# Preoperative CT improves the assessment of stability in trochanteric hip fractures

*From Stavanger University Hospital, Stavanger, Norway*

Correspondence should be sent to T. A. Woldeyesus
thomas.abel.woldeyesus@sus.no

Cite this article:
*Bone Jt Open* 2024;5(6):
524–531.

DOI: 10.1302/2633-1462.56.BJO-2023-0177.R1

**T. A. Woldeyesus,[1,2] J-E. Gjertsen,[2,3] I. Dalen,[4,5] T. Meling,[1] M. Behzadi,[6] K. Harboe,[1,2] A. Djuv[1,2]**

[1]Department of Orthopaedic Surgery, Stavanger University Hospital, Stavanger, Norway
[2]Department of Clinical Medicine, University of Bergen, Bergen, Norway
[3]Department of Orthopaedic Surgery, Haukeland University Hospital, Bergen, Norway
[4]Department of Quality and Health Technology, University of Stavanger, Stavanger, Norway
[5]Department of Research, Stavanger University Hospital, Stavanger, Norway
[6]Department of Radiology, Haukeland University Hospital, Bergen, Norway

## Aims
To investigate if preoperative CT improves detection of unstable trochanteric hip fractures.

## Methods
A single-centre prospective study was conducted. Patients aged 65 years or older with trochanteric hip fractures admitted to Stavanger University Hospital (Stavanger, Norway) were consecutively included from September 2020 to January 2022. Radiographs and CT images of the fractures were obtained, and surgeons made individual assessments of the fractures based on these. The assessment was conducted according to a systematic protocol including three classification systems (AO/Orthopaedic Trauma Association (OTA), Evans Jensen (EVJ), and Nakano) and questions addressing specific fracture patterns. An expert group provided a gold-standard assessment based on the CT images. Sensitivities and specificities of surgeons' assessments were estimated and compared in regression models with correlations for the same patients. Intra- and inter-rater reliability were presented as Cohen's kappa and Gwet's agreement coefficient (AC1).

## Results
We included 120 fractures in 119 patients. Compared to radiographs, CT increased the sensitivity of detecting unstable trochanteric fractures from 63% to 70% ($p = 0.028$) and from 70% to 76% ($p = 0.004$) using AO/OTA and EVJ, respectively. Compared to radiographs alone, CT increased the sensitivity of detecting a large posterolateral trochanter major fragment or a comminuted trochanter major fragment from 63% to 76% ($p = 0.002$) and from 38% to 55% ($p < 0.001$), respectively. CT improved intra-rater reliability for stability assessment using EVJ (AC1 0.68 to 0.78; $p = 0.049$) and for detecting a large posterolateral trochanter major fragment (AC1 0.42 to 0.57; $p = 0.031$).

## Conclusion
A preoperative CT of trochanteric fractures increased detection of unstable fractures using the AO/OTA and EVJ classification systems. Compared to radiographs, CT improved intra-rater reliability when assessing fracture stability and detecting large posterolateral trochanter major fragments.

## Take home message
- Patients suffering from a trochanteric hip fracture may benefit from a preoperative CT.
- Preoperative CT increases the orthopaedic surgeon's ability to detect unstable trochanteric hip fractures.

Bone & Joint Open

## Introduction

Radiographs are the gold standard in diagnostics and preoperative planning of hip fractures. Most classification systems for fractures are based on radiographs, enabling orthopaedic surgeons to categorize the fracture as stable or unstable, and to choose the appropriate treatment.[1] The current recommendations are to treat stable trochanteric fractures with sliding hip screws and unstable or subtrochanteric fractures with intramedullary nails.[2,3] There is little evidence to suggest that one surgical method is superior to the other with regard to the treatment outcome of trochanteric fractures.[4,5] However, a recently published study suggests that the use of intramedullary nails reduces reoperation rates compared with sliding hip screws in the treatment of unstable trochanteric fractures.[6] Accordingly, identifying unstable fracture patterns is important to optimize treatment and to reduce risk of reoperation.

The introduction of CT in fracture diagnostics has given orthopaedic surgeons the opportunity to use 3D reconstruction in preoperative planning. CT may provide a different understanding of these fractures and possibly change the interpretation of the stability. Misdiagnosing an unstable fracture as stable may increase complications related to the preferred treatment choice of sliding hip screws for stable fractures.[6]

The most-used classification system for hip fractures is the AO/Orthopaedic Trauma Association (AO/OTA) classification. However, the interobserver agreement of sub-classifications of trochanteric hip fractures (AO31.A1.1-A3.3) has been found to be low, in part explained by the difficulty of discerning posterior trochanteric fragments on radiographs.[7] The Evans Jensen classification system for trochanteric hip fractures is commonly used, based on observations on radiographs.[1] A novel classification system, the modified Nakano, has been developed based on 3D-CT.[8] We aimed to investigate if preoperative CT improves detection of unstable trochanteric hip fractures.

## Methods

This single-centre prospective study was approved by the Regional Committee for Medical Research Ethics Western Norway, REK West (ID 2019/470). Patients aged ≥ 65 years with a trochanteric hip fracture verified on a radiograph were eligible for inclusion and provided written informed consent. Trochanteric hip fractures were defined by the AO/OTA classification system (31A1-31A3).[9]

From 7 September 2020 to 22 January 2022, we prospectively assessed the eligibility of patients with trochanteric hip fractures admitted to Stavanger University Hospital. The enrolment process is specified in Figure 1. Patients were excluded if they lacked a preoperative CT. Reasons for lacking a preoperative CT would either be due to reduced CT capacity, or a short waiting time between admission and surgery. Finally, a total of 120 cases were enrolled in the study.

In total, 47 orthopaedic surgeons in our department participated in the classification of their surgical cases. Preoperatively, the surgeons who were to operate on the patient individually assessed radiographs of the fracture according to a systematic protocol (Supplementary Material). Fractures were classified in accordance with three different classification systems: AO/OTA,[9] Modified Evans Jensen (EVJ),[10] and Modified Nakano.[8] Subsequently, the surgeons answered questions addressing the presence of specific unstable fracture patterns (Supplementary Material). The order of questions was repeated when assessing CT (coronal, sagittal, axial, and 3D reconstruction) of the same fracture. The surgeons could not change their previous answers. After three months, the surgeons reassessed the radiographs and CT images according to the same protocol. The surgeons were instructed to omit reviewing their surgical note before reassessing the fractures at three months. The system used for registration was incorporated into the local surgery planning system, which sent an automatic reminder for the three-month assessment registration. This system was well known to all surgeons participating in the study.

An expert group consisting of two consultant orthopaedic surgeons (JEG and AD) and one consultant radiologist (MB) collectively assessed all fractures. They followed the same systematic protocol as the surgeons (Supplementary Material). Their assessments of the three classification systems and specific fracture pattern observations on CT were defined as the gold standard. In cases of disagreement within the expert group, radiographs and CT images were re-evaluated and discussed until consensus was reached. In contrast to the surgeons, the expert group assessed the fractures postoperatively and did not reassess the fractures after three months. The expert group did not read or receive any information about the surgery before or during their assessment.

Stable fractures were defined as AO/OTA 31A1.1 to A2.1, EVJ 1 to 2, and Nakano 2-part, 3-part G(S), and G(B).[8-10] Unstable fractures were defined as AO/OTA 31A2.2 to A3.3, EVJ 3 to 5, and Nakano 3-part G-L, G(W), L, 4-part, and a modification called 1b (Supplementary Material) illustrating an intertrochanteric fracture line.[8-10]

Missing data were handled by including all available cases for each individual analysis. Number of observations per analysis is available in the tables.

### Statistical analysis

Comparisons of classifications using CT images versus radiographs were assessed in logistic regression models, with CT classification as the dependent variable and radiograph classification as the independent variable, from which we obtained the predicted proportions of CT classifications given the radiograph classification. Given that we had two observers per patient, we applied two-level models allowing for clustering within patients.

Sensitivity and specificity were estimated separately for evaluation of radiographs and CT images using two-level logistic regression, allowing for clustering of data for the same patient (two observers per patient), with observer classification as the dependent variable and gold-standard classification as the independent variable. Predicted probabilities of the observer classifying a fracture as, for example, unstable given gold-standard classification being unstable (sensitivity), and of the observer classifying a fracture as stable given gold-standard classification being stable (specificity), were obtained and presented with 95% confidence intervals (CIs). Differences in sensitivities and specificities using radiographs versus CT images were estimated in common models including data for

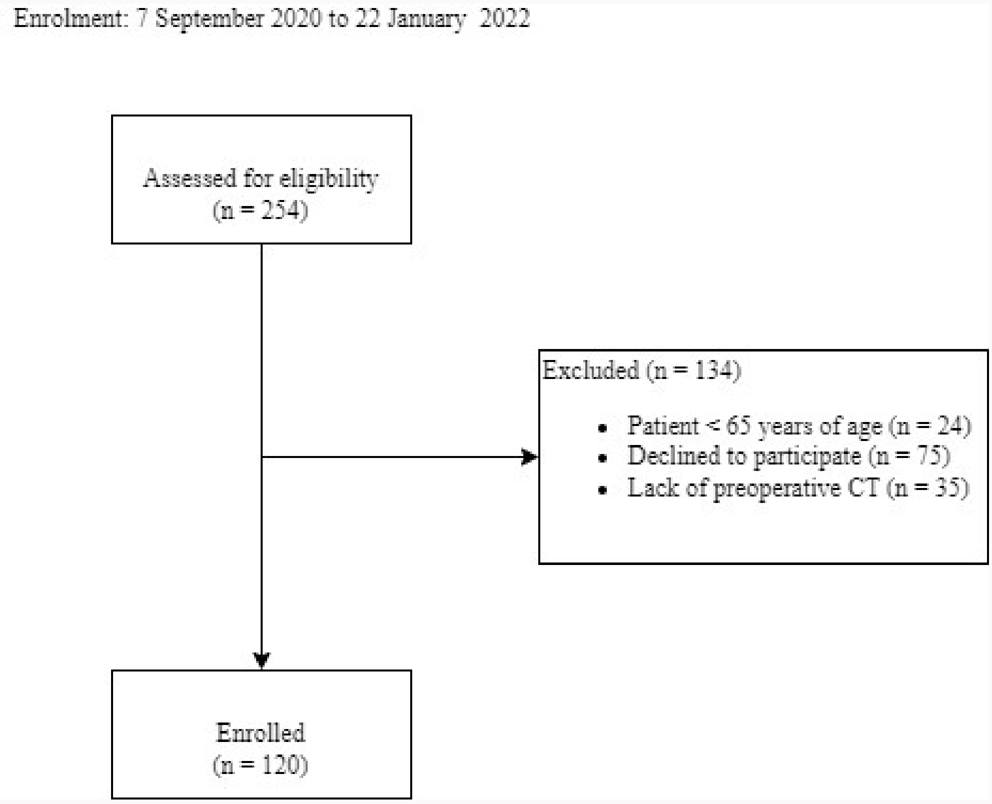


**Fig. 1**
Enrolment: flowchart of patients presenting with trochanteric hip fractures at Stavanger University Hospital, Stavanger, Norway.

both methods, with observer classification as the dependent variable and gold-standard classification, modality, and the interaction between gold standard and modality as independent variables. These models were three-level, allowing for clustering of repeated assessments (one on radiographs and one on CT) from each observer, as well as clustering of observers within patients. The p-values for differences were obtained from contrast tests.

Intra- and interobserver reliability for each modality was estimated using Cohen's kappa coefficient (κ) and Gwet's agreement coefficient (AC1), the latter being preferred over Cohen's kappa in skew marginal distribution situations where Cohen's kappa tends to be biased.[11] CIs for interobserver reliability were bootstrapped percentile intervals (B = 1,000). For intraobserver reliability, for which we had data from two observers per patient, CIs were bootstrapped percentile based on cluster bootstrap (B = 1,000), with patients as clusters. Similarly, p-values for differences between agreement coefficients for radiograph versus CT were obtained by bootstrapping differences (B = 10,000, cluster bootstrap for intraobserver agreement) from the null distribution.[12]

We used Stata version 17 (StataSE 17.0; StataCorp, USA) with functions melogit, margins, kappaetc, bootstrap with and without option cluster, and estat bootstrap.[13] The integration method for melogit was mean-variance adaptive Gauss-Hermite quadrature (the default), however in case of convergence issues for three-level models, other integration methods (plus the option difficult) were explored. Any -values < 0.05 were considered statistically significant. For intra- and interobserver reliability, the level of agreement was defined as slight (0.01 to 0.20), fair (0.21 to 0.40), moderate (0.41 to 0.60), substantial (0.61 to 0.80), and almost perfect (> 0.80).[14]

## Results

### Patient demographic data and surgeon experience

In total, 120 fractures in 119 patients were included. The mean age of the patients was 83 years (SD 7.9). Overall, 82 (68%) of the study participants were female. There were two operating surgeons present at 117 of the 120 operations, and one surgeon at three operations. This resulted in 237 assessments of the 120 fractures. Orthopaedic registrars performed 77 (32%) of the assessments, and orthopaedic consultants 160 (68%).

### Assessment of stability based on CT images versus radiographs

The operating surgeons assessed more fractures as being stable based on radiographs compared with the expert group using all three classification systems (Figures 2a and 2b). After CT evaluation, the surgeons reassessed 21 (20%), 29 (37%), and 16 (19%) of fractures as unstable using AO/OTA, EVJ, and Modified Nakano classification systems, respectively (Figure 2a). Few fractures assessed by the surgeons and expert group as unstable on radiograph were reassessed as stable after CT evaluation (Figures 2c and 2d).

### Sensitivities and specificities for CT images versus radiographs

The surgeons' ability to identify unstable trochanteric fractures on CT images compared with radiographs alone increased from 63% to 70% (p = 0.028) and from 70% to 76% (p

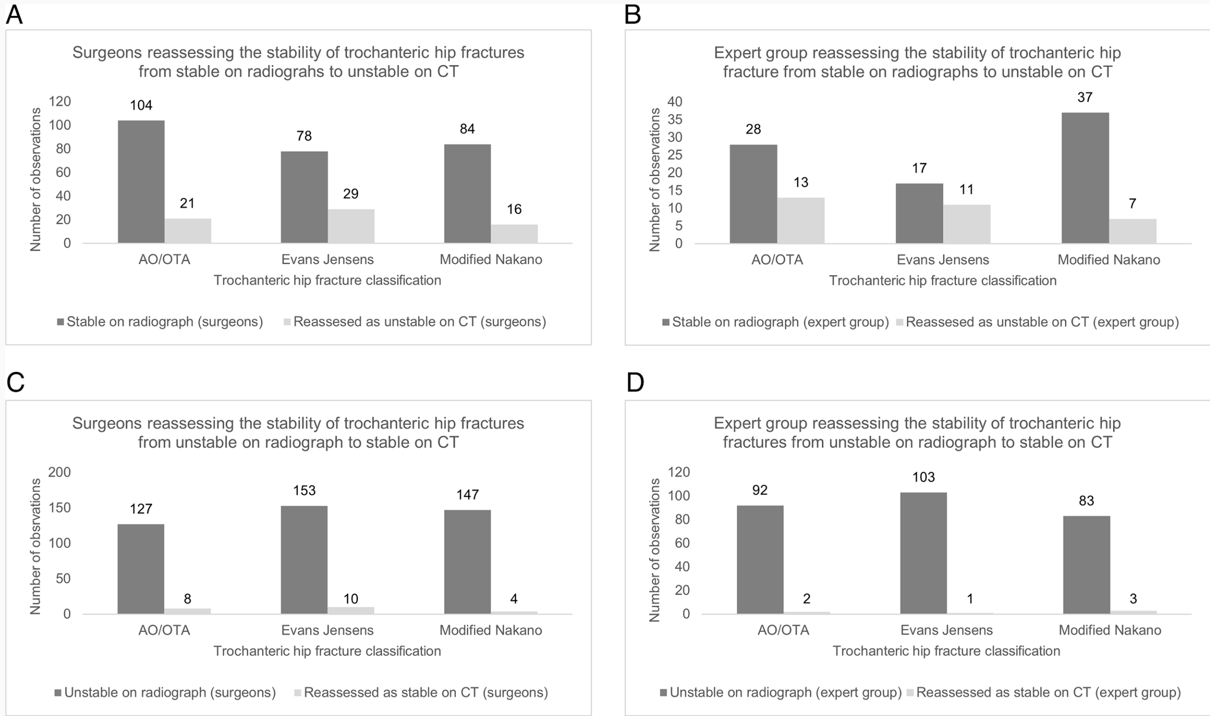


**Fig. 2**
Chart presenting the assessment of trochanteric hip fracture stability using AO/Orthopaedic Trauma Association (OTA), Evans Jensen, and modified Nakano classification. Stable fractures were defined as AO/OTA 31A1.1 – A2.1, Evans Jensen 1 to 2, and Nakano 2-part, 3-part G(S), and G(B). Unstable fractures were defined as AO/OTA 31A2.2 – A3.3, Evans Jensen 3 to 5, and Nakano 3-part G-L, G(W), L, 4-part, and a modification called 1b illustrating an intertrochanteric fracture line. a) The surgeons reassessed 20%, 37%, and 19% of fractures as unstable on CT using AO/OTA, Evans Jensens, and Modified Nakano classification systems, respectively. b) The expert group reassessed 46%, 65%, and 19% of fractures as unstable on CT using AO/OTA, Evans Jensens, and Modified Nakano classification systems, respectively. c) The surgeons reassessed 6%, 7%, and 3% of fractures as stable on CT using AO/OTA, Evans Jensens, and Modified Nakano classification systems, respectively. d) The expert group reassessed 2%, 1%, and 4% of fractures as stable on CT using AO/OTA, Evans Jensens, and Modified Nakano classification systems, respectively.

**Table I.** The surgeons' classification of trochanteric fractures as unstable or stable using three classification systems on radiographs and CT images compared to the gold standard defined by an expert group.

| Classification system | Unstable fractures* | Sensitivity (95% CI) | | p-value | Stable fractures* | Specificity (95% CI) | | p-value |
|---|---|---|---|---|---|---|---|---|
| | | Radiograph | CT | | | Radiograph | CT | |
| AO/OTA | 103 | 0.63 (0.55 to 0.71) | 0.70 (0.62 to 0.77) | 0.028 | 17 | 0.97 (0.90 to 1.00) | 0.96 (0.89 to 1.00) | > 0.999 |
| Modified Evans Jensen | 113 | 0.70 (0.63 to 0.76) | 0.76 (0.71 to 0.82) | 0.004 | 7 | 0.92 (0.74 to 1.00) | 0.84 (0.59 to 1.00) | 0.384 |
| Modified Nakano | 87 | 0.83 (0.77 to 0.90) | 0.89 (0.84 to 0.94) | 0.114 | 33 | 0.88 (0.78 to 0.98) | 0.86 (0.76 to 0.95) | 0.276 |

*Per gold standard, i.e. as evaluated by the expert group using the CT images.
AO/OTA, AO/Orthopaedic trauma Association; CI, confidence interval.

= 0.004) using AO/OTA and EVJ, respectively (Table I). No statistically significant difference in ability to identify unstable fractures was found for the Nakano classification. We did not find significant differences in specificity for detecting stable fractures for any of the classification systems using CT (Table I).

The surgeons' ability to detect a large posterolateral trochanter major fragment or a comminuted trochanter major on a CT image increased from 63% to 76% (p = 0.002) and from 38% to 55% (p < 0.001), respectively, compared with using radiograph alone (Table II). No statistically significant differences were found for the detection of intertrochanteric fracture lines, thin lateral walls, or large trochanter minor fragments, nor for the exclusion of non-existing conditions (i.e. specificities).

Intra-rater reliability was moderate to substantial (Table III), whereas inter-rater reliability was generally lower, in part falling into the category "fair" (Table IV). The intra-rater reliability for detection of instability on CT images compared with radiographs increased when the fractures were classified using EVJ (AC1 0.68 to 0.78; p = 0.049) and for detecting a large posterolateral trochanter major fragment (AC1 0.42 to 0.57; p = 0.031) (Table III).

There were no significant differences between radiographs and CT images with regard to inter-rater reliability for stability assessments, nor for detection of specific

**Table II.** Surgeons' ability to detect specific fracture morphologies on radiographs and CT images compared to the gold standard defined by an expert group.

| | | Sensitivity (95% CI) | | | | Specificity (95% CI) | | |
|---|---|---|---|---|---|---|---|---|
| Morphology | Present* | Radiograph | CT | p-value | Not present* | Radiograph | CT | p-value |
| Intertrochanteric fracture line | 30 | 0.81 (0.71 to 0.92) | 0.85 (0.75 to 0.94) | 0.566 | 90 | 0.77 (0.70 to 0.83) | 0.74 (0.68 to 0.81) | 0.532 |
| Thin lateral wall | 97 | 0.56 (0.48 to 0.63) | 0.60 (0.53 to 0.68) | 0.143 | 23 | 0.72 (0.58 to 0.86) | 0.82 (0.69 to 0.94) | 0.108 |
| Large posterolateral trochanter major fragment | 72 | 0.63 (0.53 to 0.72) | 0.76 (0.68 to 0.84) | 0.002 | 48 | 0.70 (0.59 to 0.80) | 0.65 (0.55 to 0.76) | 0.381 |
| Comminuted trochanter major fracture | 74 | 0.38 (0.29 to 0.46) | 0.55 (0.46 to 0.64) | < 0.001 | 46 | 0.64 (0.53 to 0.75) | 0.60 (0.49 to 0.71) | 0.488 |
| Large trochanter minor fragment† | 59 | 0.52 (0.42 to 0.62) | 0.52 (0.42 to 0.62) | > 0.999 | 61 | 0.97 (0.93 to 1.00) | 0.97 (0.95 to 1.00) | 0.508 |

*Per gold standard, i.e. as evaluated by the expert group using the CT images.
†Defined as a displaced or undisplaced large trochanter minor fragment extending proximal or distal of minor.
CI, confidence interval.

**Table III.** Intra-rater agreement comparing preoperative and three-month classifications for radiograph and CT.

| | Radiograph (n = 207) | | | CT (n = 205) | | | | |
|---|---|---|---|---|---|---|---|---|
| Variable | Observed agreement | κ (95% CI) | AC1 (95% CI) | Observed agreement | κ (95% CI) | p-value | AC1 (95% CI) | p-value |
| **Stability** | | | | | | | | |
| AO/OTA | 75% | 0.50 (0.35 to 0.62) | 0.52 (0.39 to 0.64) | 80% | 0.58 (0.45 to 0.68) | 0.224 | 0.62 (0.51 to 0.73) | 0.108 |
| Modified Evans Jensen | 82% | 0.60 (0.47 to 0.71) | 0.68 (0.56 to 0.78) | 86% | 0.63 (0.50 to 0.74) | 0.598 | 0.78 (0.70 to 0.86) | 0.049 |
| Modified Nakano | 85% | 0.67 (0.57 to 0.77) | 0.72 (0.63 to 0.82) | 84% | 0.63 (0.50 to 0.75) | 0.520 | 0.73 (0.63 to 0.82) | 0.850 |
| **Morphologies** | | | | | | | | |
| Intertrochanteric fracture line | 79% | 0.55 (0.43 to 0.66) | 0.61 (0.50 to 0.71) | 84% | 0.66 (0.55 to 0.76) | 0.144 | 0.69 (0.59 to 0.78) | 0.221 |
| Thin lateral wall | 81% | 0.61 (0.50 to 0.72) | 0.61 (0.50 to 0.73) | 79% | 0.57 (0.45 to 0.68) | 0.236 | 0.57 (0.46 to 0.69) | 0.286 |
| Large posterolateral trochanter major fragment | 71% | 0.42 (0.29 to 0.54) | 0.42 (0.30 to 0.55) | 77% | 0.51 (0.39 to 0.63) | 0.205 | 0.57 (0.45 to 0.68) | 0.031 |
| Comminuted trochanter major fracture | 76% | 0.50 (0.39 to 0.62) | 0.55 (0.44 to 0.66) | 76% | 0.51 (0.41 to 0.61) | 0.984 | 0.51 (0.41 to 0.62) | 0.610 |
| Large trochanter minor fragment* | 86% | 0.65 (0.52 to 0.77) | 0.78 (0.68 to 0.87) | 86% | 0.65 (0.53 to 0.76) | 0.886 | 0.76 (0.66 to 0.85) | 0.553 |

Agreement coefficients with 95% percentile bootstrap confidence interval (B = 1,000, cluster bootstrap allowing for clustering within patients). p-values for difference between agreement for CT versus radiograph obtained from clustered bootstrap resampling with B = 10,000.
*Defined as a displaced or undisplaced large trochanter minor fragment extending proximal or distal of minor.
AC1, Gwet's agreement coefficient; AO/OTA, AO/Orthopaedic Trauma Association; CI, confidence interval; κ, Cohen's kappa coefficient.

fracture morphologies at the initial evaluation (Table IV). At the three-month classification, the inter-rater reliability for detection of instability on CT images compared with radiographs increased for detecting a large posterolateral trochanter major fragment (AC1 0.28 to 0.56; p = 0.005) (Supplementary Table i).

## Discussion

The operating surgeons classified more trochanteric fractures as stable based on radiographs compared with the expert group. CT significantly increased the surgeons' ability to correctly classify the trochanteric fractures as unstable using both AO/OTA and Modified EVJ classification systems. Their ability to correctly identify the presence of specific unstable fracture patterns of a large posterolateral trochanter major fragment or a comminuted trochanter major increased using CT. Furthermore, CT increased the intraobserver reliability for stability (EVJ classification) and for detecting a large posterolateral trochanter major fragment. For the comparison of interobserver reliability between the two imaging methods,

**Table IV.** Inter-rater agreement for preoperative classifications for radiograph and CT.

| | | Radiograph (n = 111) | | | CT (n = 111) | | | |
|---|---|---|---|---|---|---|---|---|
| Variable | Observed agreement | κ (95% CI) | AC1 (95% CI) | Observed agreement | κ (95% CI) | p-value | AC1 (95% CI) | p-value |
| **Stability** | | | | | | | | |
| AO/OTA | 74% | 0.47 (0.30 to 0.64) | 0.48 (0.32 to 0.64) | 76% | 0.49 (0.30 to 0.65) | 0.860 | 0.53 (0.37 to 0.68) | 0.566 |
| Modified Evans Jensen | 77% | 0.50 (0.31 to 0.65) | 0.59 (0.43 to 0.73) | 84% | 0.57 (0.39 to 0.74) | 0.510 | 0.74 (0.60 to 0.84) | 0.066 |
| Modified Nakano | 84% | 0.65 (0.50 to 0.79) | 0.70 (0.56 to 0.83) | 82% | 0.58 (0.41 to 0.73) | 0.361 | 0.68 (0.53 to 0.81) | 0.834 |
| **Morphologies** | | | | | | | | |
| Intertrochanteric fracture line | 68% | 0.31 (0.12 to 0.47) | 0.39 (0.20 to 0.55) | 69% | 0.37 (0.19 to 0.53) | 0.562 | 0.41 (0.24 to 0.57) | 0.719 |
| Thin lateral wall | 61% | 0.23 (0.04 to 0.40) | 0.23 (0.06 to 0.41) | 61% | 0.22 (0.04 to 0.40) | 0.984 | 0.23 (0.05 to 0.40) | 0.976 |
| Large posterolateral trochanter major fragment | 72% | 0.44 (0.28 to 0.59) | 0.44 (0.29 to 0.60) | 70% | 0.38 (0.20 to 0.55) | 0.561 | 0.42 (0.26 to 0.60) | 0.922 |
| Comminuted trochanter major fracture | 66% | 0.27 (0.07 to 0.44) | 0.36 (0.17 to 0.54) | 63% | 0.26 (0.10 to 0.45) | 0.896 | 0.26 (0.11 to 0.45) | 0.407 |
| Large trochanter minor fragment* | 77% | 0.41 (0.21 to 0.60) | 0.61 (0.46 to 0.75) | 77% | 0.43 (0.23 to 0.61) | 0.854 | 0.63 (0.47 to 0.76) | 0.733 |

Agreement coefficients with 95% percentile bootstrap confidence interval (B = 1,000). p-values for difference between agreement for CT versus radiograph obtained from bootstrap resampling (B = 10,000).
*Defined as a displaced or undisplaced large trochanter minor fragment extending proximal or distal of minor.
AC1, Gwet's agreement coefficient; AO/OTA, AO/Orthopaedic Trauma Association; CI, confidence interval; κ, Cohen's kappa coefficient.

there were conflicting results between the preoperative and three-months assessments.

The present study agrees with an observational study evaluating the stability of trochanteric fractures using 3D-CT imaging that showed increased intra- and interobserver agreement compared with radiographs (AO/OTA and modified EVJ classification systems).[15] Another study evaluated the stability of trochanteric fractures intraoperatively (arthroplasty, open surgery) to assess the gold-standard AO classification and compared their evaluation with preoperative radiographs and 2D-CT images.[16] They found moderate agreement for stability using radiographs, whereas CT had excellent agreement for stability assessment.[16] However, another study using CT to classify trochanteric fractures found no significant increase in inter-rater agreement (AO/OTA and EVJ classification).[17]

Our study differs from the above-mentioned studies in terms of methodology. We engaged all orthopaedic surgeons as participants in the classification of their surgical cases. We are not aware of any study on the classification of trochanteric fractures that uses a similar approach, nor have we found any studies addressing detection of isolated predefined unstable fracture patterns on trochanteric fractures where they are compared based on different imaging methods.

We consider the chosen methodology of mimicking a normal clinical setting as a strength of this study. In real life, the surgeons treating these fractures may have considerable differences in expertise. Our results show a discrepancy in the interpretation of stability between the expert group and the operating surgeons. The latter classified fewer fractures as unstable based on the radiograph evaluation, which may be due to their varied experience in interpreting unstable fracture patterns in general. The expert group classified most of the fractures as unstable when assessing the radiographs.

The correct interpretation of trochanteric fracture stability may be limited by the predefined classification systems we have included in this study. It would have been interesting to include the simple question, "Do you consider the fracture to be stable or unstable?" This could have identified discrepancies in the interpretation of stability between the surgeons and the expert group, independent of any of the classification systems. We recommend including this question in future studies assessing the stability of fractures.

We did not find the modified Nakano classification to be applicable in the context of comparing its use with radiographs and CT. This classification system is based on CT evaluation, and the observer point is the posterior aspect of the proximal femur.[8] It is difficult to assess the posterior aspect of proximal femur on radiographs, which could explain why the results for the Nakano classification did not reach statistical significance.

The clinical relevance of this study's findings can be applied to the consequences of defining more fractures as unstable using CT. A study recently published by the Norwegian Hip Fracture Register supports the use of intramedullary nails for unstable trochanteric hip fractures (AO/OTA 31.A2-A3 and subtrochanteric fractures combined) compared with sliding hip screws in terms of reduced reoperation rate.[6] It also found that the one-year mortality rate was lower in the intramedullary nail group.[6] It may therefore be argued that introducing CT in preoperative diagnostics may further increase the use of intramedullary nails to treat these fractures, especially in cases where there is uncertainty about the stability. Plain radiographs are still advocated as the primary method of investigation for hip fractures.[2] Introducing a preoperative CT as a routine part of clinical practice for these fractures will have a considerable impact on the use of CT resources. Hospitals with good CT availability report that a preoperative CT may substitute the use of radiographs without delaying surgery.[18] However, at hospitals with more limited access to CT, there is a risk that a preoperative CT may postpone surgery, which is not desirable. In conducting this study, we made sure that preoperative CT did not delay the time to theatre. Questions can also be raised concerning potential systematic errors in stability evaluations of trochanteric fractures in previous clinical and register studies where decisions on the use of classification systems, such as AO/OTA and EVJ, have been made based on radiographs.

A limitation of the study is its single-centre design, with a rather small sample size. Furthermore, the unbalanced proportion of stable versus unstable fractures defined as the gold standard by the expert group results in imprecise estimates of specificities, and low power for the corresponding comparisons. When planning the study, we expected more fractures to be classified as stable by the expert group. Due to the length of time required to assess all fractures, the expert group defined the gold standard for all 120 fractures on one occasion. They did not make a three-month reassessment of the fractures, therefore we could not assess the expert group's intraobserver reliability. This may be considered a limitation.

In conclusion, CT images of trochanteric fractures increased detection of unstable fractures using the AO/OTA and EVJ classification systems. Compared with radiographs, CT improved intra-rater reliability when assessing fracture stability and the detection of large posterolateral trochanter major fragments.

## Supplementary material

Questionnaire provided to the surgeons/expert group, and description of the expert group's definitions of the specific unstable trochanteric fracture patterns.

## References


1. **Jensen JS**. Classification of trochanteric fractures. *Acta Orthop Scand*. 1980;51(5):803–810.
2. **No authors listed**. Clinical guideline [CG124] Hip fracture: management. National Institue for Health and Care Excellence (NICE). 2023. https://www.nice.org.uk/guidance/cg124/chapter/Recommendations (date last accessed 13 June 2024).
3. **Baldock TE**, **Dixon JR**, **Koubaesh C**, **Johansen A**, **Eardley WGP**. Variation of implant use in A1 and A2 trochanteric hip fractures: a study from the National Hip Fracture Database of England and Wales. *Bone Jt Open*. 2022;3(10):741–745.
4. **Lewis SR**, **Macey R**, **Lewis J**, **et al**. Surgical interventions for treating extracapsular hip fractures in older adults: a network meta-analysis. *Cochrane Database Syst Rev*. 2022;2(2):CD013405.
5. **Schemitsch EH**, **Nowak LL**, **Schulz AP**, **et al**. Intramedullary nailing vs sliding hip screw in trochanteric fracture management: the INSITE randomized clinical trial. *JAMA Netw Open*. 2023;6(6):e2317164.
6. **Grønhaug KML**, **Dybvik E**, **Matre K**, **Östman B**, **Gjertsen JE**. Intramedullary nail versus sliding hip screw for stable and unstable trochanteric and subtrochanteric fractures: 17,341 patients from the Norwegian Hip Fracture Register. *Bone Joint J*. 2022;104-B(2):274–282.
7. **Sharma G**, **Gn KK**, **Khatri K**, **Singh R**, **Gamanagatti S**, **Sharma V**. Morphology of the posteromedial fragment in pertrochanteric fractures: a three-dimensional computed tomography analysis. *Injury*. 2017;48(2):419–431.
8. **Shoda E**, **Kitada S**, **Sasaki Y**, **et al**. Proposal of new classification of femoral trochanteric fracture by three-dimensional computed tomography and relationship to usual plain X-ray classification. *J Orthop Surg (Hong Kong)*. 2017;25(1):2309499017692700.
9. **Marsh JL**, **Slongo TF**, **Agel J**, **et al**. Fracture and dislocation classification compendium - 2007: Orthopaedic Trauma Association classification, database and outcomes committee. *J Orthop Trauma*. 2007;21(10 Suppl):S1–133.
10. **Jensen JS**, **Michaelsen M**. Trochanteric femoral fractures treated with McLaughlin osteosynthesis. *Acta Orthop Scand*. 1975;46(5):795–803.
11. **Feinstein AR**, **Cicchetti DV**. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol*. 1990;43(6):543–549.
12. **Efron B**, **Tibshirani RJ**. *An Introduction to the Bootstrap*. 1st ed. Chapman and Hall/CRC, 1994.
13. **Klein D**. KAPPAETC: Stata Module to Evaluate Interrater Agreement, 2022. https://ideas.repec.org/c/boc/bocode/s458283.html (date last accessed 17 June 2024).
14. **Landis JR**, **Koch GG**. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174.
15. **Cho YC**, **Lee PY**, **Lee CH**, **Chen CH**, **Lin YM**. Three-dimensional CT improves the reproducibility of stability evaluation for intertrochanteric fractures. *Orthop Surg*. 2018;10(3):212–217.
16. **Isida R**, **Bariatinsky V**, **Kern G**, **Dereudre G**, **Demondion X**, **Chantelot C**. Prospective study of the reproducibility of X-rays and CT scans for assessing trochanteric fracture comminution in the elderly: a series of 110 cases. *Eur J Orthop Surg Traumatol*. 2015;25(7):1165–1170.
17. **Cavaignac E**, **Lecoq M**, **Ponsot A**, **et al**. CT scan does not improve the reproducibility of trochanteric fracture classification: a prospective observational study of 53 cases. *Orthop Traumatol Surg Res*. 2013;99(1):46–51.
18. **Liborg IK**, **Sorteberg A**, **Larsen ASF**, **Gleditsch J**, **Pettersen PM**. [CT as primary diagnosis for hip fracture]. Tidsskriftet. 2023. https://tidsskriftet.no/2023/08/debatt/ct-som-primaerdiagnostikk-ved-mistenkt-hoftebrudd (date last accessed 13 June 2024). [Website in Norwegian].


## Author information


**T. A. Woldeyesus**, MD, Orthopaedic Registrar, PhD Student
**K. Harboe**, MD, PhD, Orthopaedic Consultant, Associate Professor
**A. Djuv**, MD, PhD, Orthopaedic Consultant, Associate Professor
Department of Orthopaedic Surgery, Stavanger University Hospital, Stavanger, Norway; Department of Clinical Medicine, University of Bergen, Bergen, Norway.

**J-E. Gjertsen**, MD, PhD, Orthopaedic Consultant, Professor, Department of Clinical Medicine, University of Bergen, Bergen, Norway; Department of Orthopaedic Surgery, Haukeland University Hospital, Bergen, Norway.

**I. Dalen**, PhD, Statistician, Associate Professor, Department of Quality and Health Technology, University of Stavanger, Stavanger, Norway; Department of Research, Stavanger University Hospital, Stavanger, Norway.

**T. Meling**, MD, PhD, Orthopaedic Consultant, Department of Orthopaedic Surgery, Stavanger University Hospital, Stavanger, Norway.

**M. Behzadi**, MD, Radiologist Consultant, Department of Radiology, Haukeland University Hospital, Bergen, Norway.


## Author contributions

T. A. Woldeyesus: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Validation, Writing – original draft, Writing – review & editing.

J-E. Gjertsen: Data curation, Investigation, Supervision, Validation, Writing – original draft, Writing – review & editing.
I. Dalen: Data curation, Formal analysis, Methodology, Validation, Writing – original draft, Writing – review & editing.
T. Meling: Investigation, Validation, Writing – review & editing.
M. Behzadi: Data curation, Investigation, Methodology, Writing – review & editing.
K. Harboe: Data curation, Methodology, Software, Writing – review & editing.
A. Djuv: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing.

## Funding statement
The authors disclose receipt of the following financial or material support for the research, authorship, and/or publication of this article: T. A. Woldeyesus has been awarded a PhD fellowship financed by the Helse Vest research fund (2020) and was granted funding from the Stavanger University Hospital Reseach Grant (2021), providing funding to this study. The fracture registry of Western Norway has also provided funding for this study. No commercial benefits have been received or will be received by any of the authors during the course of the study investigation or publication.

## ICMJE COI statement
T. A. Woldeyesus reports a PhD Fellowship from Helse Vest, and Stavanger University Hospital Research Grant (2021). A. Djuv reports funding from the Fracture registry of Western Norway for this study, and support for attending meetings and/or travel from Stavanger University Hospital and participation on a quality registry and data safety monitoring board at Stavanger University Hospital, unrelated to this study. A. Djuv is also Head of the Fracture registry of Western Norway, and the unpaid Head of the fragility fracture advisory group in the Norwegian Orthopedic Association. J-E. Gjertsen reports payment for lectures from Ortomedic (the Norwegian supplier of DePuy Synthes) and LINK Norway, unrelated to this study.

## Data sharing
The data of this article is not available for public access due to data protection regulation. Data collected for this study has only been approved to be processed by the authors of this article.

## Ethical review statement
Ethical approval for this study was given by the Regional Committee for Medical Research Ethics Western Norway, REK West (ID 2019/470).

## Open access funding
Open access funding for this article was provided by the University of Bergen.

© 2024 Woldeyesus et al. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives (CC BY-NC-ND 4.0) licence, which permits the copying and redistribution of the work only, and provided the original author and source are credited. See https://creativecommons.org/licenses/by-nc-nd/4.0/