tr>
<td>Retroversion (mean degrees)</td>
<td>3.2</td>
<td>4.5</td>
<td>0.66</td>
<td>3.2</td>
</tr>
</tbody>
</table>

Endpoints:
- Clinical failure: 23 (30%) vs 16 (64%), p < 0.01
- Mortality: 7 (9%) vs 5 (20%), p = 0.14

Complications:
- None: 64 (83%) vs 14 (56%)
- Minor: 9 (12%) vs 4 (16%), p = 0.23
- Major: 4 (5%) vs 0

# Deep wound infection, pulmonary embolus, cerebro-vascular accident, congestive heart failure or myocardial infarction

out of these 25 patients (64%) proceeded to clinical failure with revision to arthroplasty. For reduction technique and implant positioning, inadequate ratings were assigned to 26 (25%) and 43 (42%) patients; clinical failure occurred in 16 out of 26 (62%) and 21 out of 43 (49%) patients respectively, table 3.

In the overall and reduction technique rating groups inadequate (varus) reduction in the AP view, as expressed by a mean Garden Index below 160°, was associated with significantly more clinical failure. Mortality also appeared higher in all 3 groups of inadequate ratings, but the number of deaths was too small to substantiate this finding. The development of postoperative complications was similar between adequately and inadequately rated patients in all 3 rating groups.

The interrater agreement between the experts on overall technique was good (kappa value 0.65 ± 0.10), agreement on fracture reduction was moderate (kappa value 0.39 ± 0.11) and agreement on implant positioning was poor (kappa value 0.16 ± 0.10).

The correlation between overall technical rating and clinical failure of IF (r = 0.265, p = 0.003) was significant. The highest correlation was found between fracture reduction rating and clinical failure at 2 year follow-up (r = 0.299, p < 0.01). For rating of implant positioning a likely, but not significant, correlation with clinical failure was
found \( r = 0.181, p = 0.07 \), which reflects the poor interrater agreement on implant positioning described above.

In the sub-analysis of patients where both experts agreed on inadequate overall technique the correlation with clinical failure became stronger and rose to 0.453 \( (p < 0.001) \) and 13 out of 15 \( (87\%) \) patients proceeded to clinical failure. When both experts assigned inadequate ratings for fracture reduction the correlation with clinical failure remained similar \( r = 0.294, p < 0.01 \) and 7 out of 8 \( (88\%) \) patients proceeded to clinical failure. When both experts assigned inadequate implant position ratings the correlation with failure did become significant \( r = 0.274, p < 0.01 \) with 9 out of 12 \( (75\%) \) patients proceeding to failure of IF.

After regression analysis only AP (varus) fracture reduction \( \text{OR} 17.5 [95\% \text{ CI}: 1.8 - 167.1], p=0.013 \) was associated with IF failure. None of the other co-factors was associated with the development of complications. Higher ASA class \( \text{OR} 14.2 [95\% \text{ CI}: 2.0 - 100.5], p=0.008 \) and inadequate implant positioning ratings \( \text{OR} 28.8 [95\% \text{ CI}: 1.5 - 523.8], p=0.023 \) had an association with mortality.
Discussion

This prospective 2-year follow-up technical rating study showed that an expert in the field of hip fracture surgery can "predict" clinical IF failure in cases where fracture reduction alone or reduction in combination with implant position is rated inadequate. Rating of implant positioning alone did not correlate with clinical outcome. In future, particularly fracture reduction in both views could objectively be rated in different ways, preferably intraoperative consultation with an expert in the operating room. Alternatives are intraoperative consultation via the intranet or internet, early postoperative local consultation or early postoperative consultation via internet communication. During daytime hours an expert or surgeon with advanced training or accreditation in fracture surgery can easily be consulted locally. The result of this study could prompt surgeons to consider intraoperative consultation or at least early rating of postoperative radiographs by their local "expert". In other circumstances, for instance during off-hours, consultation by communicating radiographs via internet to an expert is a good alternative. If operative technique is rated inadequate at consultation, and particularly if another expert agrees, change of the intraoperative treatment plan or early pre-emptive (re-) intervention should be seriously considered before clinical failure becomes apparent.

Although technical rating was performed according to a protocol based on the recent literature, it remained the opinion of an orthopaedic surgeon and a surgeon-traumatologist being regarded as hip fracture experts, who did not always agree with each other as is expressed by moderate to poor kappa values for reduction and implant positioning rating. This represents a limitation for the reproducibility of the study results and raises the issue of whether one or two experts should be consulted. Two inadequate ratings demonstrated an even stronger correlation with clinical failure than one or two inadequate ratings, although the correlation with clinical failure of the latter model was still significant. Translated into clinical practice the following strategy may be followed: in the event of a single inadequate expert rating a problem may arise in the postoperative clinical course. If a second expert rating is sought and confirms the first, a revisional procedure should be seriously considered. If disagreement occurs pre-emptive revision should still be discussed with the patient.

From the analysis of the previously conducted prospective Physiologic Status Score validation study, it was demonstrated that relatively healthy patients with a high functional demand could withstand revision of IF to arthroplasty without compromising morbidity, mortality or function at 2 year follow-up. A matched-pair
analysis with prospective data recording, however, found higher early dislocation and superficial infection rates in salvage arthroplasty patients compared with primary arthroplasty patients, but found no significant influence on function. These results would support a pre-emptive re-intervention strategy in this group of active patients, in whom most would agree that sparing of the femoral head is preferable. Patients receiving IF should be counselled that revision of IF to a form of arthroplasty may carry a risk of higher early complication rate, mostly superficial infection.

A recognized predictive factor of clinical failure is the quality of reduction, especially in the AP view. Table 3 illustrates the association between inadequate technique, poor reduction in terms of a mean Garden index below 160° and clinical failure. If varus reduction is present on the postoperative radiographs, failure of IF has been reported to be 4 times more likely to fail. This risk increased to 13 times more chance of failure if the surgeon also reported a difficult reduction. If a surgeon encounters a difficult reduction or an uncorrectable varus reduction intraoperatively, a conversion to arthroplasty would be a logical next step. The obstacle herein is the repositioning of the patient from an orthopaedic traction to a standard table and the preparation of new instruments, which prolongs the operative procedure with higher complication risk. A second session of “delayed conversion” after objective technical rating may surmount this hurdle and this approach carries the advantage of time for patient counselling between sessions.

Few studies address the influence of reduction technique in the lateral view. Our study showed no difference in mean retroversion angle in the adequately and inadequately rated groups and no correlation with clinical outcome.

A reduction in ante- or retroversion has been reported to produce inferior results to a neutral lateral view reduction, but this is not a consistent finding. Proper implant positioning is the next important technical step, but can only be done properly after adequate reduction. Meta-analysis data have shown that implants with a screw thread are preferable to pins. As agreement between the expert raters was the poorest on this issue and correlation with clinical outcome was not significant, an inadequate rating of implant positioning alone should probably not have clinical consequences. Poor implant position alone has been reported to lead to more unspecified complications, but others did not find this association.

Internal Fixation for displaced femoral neck fractures should not be delegated to junior, inexperienced staff without supervision. From the present technical rating results found in a daily practice multi-center setting, this recommendation should
have full support because overall technique was rated inadequate in 25 out of 102 cases, which were operated on. Two subanalyses of studies reporting on surgeon experience and clinical outcome in displaced FNF patients found no significant association.\textsuperscript{1,4} Interestingly, one study among 1000 patients suggested that the best results are likely to be gained by surgeons with 1-5 years of experience and a high case load (i.e. minimum 20 procedures per year). A learning curve for junior surgeons and a fall-off rate for more experienced surgeons were noted. After discriminant analysis surgical experience did not have any great effect on mortality or morbidity.\textsuperscript{1}

The present prospective study is the first to describe independent technical rating of internal fixation for displaced femoral neck fractures. The results suggest that early expert inadequate rating is a promising instrument for predicting clinical failure. (Inter)nationally recognized technical guidelines for reduction and internal fixation of these fractures should be developed by which expert surgeons could rate technical performance, which should preferably be performed intraoperatively. An inadequate rating of reduction and/or implant position could lead surgeons to implement a pre-emptive change in therapeutic strategy.

Acknowledgements We thank our technical rating experts Professor R.K. Marti PhD and Professor A.B. van Vugt PhD for their kind assistance.

References


