Simo Nortunen

STABILITY ASSESSMENT OF ISOLATED LATERAL MALLEOLAR SUPINATION-EXTERNAL ROTATION-TYPE ANKLE FRACTURES
SIMO NORTUNEN

STABILITY ASSESSMENT OF ISOLATED LATERAL MALLEOLAR SUPINATION-EXTERNAL ROTATION-TYPE ANKLE FRACTURES

Academic Dissertation to be presented with the assent of the Doctoral Training Committee of Health and Biosciences of the University of Oulu for public defence in Auditorium 1 of Oulu University Hospital (Kajaanintie 50), on 2 February 2018, at 12 noon

UNIVERSITY OF OULU, OULU 2018
Nortunen, Simo, Stability assessment of isolated lateral malleolar supination-external rotation-type ankle fractures.
University of Oulu Graduate School; University of Oulu, Faculty of Medicine; Medical Research Center Oulu; Oulu University Hospital
*Acta Univ. Oul. D 1447, 2018*
University of Oulu, P.O. Box 8000, FI-90014 University of Oulu, Finland

**Abstract**

Isolated lateral malleolar supination-external rotation (SER) -type ankle fractures without incongruity on the standard radiographs can be either stable or unstable depending on the status of the deep deltoid ligament. Neither clinical signs of injury on the medial aspect of the ankle nor the displacement of fibular fracture on mortise radiographs seem to predict stability. Therefore, stress testing has been adopted in clinical use. No “gold standard” method exists but the manual external rotation (ER) stress test is the most extensively studied. The ER stress test has some disadvantages, and other methods—such as gravity stress radiography and magnetic resonance imaging—have been suggested instead. However, the evidence to support the use of these methods is still insufficient.

The aims of this dissertation were to assess the roles of (1) morphological factors from standard radiographs of 286 patients, (2) clinical findings on the medial side of the ankle and gravity stress radiography of 79 patients, and (3) MRI of 61 patients in evaluating the stability of the ankle mortise in patients with unimalleolar SER-type fractures with no talar shift on standard radiographs. The ER stress test result was considered to be the reference for stability throughout these studies.

We found that a fracture line width < 2 mm in lateral radiographs, only two fracture fragments, and female sex are independent factors predicting a stable ankle mortise. Neither clinical signs on the medial side of the ankle nor gravity stress radiography alone predict the stability of the ankle mortise accurately. According to our MRI findings, total tears of the deep deltoid ligaments are rare, and partial tears are common in this this patient group. The reliability of the MRI assessment is only moderate.

In conclusion, patients with non-committed fractures and < 2 mm displacement on lateral radiographs have stable ankle mortises and need no further stress testing. The gravity stress radiography is an accurate test for the evaluation of the ankle mortise stability only if the clinical signs indicate a similar result with the gravity stress radiographs. The use of MRI provides no additional benefit compared to ER stress testing for stability evaluation of an SER-type ankle fracture.

**Keywords**: ankle fracture, clinical examination, clinical signs, external rotation stress test, gravity stress radiography, magnetic resonance imaging, stability evaluation, supination-external rotation
Supinaatio-ulkokiertomekanismilla syntyneet ulkokehräsluun murtumat ilman röntgenkuvassa näkyvää telaluun siirtymää voivat olla olla joko vakaita tai epävakaita nilkan sisemman niveliteen syvän lehden tilasta riippuen. Klinisessä tutkimuksessa todettujen nilkan sisäreunan vamman merkkien tai röntgenkuvauksella todettavan ulkokehräsluun murtuman virheasennon ei ole osoitettu ennustavan nivelhaarukan mahdollista epävakautta, joten nilkkaa kuormittavien tehväiviä röntgenkuvauksia on otettu kliiniiseen käyttöön. Mikään näistä kuvausmenetelmistä ei ole niin sanottu kultainen standardi, mutta ulkokiertovääntötesti (ER-testi) on tutkittu laajimmin. ER-testin käyttöön liittyviä ongelmia, joiden vuoksi niin sanottua painovoimakuvausta tai muun muassa magneettikuvausta (MRI) on ehdotettu käyttääväksi sen sijaan. Näiden menetelmien käyttöä tukeva tieteilillinen näyttö on kuitenkin vielä riittämätön.

Tämän väitöskirjatyön tarkoituksen oli tutkia (1) 286 potilaan tavallisista kuormittamattomista röntgenkuvista morfologisten tekijöiden, (2) 79 potilaan nilkan sisäreunan kliinisellä tutkimuksesta ja painovoimakuvauksesta sekä (3) 61 potilaalla MRI:n merkitystä ja tarkkuutta arvioitussa supinaatio-ulkokiertomekanismilla syntyneiden ulkokehräsluun murtumien vakautta. ER-testin tulosta käytettiin referenssinä nivelhaarukan vakaudelle kaikissa osatöissä.


Yhteenvetona voidaan todeta, että ilman ilmeistä telaluun siirtymää röntgenkuvassa yksinkertaiset supinaatio-ulkokiertomekanismilla syntyneet ulkokehräsluun murtumat ovat vakaita eikä nivelhaarukan vakauden testaaminen ole tarpeen, jos murtumaraon leveys sivukuvassa on < 2 mm. Painovoimakuvauksia on luotettava, mikäli sen tulos on sama ulkoisten vanman merkkien kanssa. Magneettikuvausten tulkinnan luotettavuus ei ole hyötyä arvioitaessa tämän nilkkamurtumatyypin vakautta.

**Asiakset:** kliininen tutkiminen, kliiniset vanman merkit, magneettikuvaus, nilkkamurtuma, painovoimakuvaukset, supinaatio-ulkokierto, ulkokiertovääntötesti, vakauden arviointi
To my family
Acknowledgements

This study was carried out at the Department of Surgery, University of Oulu, during the years 2012 to 2017.

I wish to express my sincere gratitude to my supervisor Docent Tapio Flinkkilä, M.D., for his guidance during these years, for his comprehensive knowledge and passionate interest in clinical research that have been crucial for this work to ever see the light of day. Tapio, your patient support has been priceless to me. I also owe my warmest thanks to my co-supervisors Docent Harri Pakarinen, M.D. and Docent Jukka Ristiniemi, M.D. It has been a privilege to work with both of you. Harri, your inspiration and “the Sicilian offer” was the starting point for this work but your continuous support and knowledge in the topic have been essential during these years. Jukka, your supportive attitude during this project, but also your guidance and expertise in lower limb orthopedics have been invaluable to me.

I thank Docent Kari Haukipuro, M.D., the Head of the Division of Operative Care, Professor Juhana Leppilaiti, M.D., Head of the Department of Orthopaedic Surgery, and Susanna Yli-Luukko, M.D., Head of the Department of Neurosurgery, Orthopaedics and Traumatology, for the opportunity to work and do research in the Department of Surgery. I also thank Professor Petri Lehenkari, M.D., the chairman of my Ph.D. study follow-up group, for his support during this work.

I warmly thank the reviewers Professor Hannu Aro, M.D., and Docent Rami Madanat, M.D., for their perceptive comments and constructive criticism. I want to thank Pasi Ohtonen, M.Sc., for statistical assistance and commitment throughout this study.

For significant collaboration and participation in this work, my sincere thanks go to my co-authors Heidi Haapasalo, M.D., Ph.D., Tero Kortekangas, M.D., Ph.D., Iikka Lantto, M.D., Ph.D., Sannamari Lepojärvi, M.D., Ph.D., Hannu-Ville Leskelä, M.D., Ph.D., Professor Jaakko Niinimäki, M.D. and Olli Savola, M.D., Ph.D. Your efforts were crucial for completing this study.

I want to thank Docent Pekka Hyvönen, M.D., Deputy Chief of the Department of Orthopaedic Surgery, a member of my Ph.D. study follow-up group and Ottoveli, for your guidance and continuous support throughout this work and my clinical career.

I would like to acknowledge Tuula Rauma, Niina Siira and Leena Rahko, for your assistance in both patient evaluation and invitation. I also owe my thanks to
all residents and consultant surgeons that assisted with recruiting of the patients. Of course, I am sincerely grateful to all the patients who gave their time and effort to participate in this study. I also want to thank all my great colleagues, with whom I have had the change to work during my career.

My warm thanks go to all my dear friends and their families, especially to Matti Vähäkari, Hanna Jaula, Eveliina Ronkainen, Jari Kulju, Samuli Lepojärvi and Ville Suorsa, for not helping me with this project but keeping my focus on more important things in life.

I owe my warmest thanks to my parents Elsa Palolahti and Kari Nortunen, for their love and support throughout my life. Mom, especially your continuous and unconditional help during these still ongoing a-lot-of-things-happening-at-the-same-time-years has been invaluable. I am also grateful to my siblings Sari Mattila and Jari Nortunen and their families, for being in my life. Your support is something I have always been able to count on.

Finally, my deepest gratitude and my heart go to my beloved wife and life-companion Minna, and to our fantastic boys Nuutti, Leo and Valtti. I love you guys. Thank you for being the best people in my life. Minna, without our support this project would have never come to its conclusion. I am most grateful to you my love – for everything.

This study was financially supported by the Oulu University Hospital, Competitive State Research Financing of the Expert Responsibility area of Oulu University Hospital, The Research Foundation of Orthopaedics and Traumatology in Finland, The Finnish Medical Foundation and The Finnish Medical Society Duodecim.

Oulu, December 2017

Simo Nortunen
Abbreviations

AITFL  anterior inferior tibiofibular ligament
AO    Arbeitsgemeinschaft für Osteosynthesefragen
AP    anteroposterior
CI    confidence interval
CT    computed tomography
ER    external rotation
ICC   intraclass correlation coefficient
LR    likelihood ratio
MCS   medial clear space
MRI   magnetic resonance imaging
OM    Olerud-Molander
OR    odds ratio
ORIF  open reduction and internal fixation
PITFL posterior inferior tibiofibular ligament
RCT   randomized controlled trial
SD    standard deviation
SER   supination external rotation
STTCS superior tibiotalar clear space
TFCS  tibiofibular clear space
List of original publications

This thesis is based on the following publications, which are referred throughout the text by their Roman numerals:


# Abstract

Stability of the ankle mortise in isolated lateral malleolar fractures is defined by the status of the deep deltoid ligament. Syndesmosis is a secondary stabilizer of the ankle mortise.

## 2 Review of the literature

### 2.1 Stability of the ankle mortise in isolated lateral malleolar fractures

#### 2.1.1 Stability of the ankle mortise is defined by the status of the deep deltoid ligament

#### 2.1.2 Syndesmosis is a secondary stabilizer of the ankle mortise

### 2.2 Classification of ankle fractures

#### 2.2.1 Shortcomings of Lauge-Hansen and AO/Danis-Weber classifications

#### 2.2.2 Stability-based ankle fracture classification

### 2.3 Evaluation of stability of the isolated Weber B-/SER-type lateral malleolar ankle fracture without apparent talar shift

#### 2.3.1 Clinical signs

#### 2.3.2 Standard radiographs

#### 2.3.3 Stress radiographs

#### 2.3.4 MRI

#### 2.3.5 CT

#### 2.3.6 Ultrasonography (US)

#### 2.3.7 Arthroscopy

## 3 Aims of the thesis

## 4 Materials and methods

### 4.1 Patients

### 4.2 Clinical examination and stress testing

#### 4.2.1 Clinical signs of deltoid ligament injury

#### 4.2.2 Manual ER stress radiography

#### 4.2.3 Gravity stress radiography

### 4.3 Radiographic measurements

### 4.4 Magnetic resonance imaging

---

### Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>15</td>
</tr>
<tr>
<td>Tiivistelmä</td>
<td>9</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>11</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>13</td>
</tr>
<tr>
<td>List of original publications</td>
<td>15</td>
</tr>
<tr>
<td>Contents</td>
<td>17</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>19</td>
</tr>
<tr>
<td>2 Review of the literature</td>
<td>23</td>
</tr>
<tr>
<td>3 Aims of the thesis</td>
<td>27</td>
</tr>
<tr>
<td>4 Materials and methods</td>
<td>31</td>
</tr>
</tbody>
</table>
4.5 Reliability and consensus building .......................................................... 44
4.6 Ethics ....................................................................................................... 44
4.7 Statistical methods .................................................................................. 44

5 Results
5.1 Non-stress radiography in evaluation of stability (I) ......................... 47
5.2 Diagnostic accuracy of gravity stress radiography and clinical
  signs (II) ...................................................................................................... 48
5.3 Value of MRI in stability assessment of ankle mortise (III) .............. 51

6 Discussion
6.1 General considerations ............................................................................ 53
6.2 Standard radiographs in evaluation of stability ..................................... 54
6.3 Diagnostic accuracy of gravity stress radiography and clinical
  signs .......................................................................................................... 55
6.4 Utility of MRI in stability assessment of ankle mortise ..................... 55
6.5 Clinical implications and future studies .............................................. 57

7 Conclusions
References
Appendix
Original publications
1 Introduction

Stability-based diagnostic algorithms of ankle fractures (Michelson et al. 2007, Pakarinen 2012) have been developed over the last few decades due to a fairly comprehensive consensus opinion that fractures with a stable ankle mortise can be treated conservatively but that unstable fractures are expected to have a better outcome with operative treatment (Yde & Kristensen 1980a, Yde & Kristensen 1980b, Kristensen & Hansen 1985, Phillips et al. 1985, Bauer et al. 1985b, Michelson 1995, Michelson et al. 2007, Pakarinen et al. 2011c). Supination-external rotation (SER)-type (Lauge-Hansen 1950) or AO/Danis-Weber B-type (Danis 1949, Weber 1972, Muller et al. 1979) fractures are by far the most common ankle fractures, including about 75% of all ankle fractures, and isolated fractures of the lateral malleolus alone comprise about 59% of ankle fractures (Pakarinen et al. 2011c). An isolated SER-fracture of the lateral malleolus without incongruity on standard radiographs presents a clinical problem, since the status of the deep deltoid ligament, the main stabilizer of the ankle mortise (Harper 1983, Rasmussen et al. 1983, Rasmussen 1985), remains unknown. Therefore, these injuries can be either stable or unstable. Considerable clinical and research efforts have been made over the last decade to unveil this possible dynamic instability though the clinical relevance of this concept is still to be proven (Michelson et al. 2007).

Clinical signs alone (medial swelling, tenderness, and ecchymosis) regarding the deltoid ligament are considered insufficient to reliably predict the stability of the ankle mortise in patients with isolated lateral malleolus SER-fractures (Egol et al. 2004, DeAngelis et al. 2007, McConnell et al. 2004). The displacement of the fibular fracture is also thought to be of minor importance (Michelson et al. 2007, van den Bekerom & van Dijk 2010, Pakarinen 2011c). However, very little is known about the importance of clinical signs in combination with radiological measurements.

stress test has some problems: it needs to be performed by an experienced examiner, the force applied can vary between examiners (Pakarinen et al. 2011a), it causes pain and discomfort to the patient (Schock et al. 2007), and also the examiner is exposed to radiation. Therefore, several other methods—such as gravity stress radiography (Michelson et al. 2001, Gill et al. 2007, Schock et al. 2007) and magnetic resonance imaging (MRI) (Koval et al. 2007, van den Bekerom et al. 2009)—have been proposed as an alternative for stability evaluation of the ankle mortise. The evidence to support the use of these methods is, however, insufficient.

The aims of this thesis were to assess the roles of clinical findings on the medial side of the ankle, morphological factors from standard radiographs, gravity stress radiography, and MRI in evaluating the stability of the ankle mortise in patients with isolated SER-type fractures of the lateral malleolus with no widening of the medial clear space (MCS) on standard radiographs. The manual ER stress test was used as the reference for stability throughout the studies.

Our hypotheses were that (1) ankles with minimally displaced fractures are stable and do not need further stress testing, (2) the gravity stress test can replace the ER stress test in clinical practice, and (3) the deep deltoid ligament is substantially injured only in unstable ankles.
2 Review of the literature

2.1 Stability of the ankle mortise in isolated lateral malleolar fractures

The medial malleolus, the distal tibial plafond and the lateral malleolus form a mortise in which the talus slides and rolls. Concavity of the tibial plafond and convexity of the saddle-shaped talus are responsible for the bony stability of this construct but there are several ligamentous structures that contribute considerably to stability of the ankle mortise.

2.1.1 Stability of the ankle mortise is defined by the status of the deep deltoid ligament.

The deltoid ligament complex consists of a superficial and a deep layer. The three main components of the superficial layer are the tibionavicula, tibiospring, and tibiocalcaneal ligaments; the deep layer is composed of the anterior and posterior tibiotalar ligaments (Pankovich & Shivaram 1979a). The superficial deltoid ligament is attached to the anterior colliculus, and the deep layer is primarily attached to the posterior colliculus and the intercollicular groove of the medial malleolus (Pankovich & Shivaram 1979a, Pankovich & Shivaram 1979b) (Fig. 1). The deep posterior tibiotalar ligament is the strongest ligament of the whole complex and is responsible for preventing the external rotation of the talus (Rasmussen et al. 1983, Rasmussen 1985).

Fig. 1. Medial ligaments of the tibiotalar joint.
In a simulated SER fracture model, the rotation of the talus was not affected by fibular osteotomy, the transection of the tibiofibular and the superficial deltoid ligaments. The rotational motion of the talus increased only after the deep deltoid fibers were transected, and anatomical plating of the fibula restored the kinematics only partially (Michelson et al. 1996). According to anatomical, biomechanical, and clinical studies, the posterior deep deltoid ligament is the main stabilizer of the ankle mortise (Pankovich & Shivaram 1979a, Pankovich & Shivaram 1979b, Petrone et al. 1983, Phillips et al. 1985, Michelson et al. 1992, Harper 1995, Michelson et al. 2002, Sasse et al. 1999, Earll et al. 1996).

External rotation of the talus is the primary pattern of instability. Talar shift—the widening of the space between the talus and the medial malleolus on standard radiographs is more of a visual illusion created by a two-dimensional representation of a three-dimensional situation (Harper 1983, McCullogh et al. 1980, Yablon 1977, Michelson 1995, Donken et al. 2013). Clinicians are, however, usually forced to depend on two-dimensional radiographs when trying to indirectly evaluate the status of the deep deltoid ligament.

On mortise radiographs, the width of the medial clear space (MCS) is defined as the distance between the lateral border of the medial malleolus and the medial border of the talus at the level of the talar dome (Joy et al. 1974) (Figs. 2 & 3). Absolute MCS widths ranging from 4 to 5 mm (Baird & Jackson 1987, Harper 1988, Michelson 1995, Phillips et al. 1985, McConnell et al. 2004, Egol et al. 2004, Tornetta 2000, Park et al. 2006, Gill et al. 2007, Schock et al. 2007, DeAngelis et al. 2007, Koval et al. 2007 Sanders et al. 2012, Hoshino et al. 2012) or an increase in width of 2 to 3 mm from baseline (Joy et al. 1974, Pankovich 1979, Pankovich & Shivaram 1979a, Pankovich & Shivaram 1979b) have been used as an indirect indication for total deep deltoid ligament tear. According to a biomechanical (Park et al. 2006) and MRI study (Schottel et al. 2015), the measurement of the MCS ≥ 5 mm on mortise radiographs has been found to be the most reliable criteria to predict total disruption of the deep deltoid ligament in cases with isolated lateral malleolar fractures. Similarly, nonoperatively treated patients with MCS > 5 mm initially seem to have worse functional scores (Clements et al. 2008).
Fig. 2. A standard mortise radiograph of the right ankle with an isolated SER-type lateral malleolar fracture. Measurement of (a) the medial clear space, (b) the superior tibiotalar clear space (STTCS), and (c) the tibiofibular clear space (TFCS).

Fig. 3. A mortise radiograph of the left ankle illustrating considerable widening of the medial clear space (MCS) during manual ER stress test. This patient had total tears of the deep deltoid ligaments according to MRI. A standard 30 mm metal sphere and/or standard distance can be used for calibration of the measurements.
2.1.2 **Syndesmosis is a secondary stabilizer of the ankle mortise**

The lateral malleolus of the fibula is firmly held in the fibular notch of the tibia, providing a tight elastic ankle mortise. Four syndesmotic ligaments hold the fibula in the fibular notch of the tibia: the anterior inferior tibiofibular ligament (AITFL); the posterior inferior tibiofibular ligament (PITFL); the interosseous ligament (IOL), which is the thickened continuation of the interosseous membrane (IOM); and the transverse ligament between the malleolar fossa of the fibula and the dorsal rim of the distal tibia (Ogilvie-Harris et al. 1994, Hermans et al. 2010) (Fig. 4). Clinically diagnosed SER III or IV ankle fractures have either a fracture of the posterior malleolus (34%) or a PITFL tear (66%) almost always delaminated of the posterior malleolus (Warner et al. 2015). The standard radiographs do not reveal the height or degree of the possible syndesmosis injury seen in the MRI if there is no incongruity visible (Nielsen et al. 2004, Koval et al. 2007, Cheung et al. 2009, Gardner et al. 2006). However, a differing result has been reported when the Lauge-Hansen classification was used (Hermans et al. 2012).

![Fig. 4. Lateral ligaments of the tibiotalar joint.](image)

The syndesmosis is shown to be a secondary stabilizer of the ankle mortise, as in a cadaver model, whereby total disruption of the syndesmosis complex did not change the loading characteristics of the ankle in the absence of medial injury (Boden et al. 1989). If medial injury is present loading characteristics are affected...
when disruption of the syndesmosis extends more than 4.5 cm proximal to the tibial plafond (Boden et al. 1989). The clinical relevance of a syndesmosis injury is questioned in SER-type ankle fractures, since no difference in functional outcome, pain, or radiologic outcome was found between patient groups with or without trans-syndesmotic fixation or AITFL repair after intraoperatively detected syndesmotic instability following malleolar fixation (Pakarinen et al. 2011b, Pakarinen et al. 2011d, Kortekangas et al. 2014, Kortekangas et al. 2015).

2.2 Classification of ankle fractures

Lauge-Hansen (1950) and AO/Danis-Weber (Danis 1949, Weber 1972, Muller et al. 1979) are the two most widely used classification systems for ankle fractures. Lauge-Hansen originally developed his classification to aid in closed reduction of these fractures. The foot position (supination or pronation) and the direction of the deforming force (ER, abduction or adduction) at the time of injury are the two factors on which Lauge-Hansen based his classification system (Lauge-Hansen, 1950). These four types (SER, pronation-external rotation, pronation-abduction, and supination-adduction) are further divided in stages depending on the magnitude of the deforming force. For each stage, another stabilizing structure is injured. In SER fractures, the lateral structures are tensioned, and the injury starts anteriorly with a rupture of the AITFL (SER I) and progresses around the ankle as the lateral malleolus fractures (SER II), posterior malleolus fractures, or PITFL ruptures (SER III), and finally as the medial malleolus fractures or the deltoid ligament ruptures (SER IV) (Lauge-Hansen 1950).

The AO/Danis-Weber classification system (Danis 1949, Weber 1972) was modified by the AO group (Muller et al. 1979) in the 1970s together with the rise of the new fixation methods that derived from the concept regarding lateral malleolus as the most important stabilizer of the ankle mortise (Ramsey & Hamilton 1976, Yablon et al. 1977). This rather simple classification divides ankle fractures into three types based solely on the location of the lateral malleolus fracture. The lateral malleolus fracture starts distal to the tibial plafond in type A, at the level of the tibial plafond in type B, and proximal of the tibial plafond in type C.
2.2.1 Shortcomings of Lauge-Hansen and AO/Danis-Weber classifications


The inability to produce similar fracture patterns in later cadaver studies (Michelson et al. 1997, Haraguchi & Armiger, 2009) and the inconsistency of soft tissue injury related to the observed fractures (Cedell 1967, Michelson et al. 1997, Kwon et al. 2015) has also questioned Lauge-Hansen’s original study. Then again, the AO/Danis-Weber system ignores the medial injury altogether.

2.2.2 Stability-based ankle fracture classification

For decades, the cornerstone for the treatment of ankle fractures has been the reduction and/or maintenance of the ankle mortise until fracture union. Sufficient stability of the ankle mortise is required to achieve this goal. Biomechanical studies have shown that even a millimeter of talar shift can substantially reduce the tibio-talar contact area (Ramsey et al. 1976, Clarke et al. 1991, Pereira et al. 1996) and a good reduction at the fracture union seems to predict a good clinical outcome (Phillips et al. 1985, Tunturi et al. 1983). Michelson et al. (2007) suggested after a structural analysis of the ankle fracture literature that a stability-based ankle fracture classification could be prognostic. For ankle fractures that are nowadays considered undeniably as unstable (for example bimalleolar fractures or ankle fractures with clear talar shift), the radiographic outcomes were better after surgery when the decision for surgery was made on the basis of
stability. Nonoperative treatment results have also been better with stability-based treatment (Michelson et al. 2007, Pakarinen et al. 2011c).

According to long-term follow-up studies, unimalleolar SER fractures without talar shift on standard radiographs can be treated nonoperatively with good or excellent results and with minimal risk of posttraumatic osteoarthrosis (Yde & Kristensen 1980a, Kristensen & Hansen 1985, Phillips et al. 1985, Bauer et al. 1985b, Donken et al. 2012). However, it is important to consider that these studies predate the era of stress testing. Although, Michelson et al. (2007) emphasized this, their proposition for the decision tree (Fig. 5) included stress testing for isolated lateral malleolar ankle fractures without incongruity on standard radiographs. More recently, Pakarinen et al. (2011c) found this stability-based fracture classification to be a simple and useful tool for selecting a treatment method for ankle fractures. Utilization of this classification led to the operative treatment of about half of the patients in their study (Pakarinen et al. 2011a).
2.3 Evaluation of stability of the isolated Weber B-/SER-type lateral malleolar ankle fracture without apparent talar shift

2.3.1 Clinical signs

The subjective findings of medial tenderness, swelling, and ecchymosis have been reported to indirectly imply a deltoid ligament injury and have been used as an indication for surgical treatment in the presence of an isolated fibular fracture with intact ankle mortise in non-stress radiographs (Brink et al. 1996, Malka & Taillard 1969, Port et al. 1996). However, both DeAngelis et al. (2007) and McConnell et al. (2004) have stated that medial signs alone are not accurate predictors of instability.

Egol et al. (2004) examined 101 patients with isolated SER-type fractures of the lateral malleolus with no MCS widening and found that 65% were unstable...
according to the ER stress test. They concluded that, if evaluated within 24 hours of injury, the accuracy of clinical signs alone was insufficient to predict stability of the ankle mortise. They treated all of the patients with a positive stress radiograph and no clinical signs nonoperatively, with good or excellent clinical results. However, their MCS cut-off point for instability was ≥ 4 mm, the mean MCS for unstable 66 ankles was only 4.6 mm, and for those with medial sign(s) the mean MCS was 5.3 mm. It can be argued that their conservatively treated ER stress positive but medial sign negative group of patients actually did not have an unstable ankle mortise after all, since in a cadaver study Park et al. (2006) showed that the MCS value of ≥ 5 mm is a more accurate cut-off point to reveal total tear of the deep deltoid ligament. Previous publications on the accuracy of clinical signs in patients with unimalleolar SER fractures and intact ankle mortise have been summarized in Table 1.

<table>
<thead>
<tr>
<th>Author</th>
<th>Number of patients</th>
<th>ER+ MCS cut-off value</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egol et al. 2004</td>
<td>101; ER+ 66</td>
<td>MCS ≥ 4mm</td>
<td>Medial tenderness, ecchymosis and swelling are not sensitive in predicting stability</td>
</tr>
<tr>
<td>McConnell et al. 2004</td>
<td>97; ER+ 36</td>
<td>MCS &gt; 4mm</td>
<td>Soft-tissue indicators are not accurate predictors of instability</td>
</tr>
<tr>
<td>DeAngelis et al. 2007</td>
<td>55; ER+ 23</td>
<td>MCS &gt; 4mm</td>
<td>No correlation of medial tenderness with positive stress radiographs</td>
</tr>
</tbody>
</table>

### 2.3.2 Standard radiographs

Routinely, only mortise and lateral radiographs are acquired if a fracture of the ankle is suspected after the clinical evaluation; more than two projections improve diagnostic accuracy minimally and are not cost-effective (Winkler et al. 1990, Brage et al. 1998, Cockshott et al. 1983, Vangsness et al. 1994, Brandser et al. 2000). Several recommendations of acceptable amounts of diastasis in the fibular fracture or dislocation of the talus have been made in the past, but there is little evidence to support the use of these criteria (Ali et al. 1987, Bauer et al. 1985, Dahners 1990, Lantz et al. 1991, Lehto & Tunturi 1990, Parlasca et al. 1979, Rasmussen et al. 1993, Sneppen 1969, Vandsness et al. 1994).

The biomechanical studies of Ramsey & Hamilton (1976) and Yablon et al. (1977) in the 1970s led to the concept that the talus follows the lateral malleolus
and that the amount of displacement of the lateral malleolus is the most important factor when choosing a treatment method. This concept has since been abandoned due to the findings in biomechanical and CT studies that in an isolated lateral malleolar fracture the kinematics of the joint do not change without medial side injury (Brown et al. 1994, Burns et al. 1993, Clarke et al. 1991), the displacement of the fibula does not determine talar displacement in an axially loaded ankle (Michelson et al. 1994), and the distal fibular fragment is actually anatomically positioned leaving the ankle mortise intact while the proximal fibula is internally rotated and medialized (Michelson et al. 1992).

The recent literature concentrates on the comparison between different dynamic tests or on the assessment of “new” imaging modalities such as MRI and CT for the stability evaluation of ankle fractures. To our knowledge, no studies have been performed to assess the diagnostic value of standard non-stress radiographs in stability evaluation for isolated lateral malleolar SER fractures with no widening of the MCS, using stress radiographs as a reference for stability. Choi et al. (2014) assessed preoperative standard radiographs and CTs to find factors predicting a syndesmotic instability after malleolar fixation in SER fracture patients. They divided patients into two groups based on the result of an intraoperative lateral stress test (the Cotton test) and found that a fibular fracture starting > 7 mm proximal to the anterior margin of the distal tibial articular surface and a MCS > 4.5 mm on standard radiographs predicted unstable syndesmotic injuries (Choi et al. 2014). However, it must be assumed that all the patients in their retrospective study had an unstable ankle mortise, since all were operated on even though they did not report the criteria leading to operative treatment.

2.3.3 Stress radiographs


The reported amount of applied force necessary during manual ER stress is not uniform or is only an estimate; and more often it is not reported (McConnell et al. 2004, Park et al. 2006, Tornetta 2000, Hoshino et al. 2012, Stufkens et al. 2012). A gravity stress test was proposed by Michelson et al. (2001) to reproducibly reveal deltoid ligament injury. However, it is not possible to control the flexion of the ankle compared to ER stress testing. It has since been reported that gravity stress radiography is equivalent to manual stress radiography for determining deltoid ligament injury—with less pain to the patient and no radiation exposure to the physician, and it can be performed by assistant radiographers (Gill et al. 2007, Schock et al. 2007, LeBa et al. 2015). However, an MCS cut-off value of $\geq 5$ mm to indicate an unstable ankle mortise was not used in these studies with a rather small number of patients, nor were the sensitivity or specificity of the gravity stress radiography reported (Gill et al. 2007, Schock et al. 2007, LeBa et al. 2015). Therefore, the clinical evidence to support the use of the gravity stress radiography is limited.

Weber et al. (2010) proposed the use of weightbearing radiographs as an easy, pain-free, safe, and reliable method to exclude the need for operative treatment with excellent clinical outcome in the majority of the patients seen at latest follow-up. However, this last follow-up was an interview by phone, and no radiographs were taken. Only 9% of patients had an unstable ankle mortise based on these weightbearing radiographs acquired 3–10 days after injury (Weber et al. 2010). Although a delayed visit is required to enable the patients to stand after most of the pain has ceased, the use of weightbearing radiographs seems to be well accepted by the patients (Weber et al. 2010). Hoshino et al. (2012) enrolled 36 isolated SER-type lateral malleolar fracture patients with both positive clinical signs and positive ER stress tests. They found only one (3%) patient to have an unstable ankle mortise on weightbearing radiographs and thus, to require operative treatment. They reported good to excellent functional scores after a one-year follow-up although with a follow-up rate of 71%.
If weightbearing radiographs are used to guide the appropriate treatment method, the need for surgery is substantially lower compared to ER and gravity stress tests (Weber et al. 2010, Hoshino et al. 2012, Dawe et al. 2015, Seidel et al. 2017). Based on these clinical studies, the use of weightbearing radiographs seems promising, but there is only level 3 evidence to support their use for decision-making and no basic science or, for example, MRI studies to correlate the MCS on weightbearing radiographs to the integrity of the deltoid ligament. In fact, Stewart et al. (2012) found no association between the MCS and the status of the deep deltoid ligament in an axially loaded cadaveric model.

2.3.4 MRI

Few investigators have used MRI to assess the integrity of the deep deltoid ligament to distinguish between stable and unstable injuries and to guide clinical decision-making (Koval et al. 2007, Cheung et al. 2009, Gardner et al. 2006, Berkes et al. 2012, Hermans et al. 2012). Based on these reports, the cut-off value of 5 mm, which predicts deep deltoid ligament rupture in the manual ER stress test, has been questioned. Koval et al. (2007) reported good clinical results and recommended nonoperative treatment in ER stress-positive injuries if the posterior deep deltoid ligament is only partially torn, as observed by MRI. However, the diagnostic value of MRI has not been sufficiently assessed; previous studies of SER fractures (Koval et al. 2007, Cheung et al. 2009) have only included ER stress-positive injuries, and little is known of the MRI findings of the deep deltoid ligament in both stable and unstable injuries. In these previous studies, the authors have not been able to find the correlation between the deep deltoid status seen on the MRI and the result of the ER stress radiographs; nor has the reliability of MRI in the assessment of deep deltoid injury been reported (Table 2). Studies comparing gravity stress or weightbearing radiographs and the MRI findings of the deltoid ligament have not been reported.
Table 2. Previously reported MRI findings of the deep deltoid ligament in patients with a SER fracture. A stress test (ER) was used as a reference of stability in both studies.

<table>
<thead>
<tr>
<th>Author</th>
<th>Number of patients</th>
<th>MRI results</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Koval et al. 2007</td>
<td>21 (ER+)</td>
<td>19 partial, 2 total tears</td>
<td>Most of the ER+ patients do not need operative treatment</td>
</tr>
<tr>
<td>Cheung et al. 2009</td>
<td>19 (ER+)</td>
<td>15 partial tears, 2 total tears, 1 normal</td>
<td>Most deep deltoid tears are partial, AITFL was ruptured in every case</td>
</tr>
</tbody>
</table>

2.3.5 CT

There are no clinical studies assessing the diagnostic value of CT in the stability evaluation of isolated lateral malleolus SER fractures without incongruity using stress testing or the MRI findings of the deep deltoid ligament as reference for stability. Donken et al. (2013) compared talar rotation on CT with a widening of the MCS in a cadaver model of just six specimens. They concluded that only high-grade injuries (with total transection of the deep deltoid ligament) combined with external rotation resulted in an MCS widening of 2 mm or more. However, talar rotation was severe even in low-grade injuries with normal MCS in reformatted mortise radiographs (Donken et al. 2013). The clinical relevance of this talar rotation is unknown, considering that in a weightbearing CT study of uninjured ankles the talus has been shown to rotate a mean of 10 degrees with large intersubject variation but with no widening of the MCS (Lepojärvi et al. 2016a).

2.3.6 Ultrasonography (US)

Ultrasonography’s wide availability, low cost, and dynamic nature make it a potential method for evaluating the integrity of ligaments. However, the high degree of operator dependency and challenging anatomical location of the deep deltoid ligament weaken its feasibility in the stability evaluation of the ankle mortise (Stufkens et al. 2012). Only one study comparing the value of US with MRI in the diagnosis of deep deltoid ligament injuries of 28 patients with Weber B or Weber C injuries has been reported. The researchers found US to be reliable for detecting total tears but to be unable to identify partial tears (Lechner et al. 2015), which comprise about 90% of the cases according to MRI studies (Koval
et al. 2007, Cheung et al. 2009). Overall, the literature is insufficient to evaluate the role of US in the stability assessment of the ankle mortise.

2.3.7 Arthroscopy

Arthroscopy is invasive and too costly to be used as a primary method for stability evaluation; and even its value in the stability evaluation of ankle fractures intraoperatively is unclear. The visualization of injured deep deltoid ligaments is not always possible directly after trauma, and the superficial components cannot be seen at all; partial tears of the deep deltoid ligament are especially difficult to diagnose arthroscopically (Hintermann et al. 2000, Schuberth et al. 2004). Schuberth et al. (2004) compared MCS with the deep deltoid ligament injury seen in arthroscopy and found that arthroscopy was unreliable in diagnosing unstable ankle mortise.
3 Aims of the thesis

The aims of this thesis were as follows:

1. To identify demographic factors and radiographic factors from standard radiographs that contribute to the stability of the ankle mortise in patients with isolated SER-type fractures of the lateral malleolus.
2. To assess the diagnostic accuracy of the gravity stress test and clinical findings to evaluate the stability of the ankle mortise in patients with SER-type fractures of the lateral malleolus with no widening of the medial clear space.
3. To assess the utility of MRI findings for the deep aspect of the deltoid ligament in evaluating the stability of the ankle mortise in patients who have an SER-type lateral malleolar fracture with no widening of the medial clear space.
4 Materials and methods

4.1 Patients

All skeletally mature patients (≥ 16 y) with a unilateral SER-type (Lauge-Hansen, 1950) ankle fracture with no indication of medial widening or incongruity on standard ankle radiographs who were treated within a week after injury were considered eligible for these prospective studies.

Study I (Demographic and radiographic factors)

A total of 308 prospectively collected, consecutive patients treated within a week after injury at two main trauma centers—the Oulu university hospital (between March 2012 and May 2015) and the Tampere university hospital (between February 2013 and August 2014)—were screened for Study I. Patients were excluded when they had a pathologic fracture (n = 0), concomitant fractures that contra-indicated ER stress testing (n = 1), or a previous notable ankle injury (n = 21).

The final study group comprised 286 consecutive patients that met the inclusion criteria (mean age, 45 y; range, 16–85 y), including 144 women (mean age, 50 y; range, 17–85 y) and 142 men (mean age, 40 y; range, 16–84 y). The average time delay from injury to clinical evaluation and stress test was 2.4 d (range, 0–6 d).

Study II (Accuracy of gravity stress radiography and clinical signs)

A total of 96 patients treated at Oulu University Hospital from March 2012 to April 2013 were screened for Study II. Twelve patients were excluded because of the exclusion criteria: previous ankle injury (n = 8), inability to complete the study protocol (n = 2), inability to walk without aid prior to fracture (n = 1), and peripheral neuropathy (n = 1). Two patients refused to participate, and three patients had unsatisfactory gravity stress radiographs.

The final study group comprised 79 consecutive patients (mean age 44 y, range 16–82 y): 37 women (mean age 50 y, range 16–82 y) and 42 men (mean age 39 y, range 16–79 y). The average delay from injury to clinical evaluation and
stress tests was 2.3 d (range 0–6 d). Twenty-nine patients (41%) were evaluated and tested within 24 h after injury.

**Study III (Utility of MRI for stability assessment)**

A total of 87 patients were screened for Study III. Twenty-six patients were excluded, 11 patients because of the exclusion criteria (previous ankle injury, n = 8; inability to complete the study protocol, n = 2; and peripheral neuropathy, n = 1). Five patients refused to participate. Ten patients (five with MCS of < 5 mm and five with MCS of ≥ 5 mm) were excluded, since no MRI was performed (patient withdrew consent n = 8; MRI contraindicated, n = 2).

The final study group consisted of 61 patients (27 females, 34 males; mean age: 45 ± 18 yrs, range: 16–82 yrs). The average delay to stress tests was 2.3 ± 1.6 days (range 0–6 days).

4.2 Clinical examination and stress testing

4.2.1 Clinical signs of deltoid ligament injury (II)

Two surgeons examined the patients together prior to the stress tests and assessed swelling, tenderness, and ecchymosis around the distal tip of the medial malleolus and deltoid ligament. Clinical findings were recorded as positive or negative.

4.2.2 Manual ER stress radiography (I, II, III)

In the manual ER stress test (Pankovich 1979, Pankovich and Shivaram 1979b), the tibia was stabilized with one hand, internally rotated 10-15° for a true mortise view, then the ankle was turned to neutral flexion, and an external rotation force was applied to the forefoot with the other hand (Fig. 6). The result of the manual ER stress test was considered as the reference for stability throughout this thesis.
Study I

Senior orthopedic trauma surgeons at the Oulu university hospital and the surgeon on call at the Tampere university hospital performed the ER stress tests. The manual ER stress test was the standard protocol at both trauma centers; therefore, all examiners were trained and familiar with performing the test.

Studies II and III

Two senior orthopedic trauma surgeons, or a senior orthopedic trauma surgeon and a senior orthopedic trauma resident who had completed our trauma training, performed manual ER stress tests for all patients. There were 10 surgeons involved with the ER stress tests. The principles of the manual ER stress test were summarized and rehearsed before being performed. One of the surgeons left the
room, and the other surgeon performed the manual ER stress test; a standard digital mortise radiograph was then taken. The roles were then reversed, with the second surgeon carrying out the same test.

4.2.3 Gravity stress radiography (II)

A diagnostic radiographer performed imaging for the gravity stress test as described by Michelson *et al.* (2001) (Fig. 7). Patients lay on the injured side with a mount placed under the leg, which was internally rotated 10° to 15°, and avoiding plantar flexion of the foot, a mortise view in the horizontal radiograph was acquired.

![Fig. 7. Performance of the gravity stress radiography. a) X-ray beam source. b) Digital x-ray detector panel. c) Leg support. d) 30 mm radiographic marker.](image)
4.3 Radiographic measurements

**MCS (I, II, III)**

Measurements of all stress radiographs were normalized for magnification using a constant 115 cm source-to-detector distance and a standard 30 mm radiographic marker. Standard non-stress radiographs were calibrated by a constant distance. All radiographs were analyzed with a diagnostic workstation and a high-resolution monitor. Measurements were approximated to the nearest millimeter.

The MCS was measured as the distance (mm) between the lateral border of the medial malleolus and the medial border of the talus at the level of the talar dome (Joy et al. 1974). The tests were considered positive and the ankle unstable if the MCS was \( \geq 5 \) mm and at least 1 mm greater than the superior tibiotalar clear space (Michelson 1995).

**Standard non-stress radiographs (I)**

Mortise radiographs were measured to determine the tibiofibular clear space (TFCS) at the level of the epiphyseal scar on the distal tibia, the maximum width of the fracture line (lateral diastasis), and the distance between the distal tip of the fracture and the talar dome (distal fracture height) (Fig. 8).

Measurements on lateral radiographs included the maximum width of the fracture line (posterior diastasis); the anterior and posterior fracture heights, measured perpendicular to the level of the talar dome; and the obtuse angle between the fracture line and the axis of fibula (fracture line angle) (Fig. 9). The fracture fragments were counted, except for the typical AITFL-avulsion fragment.
Fig. 8. Measurements on mortise radiographs: distal fracture height (a), TFCS (b), and lateral diastasis (c).

Fig. 9. Measurements on lateral radiographs: posterior diastasis (a), posterior fracture height (b), fracture line angle (c), and anterior fracture height (d).
4.4 Magnetic resonance imaging (III)

After splinting, all patients underwent MRI examination on a 3.0-tesla (T) MAGNETOM Skyra (Siemens Healthcare Sector, Erlangen, Germany) MR unit using coil elements of body matrix and spine matrix coils. The patients were placed supine, and the splinted extremity was neutrally positioned. The MRI sequences consisted of trans-axial T1-weighted turbospin-echo, sagittal TIRM (turbo inversion recovery magnitude), coronal intermediate-weighted, trans-axial intermediate-weighted, fat-saturated, and isotropic intermediate-weighted 3D space sequences. The injury of the anterior and posterior deep deltoid ligament was graded on a four-grade scale. The evaluation was based on the visibility of the normal fatty striation between ligament fibers, fluid signal between otherwise intact fibers, continuity of the ligament fibers, or the absence of the ligament fibers.

The traditional three-grade scale (normal, partial, total) appeared to be too insensitive, since, in the 3T MRI images, partial disruption of the fibers was clearly visible and could be separated from contusion or distension oedema between otherwise intact ligament fibers (Chhabra et al. 2010). Based on these images, ligaments were judged as intact with normal fatty striation (Fig. 10), oedema between continuous fibers (oedema, Fig. 11), fluid signal transecting some of the ligament fibers (partial tear, Fig. 12), or complete disruption of the fibers (total tear or avulsion, Fig. 13). No sign of previous trauma (e.g., thickening, thinning, low signal intensity scarring) was found in the deltoid ligament complex or other ankle ligaments. A tear of the deep deltoid was considered total when both parts of the ligament were completely torn.
Fig. 10. Coronal intermediate-weighted 3.0-T MRI of a normal deep deltoid ligament (white arrow). This patient did not participate in our study, since we did not find a single uninjured deep deltoid ligament by MRI.

Fig. 11. Coronal intermediate-weighted 3.0-T MRI of a oedematic deep deltoid ligament. Oedema between continuous fibers (white arrow).
Fig. 12. Coronal intermediate-weighted 3.0-T MRI of a partial tear of the deep deltoid ligament, fluid signal transecting some of the ligament fibers (white arrow).

Fig. 13. Coronal intermediate-weighted 3.0-T MRI of a total tear of the deep deltoid ligament (white arrow).
4.5 Reliability and consensus building

Two senior orthopedic trauma surgeons analyzed all standard mortise and lateral non-stress radiographs in separate sessions and were blinded to each other’s measurements (I). Two orthopedic surgeons, blinded to the results of the ER stress testing and to the clinical findings, independently evaluated the gravity stress radiographs to calculate the reliability (II). Disagreements were resolved by a consensus decision. Clinical findings on the medial aspect of the ankle were recorded in consensus (II). Two senior musculoskeletal radiologists independently analyzed the anterior and posterior deep deltoid ligament on the MRI and were blinded to the result of the ER stress tests (III). In cases of disagreement, a consensus decision was regarded as the final grade of injury (III). The results of the initial ER stress test performers and the results of the re-analysis, which was done twice in separate sessions by the author, were used to calculate the inter-observer reliability of the ER stress radiography (II, III).

4.6 Ethics

The ethical committee of Oulu University Hospital approved all study protocols (I-III). All patients received oral and written information about the trial and provided informed consent to participate.

4.7 Statistical methods

Simple between-group comparisons were analyzed either by t-test, one-way analysis of variance (one-way ANOVA) or \( \chi^2 \)-test. Two-tailed \( p \)-values < 0.05 were considered significant. The inter-observer reliabilities were calculated as the percentage agreement between the assessors and the kappa coefficient (Landis & Koch 1977, Fleiss 1981) or with intraclass correlation coefficients (ICCs) (Zidan et al. 2015).

In Study I, the receiver operating characteristic (ROC) analyses were performed for continuous variables to test for diagnostic accuracy and to determine optimal thresholds. Variables with areas under the ROC-curve (AUC) > 0.75 were considered adequate for further investigation. A sensitivity > 90% was used as the criterion for an optimal threshold. Variables with \( p \)-values < 0.2 were manually entered, one by one, into the multiple logistic regression analysis to model the stability of the ankle mortise. A variable remained in the model when
its $p$-value was < 0.05 or when it had a significant impact on the -2 Log likelihood value. The probability of ankle mortise stability was calculated in different situations with a logit function (logit = ln[odds]).

In Study II, sensitivities, specificities, and positive and negative likelihood ratios (LRs) (Jaeschke et al. 1994, Sackett et al. 2000) with 95% confidence intervals (CIs) of clinical findings and the gravity stress test were calculated using the manual ER stress test as a reference standard. LRs were used to calculate post-test probabilities when the incidence of unstable injuries in our study group was used for pre-test probability. (Please see the appendix for more information concerning Bayesian statistics and the interpretation of LRs.)
5 Results

5.1 Non-stress radiography in evaluation of stability (I)

The group of 217 (75.9%) stable ankles had a mean MCS of 3.5 mm (SD 0.6, range 2.0–4.0). The group of 69 unstable ankles had a mean MCS of 5.7 mm (SD 1.1, range 5.0–11.0).

These groups showed significant differences in patient sex (p = 0.001), TCFS (p < 0.001), posterior diastasis (p < 0.001), posterior fracture height (p = 0.031), fracture line angle (p = 0.017), and number of fragments (p = 0.001) (Table 3). The only continuous radiological variable with an adequate AUC value (0.78, 95% CI: 0.71–0.84) in the ROC analysis was the posterior diastasis, measured in lateral radiographs. The selected threshold of < 2 mm for posterior diastasis corresponded to a sensitivity of 0.94 and a specificity of 0.39.

Table 3. Comparison between stable and unstable groups.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Stable</th>
<th>Unstable</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>217</td>
<td>69</td>
<td>-</td>
</tr>
<tr>
<td>Sex (F/M)</td>
<td>121/96</td>
<td>23/46</td>
<td>0.001</td>
</tr>
<tr>
<td>Age (years)</td>
<td>46 ± 18</td>
<td>42 ± 17</td>
<td>0.44</td>
</tr>
<tr>
<td>Mortise radiographs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TFCS (mm)*</td>
<td>3.2 ± 1.0</td>
<td>3.9 ± 1.3</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Lateral diastasis (mm)*</td>
<td>1.5 ± 1.1</td>
<td>1.6 ± 1.7</td>
<td>0.60</td>
</tr>
<tr>
<td>Distal fracture height (mm)*</td>
<td>−1.9 ± 4.4</td>
<td>−0.8 ± 5.8</td>
<td>0.17</td>
</tr>
<tr>
<td>Lateral radiographs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posterior diastasis (mm)*</td>
<td>1.8 ± 1.1</td>
<td>3.3 ± 1.5</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Anterior fracture height (mm)*</td>
<td>−0.7 ± 5.8</td>
<td>0.0 ± 7.3</td>
<td>0.47</td>
</tr>
<tr>
<td>Posterior fracture height (mm)*</td>
<td>33.0 ± 12.3</td>
<td>37.0 ± 13.7</td>
<td>0.031</td>
</tr>
<tr>
<td>Fracture line angle (°)*</td>
<td>146.8 ± 7.8</td>
<td>149.3 ± 7.5</td>
<td>0.017</td>
</tr>
<tr>
<td>Number of fracture fragments (2 / &gt; 2)</td>
<td>211/6</td>
<td>60/9</td>
<td>0.001</td>
</tr>
</tbody>
</table>

* Values are presented as mean ± standard deviation

The multiple logistic regression analysis showed that a posterior diastasis < 2 mm (OR 10.8, 95% CI: 3.7–31.5), only two fracture fragments (OR 7.3, 95% CI: 2.1–26.3) and the female sex (OR 2.5, 95% CI: 1.4–4.6) were independent factors for predicting a stable ankle mortise (Table 4).
Table 4. Multiple logistic regression analysis results for factors related to ankle mortise stability.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Odds ratio</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posterior diastasis on lateral radiograph &lt; 2 mm</td>
<td>10.8</td>
<td>3.7 to 31.5</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Only two fracture fragments</td>
<td>7.3</td>
<td>2.1 to 26.3</td>
<td>0.002</td>
</tr>
<tr>
<td>Sex (female)</td>
<td>2.5</td>
<td>1.4 to 4.6</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Among the 85 patients with fractures that had a posterior diastasis < 2 mm and only two fracture fragments, the probability of a stable ankle mortise was 0.98 for 48 women (16.8% of all patients) and 0.94 for 37 men (12.9% of all patients). In contrast, among six (2.1%) male patients with fractures that had a posterior diastasis ≥ 2 mm and three fracture fragments, the probability of an unstable ankle mortise was 0.83.

ICC for posterior diastasis was 0.975 (95% CI: 0.968–0.980), indicating excellent inter-observer reliability. Two or three fracture fragment measurements were in perfect agreement between the two observers.

5.2 Diagnostic accuracy of gravity stress radiography and clinical signs (II)

The gravity stress test had a sensitivity of 0.87 and a specificity of 0.85 for diagnosing the instability of the ankle mortise. The positive LR of the gravity stress test was 5.81 (95% CI: 2.75–12.27) and the negative LR was 0.15 (95% CI: 0.07–0.35). The ER stress test results were positive in 39 cases (49%), and the gravity stress test results were positive in 40 cases (51%) (Table 5). For medial tenderness and its combinations with swelling and ecchymosis, positive LR values ranged from 2.74 to 3.25 and negative LR values from 0.50 to 0.62. For swelling and ecchymosis around the medial malleolus and their combination, positive LR values varied from 1.41 to 1.65 and negative LR values from 0.38 to 0.47 (Table 6).

The k coefficient for the gravity and ER stress tests was 0.72, indicating substantial reliability. Inter-observer reliability was 91% for the ER stress test and 95% for the gravity stress test, and k coefficients were 0.82 (substantial reliability) and 0.90 (almost perfect reliability), respectively.
Table 5. Cross-tabulation between ER and gravity stress tests.

<table>
<thead>
<tr>
<th>ER stress test</th>
<th>Gravity stress test</th>
<th>Negative (&lt; 5 mm)</th>
<th>Positive (≥ 5 mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative (&lt; 5 mm)</td>
<td>34</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Positive (≥ 5 mm)</td>
<td>5</td>
<td>34</td>
<td></td>
</tr>
</tbody>
</table>

The incidence of ankle mortise instability in our study population was 0.49. The probability of an ER stress test positive fracture was 0.85 after a positive gravity stress test and 0.13 after a negative test. If medial tenderness was present, this probability increased to 0.94–0.95 after a positive gravity stress test and decreased to 0.29–0.32 after a negative gravity stress test. Without clinical findings, the probability of an ER stress test positive fracture was 0.77 after a positive gravity stress test and 0.08 after a negative gravity stress test (Table 6).

No significant associations were found between the ER stress test results and the gravity stress test results (p = 0.11), medial tenderness (p = 0.83), or delay between injury and evaluation (p = 0.21). Delay between injury and evaluation was associated with swelling (p = 0.01) and ecchymosis (p = 0.004). Patient age was not associated with medial tenderness (p = 0.58) or swelling (p = 0.32) but was associated with ecchymosis (p = 0.005).
Table 6. Sensitivity, specificity, positive (LR+) and negative (LR-) likelihood ratios (with 95% CIs), and post-test probability values for clinical signs and gravity stress test for an ER stress test positive injury (pre-test probability 0.49).

<table>
<thead>
<tr>
<th>Test (n)*</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>LR+ (95% CIs)</th>
<th>LR- (95% CIs)</th>
<th>Post-test probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gravity</td>
<td>0.87</td>
<td>0.85</td>
<td>5.8 (2.75 to 12.27)</td>
<td>0.15 (0.07 to 0.35)</td>
<td>Gravity +</td>
</tr>
<tr>
<td>Clinical tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Clinical+</td>
</tr>
<tr>
<td>Swelling (n = 57)</td>
<td>0.85</td>
<td>0.4</td>
<td>1.41 (1.06 to 1.9)</td>
<td>0.38 (0.17 to 0.88)</td>
<td>0.89</td>
</tr>
<tr>
<td>Medial Tenderness (n = 33)</td>
<td>0.62</td>
<td>0.78</td>
<td>2.74 (1.46 to 5.12)</td>
<td>0.5 (0.32 to 0.76)</td>
<td>0.94</td>
</tr>
<tr>
<td>Ecchymosis (n = 49)</td>
<td>0.77</td>
<td>0.53</td>
<td>3.62 (1.12 to 2.48)</td>
<td>0.75 (0.23 to 0.84)</td>
<td>0.9</td>
</tr>
<tr>
<td>Swelling and Tenderness (n = 29)</td>
<td>0.56</td>
<td>0.83</td>
<td>2.22 (1.56 to 6.67)</td>
<td>0.53 (0.36 to 0.78)</td>
<td>0.93</td>
</tr>
<tr>
<td>Swelling and Ecchymosis (n = 47)</td>
<td>0.74</td>
<td>0.55</td>
<td>1.65 (1.12 to 2.44)</td>
<td>0.47 (0.26 to 0.85)</td>
<td>0.9</td>
</tr>
<tr>
<td>Tenderness and Ecchymosis (n = 26)</td>
<td>0.49</td>
<td>0.83</td>
<td>2.78 (1.32 to 5.87)</td>
<td>0.62 (0.44 to 0.87)</td>
<td>0.94</td>
</tr>
<tr>
<td>All Clinical Signs (n = 25)</td>
<td>0.49</td>
<td>0.85</td>
<td>3.25 (1.45 to 7.26)</td>
<td>0.6 (0.43 to 0.84)</td>
<td>0.95</td>
</tr>
</tbody>
</table>

* number of positive findings
5.3 Value of MRI in stability assessment of ankle mortise (III)

The ER stress test was negative (stable) in 30 (49%) cases and positive (unstable) in 31 (51%) cases, as determined by at least one surgeon. In re-analysis, we found 28 (46%) cases with MCS < 5 mm and 33 (54%) cases with MCS ≥ 5 mm. For these two groups, the mean MCS ± SD and range were 3.9 ± 0.3 mm (range 3.0 to 4.0) and 6.0 ± 1.3 mm (range 5.0 to 11.0), respectively. Injury to the anterior part of the deep deltoid ligament was assessed as oedematous in 23 cases, partially torn in 33 cases, and totally torn in five cases. Injury to the posterior part was oedematous in 24, partially torn in 35, and totally torn in two cases. In both of these cases with a total posterior deep deltoid tear, there was also a complete tear of the anterior deep deltoid (Table 7).

Table 7. Cross-tabulation of the MRI findings of the anterior and posterior deep deltoid ligament.

<table>
<thead>
<tr>
<th>Anterior deep deltoid ligament</th>
<th>Posterior deep deltoid ligament</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Oedema</td>
</tr>
<tr>
<td>Oedema</td>
<td>14</td>
</tr>
<tr>
<td>Partial tear</td>
<td>10</td>
</tr>
<tr>
<td>Total tear</td>
<td>0</td>
</tr>
</tbody>
</table>

None of patients with MCS < 5 mm had a normal anterior or posterior deep deltoid ligament by MRI, and 19 cases had a partial tear of either ligament (Table 8). In patients with MCS ≥ 5 mm, 26 cases had a partial tear, and two cases had total tears (Table 8). We did not find a single uninjured deep deltoid ligament by MRI.

Patients with a total, deep deltoid ligament tear (n = 2) in the MRI had MCS values of 7.0 and 11.0 mm, whereas for patients with a partial tear of the deep deltoid ligament (n = 45), the mean MCS was 5.0 ± 1.1 mm (range: 4.0 to 8.0). In cases of oedematous deep deltoid ligament (n = 14), the mean MCS was 4.6 ± 1.3 mm (range: 3.0 to 8.0). The MCS increased according to the severity of the observed deep deltoid ligament injury in the MRI (p < 0.001) (Fig. 14).

The inter-observer agreement of the ER stress tests was 94%, and the kappa value was 0.87, indicating excellent inter-observer reliability. The inter-observer agreement of the MRI findings between the two musculoskeletal radiologists was 72% for the posterior deep deltoid and 56% for the anterior deep deltoid, and the
kappa values were 0.46 and 0.22, indicating moderate and fair inter-observer reliability, respectively.

Table 8. The MRI findings of the deep deltoid ligament according to medial clear space (MCS) group (< 5 mm and ≥ 5 mm).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Frequency (no.)</th>
<th>Percent (%)</th>
<th>Within MCS group</th>
<th>Of all patients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MCS &lt; 5 mm</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oedema</td>
<td>9</td>
<td>32.1</td>
<td>14.8</td>
<td></td>
</tr>
<tr>
<td>Partial tear</td>
<td>19</td>
<td>67.9</td>
<td>31.1</td>
<td></td>
</tr>
<tr>
<td>Total tear</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td><strong>MCS ≥ 5 mm</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oedema</td>
<td>5</td>
<td>15.2</td>
<td>8.2</td>
<td></td>
</tr>
<tr>
<td>Partial tear</td>
<td>26</td>
<td>78.8</td>
<td>42.6</td>
<td></td>
</tr>
<tr>
<td>Total tear</td>
<td>2</td>
<td>6.1</td>
<td>3.3</td>
<td></td>
</tr>
</tbody>
</table>

Partial tear = Either anterior or posterior part of the deep deltoid ligament was partially ruptured.

Total tear = Both anterior and posterior deep deltoid ligament were totally ruptured.

Fig. 14. A boxplot diagram of the medial clear space with different MRI findings of the deep deltoid ligament (p < 0.001).
6 Discussion

6.1 General considerations

All of our studies benefitted from prospectively collected and reasonably large patient cohorts compared to previous studies (Michelson et al. 2001, Gill et al. 2007, Schock et al. 2007, Koval et al. 2007, Cheung et al. 2009). The blinded performance of stress tests, and their blinded analysis with respect to fracture morphology, clinical signs, and MRI findings, were also clear strengths. The inter-observer reliability was excellent for the analysis of all radiographs: for both the standard and the stress radiographs. The inclusion of two main trauma centers in Study I was an asset itself, but it also provided enough participants to enable multiple regression analyses with reasonably narrow CIs. Study II was the first study to report the performance statistics of the gravity stress test, and the use of LRs were strengths, since LRs are not affected by the incidence of unstable ankles but predictive values are (Bland & Altman 1986). Study III allowed us to demonstrate for the first time an association between the ER stress stability and the result of the MRI assessment. Also, the use of a 3.0 T MRI device was a strength compared to previous studies using a 1.5 T device (Koval et al. 2007, Cheung et al. 2009).

Our studies had some limitations. The mechanism of injury was not systematically recorded. However, isolated SER-type lateral malleolar ankle fractures commonly derive from low-energy mechanism (Briet et al. 2017) and we excluded high-energy polytrauma patients. ER stress tests were performed before gravity testing, and this might have had an effect on the results of the gravity tests. Clinical signs were recorded only as a consensus decision of two surgeons and the clinical examination was not always performed within 24 hours after injury. The gravity stress radiography was performed only once per patient, but it is unlikely that the result would have changed after repeated radiographs due to the stationary setting. The ER stress radiographs in all three studies were analyzed at least twice but only by a single investigator. However, the gravity stress radiographs in Study II were analyzed by two surgeons with excellent inter-observer reliability. Since the measurement of MCS is identical in both mortise projections, it is unlikely that the reliability of the ER stress test analysis would have differed from that of the gravity stress radiographs in any of the studies. The CIs were rather wide in Study II, indicating that the results must be evaluated
cautiously. The low rate of total tears of the deep deltoid ligament in Study III limited the statistical significance of the results but was in line with previous studies (Koval et al. 2007, Cheung et al. 2009). We used a previously unpublished MRI classification since the MRI findings of the deep deltoid ligament were more distinguishable in the 3.0 T MRI (Chhabra et al. 2010) compared to the 1.5 T MRI (Mengiardi et al. 2007). This could be accounted as a limitation although our four-grade scale probably weakened the association between the stress test stability and the MRI finding of the deep deltoid ligament compared to the previously reported two- or three-grade scales (Schottel et al. 2015, Koval et al. 2007, Cheung et al. 2009).

6.2 Standard radiographs in evaluation of stability

We found that the displacement of the fibular fracture on lateral non-stress radiographs, the number of fracture fragments, and the sex were independent risk factors for instability in SER-type isolated lateral malleolus fractures without medial widening of the ankle mortise. Non-comminuted fractures with minimal displacement (< 2 mm) on lateral radiographs comprised 30% of our study population. Our finding, that these ankles could be diagnosed as stable without further stress testing, supported our hypothesis. As previously reported by van den Bekerom & van Dijk (2010), we found no association between the fracture line width on the mortise radiographs and the ankle mortise stability. However, the association was strong between the latter and the posterior diastasis on lateral radiographs in our study. This may be explained by the rotational nature of the injury, as the distal fracture fragment follows the externally rotating talus, while the proximal fibula is medialized and internally rotated (Michelson et al. 1992, Donken et al. 2013). Therefore, the mortise view reveals this proximal medialization and internal rotation, but the posterior diastasis on lateral radiographs should be a better sign of external rotation of the talus.

In our study, men were on average 10 years younger, had more comminuted fractures, and their risk for an unstable injury was also higher. It is likely that younger males have better bone quality than older males and females, and they are probably more active and more exposed to higher trauma energy. Therefore, sex may be just a surrogate factor for trauma energy, since men with better bone quality are likely to sustain more severe ligamentous injuries prior to fibular fracture (Choi et al. 2014).
6.3 Diagnostic accuracy of gravity stress radiography and clinical signs

Previous studies with fewer patients have suggested that the ER stress test could be replaced by gravity stress radiography (Michelson et al. 2001, Gill et al. 2007, Schock et al. 2007). Our results show that the gravity stress and the ER stress test results were in agreement in the majority of cases, but the LRs of the former were only moderate. In fact, our study was the first to report the diagnostic test performance statistics of gravity stress radiography. Our results did not confirm our hypothesis that ER stress test could be replaced by the gravity stress test. We agree with previous studies (Egol et al. 2004, DeAngelis et al. 2007, McConnell et al. 2004, Clements et al. 2008) that medial tenderness, ecchymosis and swelling alone are insufficient for predicting the instability of the ankle mortise. However, when the result of the gravity stress test was in agreement with the medial signs, indirectly suggesting a deltoid ligament injury, an equivalent result with the ER stress test was highly probable. Delay between injury and clinical evaluation had an effect on swelling and ecchymosis but not on medial tenderness when evaluated within a week from injury. Medial tenderness was not affected by patient age and seemed to be the most accurate of the clinical signs in predicting the stability of the ankle mortise. Therefore, the utility of these clinical signs is similar whether evaluated within a week of injury or, as described in previous studies (Egol et al. 2004, DeAngelis et al. 2007, McConnell et al. 2004), within 24 hours of injury.

6.4 Utility of MRI in stability assessment of ankle mortise

Our 3.0 T MRI findings show that the majority of patients with an isolated lateral malleolus SER-type fracture without static widening of the MCS had a partially torn deep deltoid ligament regardless of the ER stress test results. In fact, according to our MRI results, the deep deltoid ligament was affected in every case. Our findings that the AITFL was torn in every case and that total tears of the deep deltoid complex were rare even in ER stress positive ankles are in line with previous reports (Koval et al. 2007, Cheung et al. 2009). We found an association between the MCS under ER stress and the MRI finding of the deep deltoid ligament; this has not been reported previously, probably due to the lack of a patient group with a stable ankle mortise and small sample sizes in previous studies (Koval et al. 2007, Cheung et al. 2009). Although these findings confirm
our hypothesis that the deep deltoid ligament is substantially injured only in unstable ankles, they also contribute to the paradigm that the stability of the ankle mortise is a continuum rather than a binary (stable/unstable) phenomenon.

Schottel et al. (2015) assessed the ability of the ER stress test to distinguish stable SER II/III-injuries from unstable SER IV-injuries. Their results were reported almost simultaneously with ours, but they used the MRI as a reference standard and thus concluded that the manual ER stress test had a poor ability to predict deep deltoid ligament integrity. They suggested that an MRI study or intraoperative stress testing should be performed on the patients with a manual stress MCS of 4.0 to 5.5 mm—a rather costly and time-consuming policy considering that two-thirds of our study population fit into this category. The clinical use of the MRI is further challenged by the moderate inter-observer agreement of the MRI assessment in both of these studies. However, according to these MRI studies, the cadaver study based stress MCS cut-off value of 5 mm (Park et al. 2006) seems to overestimate the total tears of the deep deltoid ligament.

The use of delayed weight bearing radiographs were proposed after the onset of our study as a low-cost and pain-free method to evaluate the stability of the ankle mortise and to guide the decision-making between the operative and nonoperative treatment of unimalleolar SER fractures without static incongruity (Weber et al. 2010, Hoshino et al. 2012). Unstable (SER IV) ankles seem to be rare on weightbearing radiographs, comprising only 3 to 9% of ER stress test positive ankles (Hoshino et al. 2012, Weber et al. 2010). Interestingly, we found that only 6% of both manual ER stress and gravity stress (unpublished data on the latter) positive patients had a total tear of the deep deltoid ligament according to the MRI assessment. These numbers are surprisingly similar and support the argument that the use of both the manual ER and the gravity stress tests might seriously overestimate the injury of the deep deltoid ligament (Weber et al. 2010, Hoshino et al. 2012, Dawe et al. 2015). However, there are no studies comparing the results of weightbearing radiographs and the MRI. The use of weightbearing radiographs can be criticized due to its feature of assessing rotational injury with axial loading. Donken et al. (2013) reported the talar rotation on CT to be severe even in cadaver models with low-grade injuries; however, the widening of the MCS manifested in their non-weightbearing SER IV models only when external rotational force was applied. The axial loading can reduce the talus under the concave tibia to some degree even in the presence of a notable deep deltoid ligament tear (Stormont et al. 1985, Richter et al. 2003, Tochigi et al. 2006).
Furthermore, an association between the MCS and the status of the deep deltoid ligament was not found in a weightbearing cadaveric model (Stewart et al. 2012).

6.5 Clinical implications and future studies

In conclusion, our study suggests that isolated non-comminuted SER fractures of the lateral malleolus without talar shift on standard radiographs are stable and thus need no stress testing if the posterior diastasis on lateral radiographs is < 2 mm. We further conclude that an equivalent result in both the gravity stress test and the ER stress test is highly probable but only if clinical signs on the medial side of the ankle are in agreement with the result of the stress test. The MRI provides no additional value compared to ER stress testing in the evaluation of ankle mortise stability in SER fracture patients without incongruity on standard radiographs.

Previous biomechanical studies have shown that even a millimeter of talar shift can substantially reduce the tibio-talar contact area (Ramsey et al. 1976, Pereira et al. 1996). The long-term effect of this subtle talar shift on the development of osteoarthritis in ankles without static widening of the MCS is unknown and will be difficult to prove since posttraumatic osteoarthrosis of the ankle might take up to 20 years to develop (Horisberger et al. 2009). Intelligibly, finding clinically relevant differences in the outcomes of these overall well-healing stress test positive ankles with only relative instability of the ankle mortise is challenging. The evidence is still lacking regardless of the method used to reveal a possible dynamic instability of the ankle mortise, since no significant differences were found in the functional scores or pace of recovery after one-year follow-up in the only randomized study comparing the operative and nonoperative treatment of patients with a positive stress test, the manual ER stress test in this case (Sanders et al. 2012). However, there were more patients with radiological misalignment and/or delayed fracture union in the nonoperative group. The ORIF of these ER stress positive ankles was not cost effective in the one-year time horizon but would be if it could decrease the lifetime incidence of posttraumatic ankle arthrosis by >3% (Slobogean et al. 2012).

Our MRI study showed that total tears of the deep deltoid complex were rare in patients with isolated lateral malleolar Weber B-/SER-type fractures in the absence of talar shift on standard radiographs. Consequently, it is not rational to use the MRI as a screening test. According to the MRI findings, both the ER and the gravity stress tests overestimated the number of total deep deltoid tears, and therefore a positive result in either of the above stress tests may lead to an
unnecessary ankle fracture operation. Similarly, only a small portion of ankles with positive clinical signs had a total tear of the deep deltoid ligaments. Considering our findings and the previous literature, attempts to ultimately reduce the number of operatively treated stress test positive patients but also better targeting of the diagnostic resources seem justified. Not all isolated lateral malleolar fractures without static talar shift need stress testing, in contrast to the previous suggestion by Michelson et al. (2007), since we found no patient with < 2 mm of posterior diastasis on lateral radiographs or negative clinical signs to have a total tear of the deep deltoid ligament complex.

There are some patients with considerable talar shift that is only evident in a stress test. These patients probably have total tears of the deltoid ligament complex and are the true candidates for operative treatment. An argument for the use of the ER or the gravity stress test can also be made due to recent evidence that do not indicate the need for further clinical follow-up or even favor the use of cast immobilization over orthosis after a negative ER stress test (Kortekangas 2017). The deployment of these findings in clinical practice would reduce the burden and costs to healthcare systems. However, studies using weightbearing radiographs as the reference for stability have had similar results with a considerably smaller proportion of operatively treated patients (Weber et al. 2010, Hoshino et al. 2012, Hastie et al. 2015, Seidel et al. 2017). Although the need for a second visit to obtain the weightbearing radiographs is necessary within a couple of weeks after injury, the savings would probably be considerably greater not only due to a 10-fold decrease in the number of initial operations but also due to a similar decrease in the number of secondary operations and complications, respectively (Dawe et al. 2015).

The need for well designed and conducted clinical ankle fracture studies with adequate power and long-term follow-up is obvious. There is insufficient short-term evidence to support the use of any tests to reveal the potential dynamic instability or the operative treatment of isolated lateral malleolar Weber B-type fractures (Sanders et al. 2012, Mittal et al. 2017). Actually, due to the lack of high quality studies it is unclear whether surgery is superior to nonoperative treatment in the long term regardless of the ankle fracture type (Donken et al. 2012). The incidence of total tears of the deep deltoid ligaments in patients with an isolated unimalleolar SER fracture seems to be quite similar according to our MRI results and in studies using weightbearing radiographs (Weber et al. 2010, Hoshino et al. 2012). Therefore, a diagnostic study that compares these two methods would be beneficial.
7 Conclusions

The following conclusions can be made based on the results of this thesis:

1. For a fracture line width < 2 mm in lateral radiographs, only two fracture fragments and female sex are independent factors that predict a stable ankle mortise. Patients with non-comminuted fractures and < 2 mm diastasis on lateral radiographs have stable ankle mortises and need no further stress testing.

2. When gravity stress test results are in agreement with clinical findings, the result is likely to predict stability of the ankle mortise with an accuracy equivalent to ER stress test results.

3. The use of MRI provides no additional benefit compared to ER stress testing for decision-making between the operative and nonoperative treatment of a SER-type ankle fractures.
References


Diagnostic test accuracy is often described based on its discriminative ability, usually in terms of sensitivity and specificity. The information conveyed by these parameters is not very clinically useful, but the two parameters can be combined to calculate the likelihood ratio, which is easier to interpret for clinical use. Likelihood ratios can also easily be converted to predictive values (probabilities) to rule-in or rule-out a diagnosis. The probability estimated before any testing (i.e., the prevalence of the condition or pre-test probability) will either increase or decrease after the test (the post-test probability of the condition). Bayes theorem, a mathematical relationship, allows calculation of the probability of a condition if the pre-test probability and the likelihood ratio of the test are known. When the likelihood ratios of every test are known, it is easy to calculate the probability of the condition after consecutive tests simply by multiplying the pre-test odds with the likelihood ratio of every test.

Definitions of terms and equations used with regard to likelihood ratios in diagnostic tests:

Pre-test probability = the probability of disease before any clinical examination or tests. (In this study, the incidence of ER stress test (reference)-positive patients was used.)

Probability (P) = ODDS/ODDS + 1, ODDS = P/1 − P.

Likelihood ratio (LR) = probability of finding in patients with disease / probability of same finding in patients without disease.

Positive LR = probability of an individual with the condition having a positive test / probability of an individual without the condition having a positive test.

= sensitivity / (1 − specificity)

Negative LR = probability of an individual with the condition having a negative test / probability of an individual without the condition having a negative test

= (1 − sensitivity) / specificity

Post-test ODDS = pre-test ODDS × LR (of test 1) × LR (test 2) × … LR (test…)

71
<table>
<thead>
<tr>
<th>LR</th>
<th>Interpretation (according to Jaeschke et al. 1994)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 10</td>
<td>Large and often conclusive increase in the likelihood of disease</td>
</tr>
<tr>
<td>5–10</td>
<td>Moderate increase in the likelihood of disease</td>
</tr>
<tr>
<td>2–5</td>
<td>Small increase in the likelihood of disease</td>
</tr>
<tr>
<td>1–2</td>
<td>Minimal increase in the likelihood of disease</td>
</tr>
<tr>
<td>1</td>
<td>No change in the likelihood of disease</td>
</tr>
<tr>
<td>0.5–1.0</td>
<td>Minimal decrease in the likelihood of disease</td>
</tr>
<tr>
<td>0.2–0.5</td>
<td>Small decrease in the likelihood of disease</td>
</tr>
<tr>
<td>0.1–0.2</td>
<td>Moderate decrease in the likelihood of disease</td>
</tr>
<tr>
<td>&lt; 0.1</td>
<td>Large and often conclusive decrease in the likelihood of disease</td>
</tr>
</tbody>
</table>
Original publications

This thesis is based on the following publications, which are referred to throughout the text by their Roman numerals:


Reprinted with permission from Wolters Kluwer (I, III) and the British Editorial Society of Bone & Joint Surgery (II).

Original publications are not included in the electronic version of the dissertation.
1431. Lavander, Päivi (2017) Nimikesuojattujen ja laillistettujen ammattihenkilöiden työnjakoin yliopistosairaalain muuttuvassa toimintaympäristössä
1434. Hulkko, Anja (2017) The association of lifetime antipsychotic and other psychiatric medications with cognition in schizophrenia : the Northern Finland Birth Cohort 1966 Study
1435. Ramsay, Hugh (2017) Predictors of psychosis risk and neurocognitive deficits
1436. Kuitunen, Hanne (2017) DLBCL, primary and secondary central nervous system involvement, treatment and prophylaxis
1437. Filatova, Svetlana (2017) Incidence of schizophrenia and associations of schizophrenia and schizotypy with early motor developmental milestones
1438. Kärjäjärvi, Aki (2017) Non-alcoholic fatty liver disease (NAFLD) : perspectives to etiology, complications and lipid metabolism
1440. Hagnäs, Magnus (2018) The association of cardiorespiratory fitness, physical activity and ischemic ECG findings with coronary heart disease-related deaths among men
1441. Huhtaniska, Sanna (2018) The association between antipsychotic and benzodiazepine use with brain morphology and its changes in schizophrenia
1442. Sundquist, Elias (2018) The role of tumor microenvironment on oral tongue cancer invasion and prognosis

Book orders:
Granum: Virtual book store
http://granum.uta.fi/granum/
Simo Nortunen

STABILITY ASSESSMENT OF ISOLATED LATERAL MALLEOLAR SUPINATION-EXTERNAL ROTATION-TYPE ANKLE FRACTURES