Limited intrarater and interrater reliability of acute ligamentous ankle injuries on 3 T MRI

Thomas P A Baltes 1,2,3,4, Javier Arnaiz, 5 Maryam R Al-Naimi, 5 Omar Al-Sayrafí, 6 Celeste Geertsema, 6 Liesel Geertsema, 6 Toni Evans, 5 Pieter D’Hooghe, 7 Gino M M J Kerkhoffs, 2,3,4, Johannes L Tol 2,3,4,6

ABSTRACT

Objectives To determine the diagnostic reliability of the Schneck grading system for acute ligamentous injuries of (1) the three major ligamentous ankle complexes, (2) the individual ankle ligaments and (3) the Sikka classification for syndesmotic injury.

Methods All acute ankle injuries in adult athletes (≥18 years), presenting to the outpatient department of a specialised Orthopaedic and Sports Medicine Hospital, within 7 days postinjury were screened for inclusion. Ankle injuries were excluded if imaging demonstrated a frank ankle fracture or if the 3 T MRI study could not be acquired within 10 days postinjury. Two radiologists graded the three major ligamentous complexes (lateral ankle complex, deltoid complex and syndesmosis complex) and their comprising individual ligaments according the four-grade Schneck grading system. Syndesmotic injuries were classified according the four-grade Sikka classification for consequence injury of the individual syndesmosis ligaments and the deltoid complex. Agreement and kappa (K) statistics were calculated to determine intrarater and interrater reliability.

Results Between September 2016 and September 2018, a total of 92 MR scans were obtained (87 patients). Interrater and intrarater reliability of the Schneck grading system was moderate to substantial for the lateral ankle complex (K=0.47–0.76), fair to almost perfect for the syndesmosis complex (K=0.37–0.89) and fair to moderate for the deltoid complex (K=0.14–0.51). For the individual ligaments, kappa values ranged from moderate to substantial for the anterior talofibular ligament (ATFL) (K=0.55–0.73), fair to substantial for the calcaneofibular ligament (K=0.31–0.62) and fair to almost perfect for the anteroinferior tibiotalar ligament (AITFL) (K=0.36–0.89). Diagnostic reliability of the Sikka classification ranged from moderate to almost perfect (K=0.51–0.95).

Conclusions Grading of the three major ligamentous complexes and of the individual ankle ligaments according the Schneck grading system resulted in limited diagnostic reliability. When dichotomised for the presence of complete discontinuity, the interrater reliability of the Schneck grading system improved to substantial and almost perfect for the ATFL and AITFL, respectively. Classification of syndesmosis injury according the Sikka classification resulted in moderate interrater reliability.

INTRODUCTION

Acute ankle sprains are among the most common sport-related injuries.1 The lateral ankle ligaments are most frequently injured (0.93/1000 athlete exposures), followed by the syndesmosis (0.38/1000 athlete exposures) and deltoid ligaments (0.06/1000 athlete exposures).2 In athletes, MRI is increasingly used for initial diagnosis and prognosis of ligamentous ankle injuries.3–5 To categorise and translate these MR findings into clinical practice, standardised grading systems with high diagnostic reliability are warranted.

Diagnostic reliability of standardised grading systems for acute ligamentous ankle injuries have been described in various studies.5–8 However, only the grading system used by Roemer et al included grading of injury in multiple ligamentous complexes (lateral ankle complex, deltoid complex and syndesmosis complex).9 In this study, two radiologists determined intrarater and interrater reliability of a five-grade system for acute and chronic ligamentous ankle injury, based on 30 MR scans (1.5 T).9 The main limitation in this study was that it reported diagnostic reliability per ligamentous complex (eg, lateral ligaments) and not per individual ligament (eg, anterior talofibular ligament (ATFL)), leaving diagnostic reliability of scoring acute injury of individual ankle ligaments on 3 T MRI unknown.

Reliability of prognostic scoring for syndesmotic injury has been evaluated in two previous studies.7,10 In a retrospective cohort study by Howard et al, prognostic scoring of syndesmotic injury in 16 NFL players resulted in fair to almost perfect interobserver reliability. However, except for syndesmotic joint width, no association between prognostic scoring and time to return to play was established. Sikka et al evaluated a prognostic syndesmosis injury classification in a retrospective cohort study on 36 NFL players with MRI-confirmed (1.5 T) syndesmosis injury. The main limitation in this
The study was that it lacked evaluation of the classifications’ inter-rater reliability.

Given these two limitations, a diagnostic reliability study on grading of the individual ankle ligaments, the ligamentous complexes and classification of syndesmotic injury severity is warranted. Therefore, the aim of this study was to determine the diagnostic reliability of the Schneck grading system for acute ligamentous injuries of (1) the three major ligamentous ankle complexes, (2) the individual ankle ligaments and (3) the Sikka classification for syndesmosis injury.

METHODS

Patient selection

Patients presenting to the outpatient department of a specialised Orthopaedic and Sports Medicine Hospital within 7 days after an acute ankle injury were asked to participate in this study. Inclusion criteria were: acute ankle injuries in adult athletes (≥18 years), participating in sports at a professional or recreational level and presenting within 7 days of injury. Ankle injuries were excluded if imaging studies demonstrated a frank ankle fracture or if the 3 T MRI study could not be acquired within 10 days postinjury. After clinical history and physical examination was performed by a Sports Medicine Physician or Orthopaedic Surgeon, MR images were obtained. Written informed consent was obtained from all patients at time of inclusion.

MRI

All MR scans were obtained using a wide-bore 3.0 T MRI system (GE Discovery, GE Healthcare) with an 8-channel receive only Foot & Ankle array (Invivo, Philips Healthcare, Best, The Netherlands). In the sagittal plane T1-weighted (repetition time (TR) 400–680 ms; echo time (TE) 10–11 ms; 3.0 mm slice thickness; 0.5 mm interslice gap; 416×288 pixel matrix; two excitations (NEX)) 16 cm² field of view (FOV); Echo train length (ETL) 3) and Proton Density Fat Saturated (PD-S) (TR 2500–3200 ms, TE 32–35 ms, 3.0 mm slice thickness, 0.5 mm interslice gap; 352×526 pixel matrix; 2 NEX; 20 cm² FOV; ETL 8) sequences were obtained. In the axial plane T2-weighted (TR 5500–6700 ms; TE 72–80 ms; 3.5 mm slice thickness, 0.5 mm interslice gap; 320×224 pixel matrix; 2 NEX; 13 cm² FOV; ETL 16) and PD-FS (TR 2900–4000; TE 35–39 ms; 3.5 mm slice thickness, 0.5 mm interslice gap; 320×224 pixel matrix; 2 NEX; 13 cm² FOV; ETL 6) sequences were acquired. In the coronal plane a PD FS (TR 2700–3400; TE 35–38 ms; 3.5 mm slice thickness; 0.5 mm interslice gap; 320×224 pixel matrix; 2 NEX; 16 cm² FOV; ETL 6) sequence was obtained.

Table 1  Schneck grading system and Sikka classification for grading of the ligamentous complexes and individual ligaments

| Schneck grading system | Grade 0 | No abnormality of ligament |
|------------------------|---------|--------------------------|
| Grade 1                |         | Periligamentous high signal/oedema on proton density-weighted sequences |
| Grade 2                |         | Partial discontinuity but preserved remnant fibres |
| Grade 3                |         | Complete discontinuity |
| Sikka classification*  |         |                           |
| Grade I                |         | Isolated injuries to the ATIFL |
| Grade II               |         | Injury to the ATIFL, interosseous ligaments and interosseous membrane |
| Grade III              |         | Injury to the ATIFL, interosseous ligament, interosseous membrane and PITFL |
| Grade IV               |         | Injury to the ATIFL, interosseous ligament, interosseous membrane, PITFL and deltoid ligament |

*Ligamentous injury was defined as partial or complete tear of the respective ligament.

ATIFL, anteroinferior tibiofibular ligament; PITFL, posteroinferior tibiofibular ligament.

one radiologist (R1) repeated the scoring process. To minimise recall bias, the radiologist repeated scoring after a period of 28 days.

Grading system for ligamentous complexes and individual ligaments

The ligamentous complexes were graded as normal (grade 0) or in accordance with the highest graded acute lesion (grade 1–3) in one of its comprising individual ligaments. All individual ligaments were graded according to the four grade Schneck grading system11 (table 1).

Classification of syndesmotic injury

In patients with an observed syndesmotic injury, the severity of the syndesmotic injury was classified in accordance to the classification proposed by Sikka et al10 (table 1).

Grading of ligamentous complexes

The three major ligamentous complexes (lateral ankle complex, deltoid complex (subdivided in deep deltoid and superficial deltoid) and syndesmosis complex) were graded according the four-grade Schneck grading system.

Grading of individual ligaments

The following individual ankle ligaments were graded according to the four grade Schneck grading system:

- Lateral ligaments; including the the anterior talofibular ligament (ATFL), calcaneofibular ligament (CFL) and posterior talofibular ligament (PTFL) (figure 1).
- Syndesmosis ligaments; including the anteroinferior tibiofibular ligament (ATIFL), interosseous ligament (IOL), interosseous membrane (IOM), posteroinferior tibiofibular ligament (PITFL) and transverse tibiofibular ligament (TTFL) (figure 2).
- Medial ligaments; subdivided in the superficial deltoid ligaments (tibionavicular [TN], tibiospring [TS], tibiocalcaneal [TC] and posterior tibiotalar [PT], respectively) and the ligaments comprising the deep portion of the deltoid (anterior tibiotalar [ATT] and posterior tibiotalar [PTT]) (figure 3).
Presence versus absence of acute ligamentous lesions
To assess the intrarater and interrater reliability for the presence or absence of acute ligamentous lesions in the ligamentous complexes and individual ligaments, the MRI grading system was evaluated as dichotomous outcomes;

► Grade 0: was considered absence of an acute lesion.
► Grade 1–3: was considered presence of an acute lesion.

Presence versus absence of complete discontinuity
To assess the intrarater and interrater reliability for the presence or absence of complete discontinuity in the ligamentous complexes and individual ligaments, the MRI grading system was evaluated as dichotomous outcomes;

► Grade 0–2: was considered absence of complete discontinuity.
► Grade 3: was considered as presence of complete discontinuity.

Statistical analysis
Descriptive statistics was used to present patient demographics (age, time to MRI, sports) and the number and distribution of lesions graded by the individual observers. Continuous variables were presented as mean with SD for data with a normal distribution and as median with IQR in case of non-normal distribution. Categorical data were presented as frequencies and proportions.

Intrarater and interrater reliability of the Schneck grading system \(^{10}\) (ligamentous lesions; grade 0–3) and Sikka classification system \(^{11}\) (syndesmosis injury; grade I–IV) were determined using linear weighted kappa statistics on an ordinal scale (K). Intrarater and interrater reliability for dichotomised data were determined using unweighted kappa statistics.

Overall agreement was calculated for dichotomous observations and weighted agreement was calculated for ordinal variables. We calculated prevalence (P) and bias index from cross tabulations for the dichotomous variables. Prevalence was defined as percentage (%) of included ankle injuries with positive findings. Bias index was defined as the extent to which the radiologists disagreed on the proportion of positive (or negative) findings. \(^{12}\)

Reliability was considered poor if <0, slight 0–0.20, fair 0.21–0.40, moderate 0.41–0.60, substantial 0.61–0.80 and almost perfect if 0.81–1.00. \(^{13}\) Statistical analysis was performed using Statistical Package for Social Sciences (SPSS V21.0, Chicago, Illinois, USA). Weighted agreement was calculated using Stata Statistical Software, Release 11 (Stata, College Station, Texas, USA).

RESULTS
Baseline characteristics
Between September 2016 and September 2018, a total of 115 acute ankle injuries (110 athletes) were assessed for eligibility (figure 4). Ninety-two ankles were included. Of these 92 imaged acute ankle injuries, 4 were subsequent contralateral ankle injuries and 1 case of reinjury (>1 year). The median age at time of injury was 23 years (IQR 20–27), with a range from 18 to 42 years. The median time from injury to MRI was 3 days (IQR 1–5). Of the 87 included athletes, 47% played football, 14% volleyball, 14% basketball, 11% futsal, 5% athletic and 7% participated in other sports.

Grading of ligamentous complexes
The distribution of acute ligamentous complex lesions (Schneck grades 1–3) as graded by both radiologists, ranged from: lateral complex 72.8%–87.0%; syndesmosis complex 18.5%–68.5%; deep deltoid complex 4.4%–62.0%; superficial deltoid complex 5.4%–43.5% (table 2; online supplemental appendix). Intrarater and interrater reliability of the Schneck grading system for the three ligamentous complexes was, respectively, moderate to almost perfect (K=0.51–0.89) and slight to moderate (K=0.14–0.47).

Classification of syndesmotic injury
The distribution of the Sikka classification for both radiologists is reported in the online supplemental appendix (table 2; online supplemental appendix). Use of this classification system in patients with syndesmosis injury resulted in almost perfect intrarater reliability (k=0.95) and moderate interrater reliability (K=0.51).

Grading of individual ligaments
The distribution of acute ligamentous lesions (Schneck grades 1–3) per individual ligament, as graded by both radiologists, is detailed in the online supplemental appendix (table 2, online supplemental appendix). Grading of the individual lateral ankle ligaments resulted in substantial intrarater reliability (K=0.62–0.73) and slight to moderate interrater reliability (K=0.14–0.55). For
the individual syndesmosis ligaments intrarater reliability ranged from substantial to almost perfect (K=0.63–0.94) and interrater reliability ranged from poor to moderate (K=−0.02 to 0.56). Intra rater reliability for the deltoid ligaments ranged from fair to substantial (K=0.27–0.69) and interrater reliability ranged from slight to fair (K=0.01–0.24).

**Presence versus absence of acute ligamentous lesions**

Acute ligamentous lesions (Schneck grades 1–3) were most frequently scored in the ATFL (71.7%–83.7%), CFL (50.0%–82.6%) and AITFL (18.5%–68.5%) (table 3, online supplemental appendix). Grading of the ATFL and CFL for acute lesions of the individual ligaments resulted in substantial intrarater reliability (K=0.61–0.70) and fair interrater reliability (K=0.26–35), respectively. Almost perfect intrarater reliability (K=0.82) and slight interrater reliability was observed for acute lesions of the AITFL (K=0.17). Intrarater and interrater reliability of the ligamentous complexes ranged from fair to almost perfect (K=0.37–0.82) and slight to fair (K=0.05–0.31).

**Presence versus absence of complete discontinuity**

Complete ligamentous discontinuity (Schneck grade 3) was scored most frequently in the ATFL (39.1%–48.9%), CFL (9.8%–53.3%) and AITFL (13.0%–14.1%) (table 3; online supplemental appendix). Moderate to substantial intrarater reliability was established for complete discontinuity of the ATFL and CFL (K=0.56–0.68) and interrater reliability ranged from fair to substantial (K=0.29–0.63). Almost perfect intrarater and interrater reliability was observed for the AITFL (K=0.91–0.95). Intrarater and interrater reliability of the ligamentous complexes ranged from substantial to almost perfect (K=0.66–1.00) and poor to moderate (K=−0.03–0.87).

![Figure 3](image-url) **Figure 3** Acute deltoid injury. (A) Coronal PD-FS image showing complete discontinuity (Schneck grade III) of the tibiocalcaneal ligament (TC; arrow). (B) Axial PD-FS image showing tibionavicular ligament (TN; arrow) detachment (Schneck grade III). (C) Coronal PD-FS image showing complete discontinuity (Schneck grade III) of the posterior tibiotalar ligament (PTT; arrow). PD-FS, proton density fat saturated

![Figure 4](image-url) **Figure 4** Flowchart. *In 110 athletes. †In 87 athletes.
Table 2  
Intrarater and interrater reliability of grading ligamentous complexes and individual ligaments according the four grade Schneck grading system and classification of syndesmosis injury according the four grade Sikka classification system

| Ligament Complex | Intrarater N (lesions) | Kappa (95% CI) Weighted agreement (%) | Interrater N (lesions) | Kappa (95% CI) Weighted agreement (%) |
|------------------|------------------------|--------------------------------------|------------------------|--------------------------------------|
| Lateral complex  |                        |                                      |                        |                                      |
| Schneck grading of ligamentous complexes |                        |                                      |                        |                                      |
| Lateral | 67 | 0.76 (0.65 to 0.86) | Substantial 88.77 | 67 | 0.47 (0.34 to 0.60) | Moderate 76.81 |
| Deltoid complex | 10 | 0.51 (0.18 to 0.84) | Moderate 94.57 | 10 | 0.14 (0.04 to 0.24) | Slight 68.84 |
| Superficial | 5 | 0.65 (0.30 to 1.00) | Substantial 97.83 | 5 | 0.19 (0.03 to 0.35) | Slight 80.80 |
| Deep | 6 | 0.42 (−0.03 to 0.87) | Moderate 96.38 | 6 | 0.07 (−0.03 to 0.16) | Slight 72.83 |
| Syndesmosis complex | 18 | 0.89 (0.80 to 0.99) | Almost perfect 97.10 | 18 | 0.37 (0.23 to 0.50) | Fair 71.74 |
| Sikka classification | 16 | 0.95 (0.90 to 1.00) | Almost perfect 99.18 | 16 | 0.51 (0.37 to 0.66) | Moderate 89.40 |
| Schneck grading of individual ligaments |                        |                                      |                        |                                      |
| Lateral ankle ligaments |                        |                                      |                        |                                      |
| ATFL | 66 | 0.73 (0.62 to 0.84) | Substantial 87.68 | 66 | 0.55 (0.43 to 0.68) | Moderate 80.07 |
| CFL | 46 | 0.62 (0.49 to 0.75) | Substantial 84.78 | 46 | 0.31 (0.20 to 0.41) | Fair 63.41 |
| PTFL | 2 | N/A | N/A | 2 | 0.14 (−0.05 to 0.33) | Slight 84.78 |
| Superficial deltoid ligaments |                        |                                      |                        |                                      |
| TN | 4 | 0.56 (0.10 to 1.00) | Moderate 97.83 | 4 | 0.16 (0.02 to 0.30) | Slight 83.70 |
| TS | 5 | 0.69 (0.35 to 1.00) | Substantial 98.19 | 5 | 0.24 (0.06 to 0.42) | Fair 86.59 |
| TC | 3 | 0.32 (−0.19 to 0.83) | Fair 97.83 | 3 | 0.07 (−0.05 to 0.19) | Slight 82.97 |
| PT | 3 | 0.38 (−0.17 to 0.94) | Fair 96.74 | 3 | 0.01 (−0.08 to 0.10) | Slight 80.98 |
| Deep deltoid ligaments |                        |                                      |                        |                                      |
| ATT | 6 | 0.48 (0.02 to 0.95) | Moderate 97.10 | 6 | 0.07 (−0.02 to 0.17) | Slight 73.91 |
| PTT | 2 | 0.27 (−0.20 to 0.74) | Fair 97.28 | 2 | 0.06 (−0.02 to 0.14) | Slight 85.87 |
| Syndesmosis ligaments |                        |                                      |                        |                                      |
| AITFL | 18 | 0.89 (0.80 to 0.99) | Almost perfect 97.10 | 18 | 0.36 (0.22 to 0.49) | Fair 71.74 |
| IOL | 12 | 0.64 (0.43 to 0.86) | Substantial 93.48 | 12 | 0.56 (0.39 to 0.73) | Moderate 89.13 |
| IOM | 7 | 0.63 (0.40 to 0.87) | Substantial 93.48 | 7 | 0.53 (0.32 to 0.73) | Moderate 92.93 |
| PITFL | 10 | 0.94 (0.87 to 1.00) | Almost perfect 99.28 | 10 | 0.45 (0.16 to 0.75) | Moderate 92.93 |
| TTFL | 2 | N/A | N/A | 2 | 0.02 (−0.04 to 0.00) | Poor 96.74 |

The total valid lesions for both radiologists (R1a, R1b, R2) out of an overall total of 92 MR scans are presented (N). Reliability for grading (Schneck) and classification (Sikka) are presented as weighted-kappa (K) and weighted agreement. All values are presented with 95% confidence interval (95% CI).

AITFL, anteroinferior tibiofibular ligament; ATFL, anterior talofibular ligament; ATT, deep anterior tibiotalar; CFL, calcaneofibular ligament; IOL, interosseous ligament; IOM, interosseous membrane; N/A, not applicable; PITFL, posteroinferior tibiofibular ligament; PTFL, posterior talofibular ligament; PTT, deep posterior tibiotalar; TC, tibiocalcaneal; TN, tibionavicular; TS, tibiospring; TTFL, transverse tibiofibular ligament.
Table 3  Dichotomisation of Schneck grading system for (1) presence versus absence of acute lesions and (2) presence versus absence of complete discontinuity of the individual ankle ligaments

| Intrarater | Presence of acute ligamentous lesions | Interrater | Presence of complete discontinuity |
|------------|--------------------------------------|------------|-----------------------------------|
|            | N (lesions) Kappa (95% CI) Landis and Koch Agreement (%) | N (lesions) Kappa (95% CI) Landis and Koch Agreement (%) | N (lesions) Kappa (95% CI) Landis and Koch Agreement (%) |
| Lateral ankle complex | 67 0.73 (0.57 to 0.89) Substantial 89.10 | 67 0.31 (0.10 to 0.53) Fair 77.20 | 39 0.71 (0.56 to 0.85) Substantial 85.90 |
| ATFL | 66 0.70 (0.54 to 0.87) Substantial 88.00 | 66 0.35 (0.14 to 0.57) Fair 77.20 | 38 0.68 (0.53 to 0.84) Substantial 84.80 |
| CFL | 46 0.61 (0.45 to 0.77) Substantial 80.40 | 46 0.26 (0.11 to 0.41) Fair 63.00 | 17 0.56 (0.32 to 0.80) Moderate 89.10 |
| PTTL | 2 0 N/A | N/A | 17 0.39 (0.00 to 0.83) Fair 96.70 |
| Superficial deltoid complex | 5 0.52 (0.15 to 0.89) Moderate 94.60 | 5 0.14 (0.03 to 0.25) Slight 62.00 | 12 0.69 (0.46 to 0.92) Substantial 93.50 |
| TS | 5 0.58 (0.20 to 0.95) Moderate 95.70 | 5 0.20 (0.04 to 0.35) Fair 70.70 | 10 0.94 (0.83 to 1.00) Almost perfect 98.90 |
| TC | 3 0.38 (−0.17 to 0.94) Fair 96.70 | 3 0.01 (−0.11 to 0.13) Slight 71.70 | 10 0.83 (0.73 to 0.93) Almost perfect 98.90 |
| PT | 3 0.38 (−0.17 to 0.94) Fair 96.70 | 3 0.05 (−0.03 to 0.12) Slight 42.40 | 2 0.37 (−0.03 to 0.76) Fair 93.50 |
| Deep deltoid complex | 6 0.42 (0.00 to 0.83) Moderate 94.60 | 6 0.06 (−0.02 to 0.14) Slight 45.70 | 6 0.37 (−0.03 to 0.76) Fair 93.50 |
| ATT | 6 0.38 (−0.17 to 0.94) Fair 96.70 | 6 0.08 (−0.03 to 0.19) Slight 68.50 | 2 0.37 (−0.03 to 0.76) Fair 93.50 |
| Syndesmosis complex | 18 0.82 (0.67 to 0.97) Almost perfect 94.60 | 18 0.17 (0.06 to 0.27) Slight 48.90 | 13 0.67 (0.42 to 0.91) Substantial 93.50 |
| AITFL | 18 0.82 (0.67 to 0.97) Almost perfect 94.60 | 18 0.17 (0.06 to 0.27) Slight 48.90 | 12 0.69 (0.46 to 0.92) Substantial 93.50 |
| IOL | 7 0.67 (0.42 to 0.91) Substantial 93.50 | 7 0.39 (−0.16 to 0.80) Slight 95.70 | 9 0.83 (0.73 to 0.93) Almost perfect 98.90 |
| PITFL | 10 0.94 (0.83 to 1.00) Almost perfect 98.90 | 10 0.54 (0.23 to 0.86) Moderate 93.50 | 2 0.94 (0.83 to 1.00) Almost perfect 98.90 |
| TTFL | 2 0 N/A | N/A | 2 0.94 (0.83 to 1.00) Almost perfect 98.90 |

The total valid lesions for both radiologists (R1a, R1b R2) out of an overall total of 92 MR scans are presented (N). Reliability for grading (Schneck) dichotomised for the presence of acute lesions or complete discontinuity are presented as unweighted-kappa (K) and overall agreement. All values are presented with 95% CI.

ATFL, anteroinferior tibiofibular ligament; ATFL, anterior talofibular ligament; AITFL, anterior tibiofibular ligament; ATT, deep anterior tibiotalar; CFL, calcaneofibular ligament; IOL, interosseous membrane; IOM, interosseous ligament; IOL, interosseous ligament; NA, not applicable; PITFL, posteroinferior tibiofibular ligament; PT, posterior tibiotalar; PTTL, posterior talofibular ligament; PTT, deep posterior talostral; TC, tibiocalcaneal; TN, tibionavicular; TS, tibiospring; TTFL, transverse tibiofibular ligament.
DISCUSSION

In this study, we reported the reliability of the Schneck grading system and the Sikka classification for acute ligamentous ankle injuries on 3 T MRI. Grading of the ligamentous complexes according the Schneck grading system and classification of syndesmotic injury according the Sikka classification resulted in slight to almost perfect reliability. Grading of the individual ankle ligaments according the Schneck grading system resulted in limited reliability. When dichotomised for the presence or absence of complete discontinuity, the interrater reliability of the ATFL andAITFL improved to substantial and almost perfect, respectively.

Grading of ligamentous complexes: comparison with previous literature

For grading of acute ligamentous complex injuries, only two previous studies have reported on the diagnostic reliability of a standardised grading approach. In a prospective study Gaebler et al presented the diagnostic reliability of a grading approach to acute injury of the lateral ligamentous complex. Applied on 0.5 T and 1.0 T MRI, grading resulted in good intrarater reliability ($\kappa = 0.65$) and fair interrater reliability ($\kappa = 0.40$). A more recent study by Roemer et al reported almost perfect intrarater reliability ($\kappa = 0.88$, $\kappa = 0.94$, $\kappa = 0.81$, respectively) and moderate to almost perfect interrater reliability ($\kappa = 0.85$, $\kappa = 0.90$, $\kappa = 0.41$, respectively) for grading of acute and chronic injury of the three ligamentous complexes, which is comparable to our findings.

For grading of all the individual ankle ligaments, no comparable study on acute ankle injuries has been published. For the individual lateral ankle ligaments, one study investigated the reliability of 3 T MRI for acute injury of the ATFL. In this study the diagnostic accuracy and diagnostic reliability for addition of the ‘bright-rim sign’ to the standard diagnostic criteria of ATFL injury were determined. Interrater reliability for acute injury of the ATFL varied widely, depending on the applied definition ($\kappa = 0.48–0.93$).

In chronic lateral ankle ligament injuries, two studies have investigated the reliability of scoring injury to the ATFL. Kim et al reported excellent interrater reliability (Intraclass Correlation Coefficient $\kappa = 0.915$) for detection of presence or absence of injury to the ATFL. In contrast to our study, no grading of injury severity was applied. In another study, grading of chronic injury was reported in a cohort of patients with chronic lateral ankle instability. Grading on a four-grade scale for chronic injury resulted in substantial intrarater reliability ($K = 0.68–0.75$) and moderate to almost perfect interrater reliability ($K = 0.55–0.87$).

For acute deltoid injury, the interrater reliability on 3 T MRI has been reported in one study. In this study, diagnostic reliability of MRI was investigated in a cohort of patients with lateral malleolar fractures secondary to a Supination-External Rotation trauma. The interrater reliability for partial and complete discontinuity of the deep deltoid ligaments ranged from fair to moderate ($K = 0.46$; $K = 0.22$), which is better than that observed in our cohort. In addition, the increased prevalence of deltoid injury in this cohort could potentially have decreased the reported kappa-values.

For syndesmotic injuries, three previous studies have reported on diagnostic reliability of MRI. The main difference with our population is that these studies only included patients with MRI-confirmed syndesmotic injury or with acute ankle fractures and thus an increased prevalence. As increased prevalence comes with high chance agreement, the reported kappa values for syndesmotic injury might be lower than the true kappa value in a non-selected population.

In the study by Hermans et al, patients with an acute ankle fracture underwent 1.5 T MRI. Grading of acute syndesmotic injury demonstrated substantial and almost perfect interobserver reliability for the ATFL ($K = 0.61$) and PITFL ($K = 0.83$). Addition of a 45º oblique MRI-plane improved the interrater reliability for the ATFL to almost perfect ($K = 0.92$). As our multiplane MRI-sequence lacked such a plane, the addition could hypothetically further improve the diagnostic reliability for low-grade injuries of the ATFL.

Presence versus absence of complete discontinuity

In daily clinical practice, the presence of periligamentous oedema or partial discontinuity is less consequential, as the decision for surgical intervention is based on the presence of complete ligamentous discontinuity. Simplified scoring for the presence of complete discontinuity or acute lesions might therefore be more clinically relevant in this setting. Therefore, the four-grade Schneck grading was dichotomised for presence of acute lesions and presence of complete discontinuity. This resulted in improved interrater reliability for complete discontinuity of the ATFL (substantial) andAITFL (almost perfect). As complete discontinuity of these two ligaments has major ramifications in selected patients (eg, athletes), the improved reliability of dichotomised grading might be preferential in the clinical setting.

Strength and limitations

To our knowledge, this study is the largest prospective cohort study on diagnostic reliability of grading acute injuries of all three ligamentous complexes. Its strength lies in its prospective design, broad inclusion criteria (all acute ankle injuries) and use of 3 T MRI. Despite these facts, the study has some limitations. First, the reported reliability of the deltoid ligaments and posterior syndesmosis ligaments (PITFL and TTFL) should be interpreted with caution as the low prevalence of injury potentially influenced the k-values and corresponding CIs. An even larger cohort of athletes might improve the obtained k values and narrow the CI further. Secondly, although discrepancies in MRI grading are inherent, considerable bias was observed for the dichotomization strategy in which ligaments were graded either normal (grade 0) or as having an acute lesion (grade 1–3). The bias indices normalised for the dichotomisation strategy in which ligaments were either not completely discontinuous (grade 0–2) or completely discontinuous (grade 3). This suggests a bias of the second radiologist towards scoring low-grade injuries. Potentially, increased interrater reliability could have been achieved with a more elaborate calibration session. However, the limited calibration of both radiologists should be considered a strength of this study, as it represents daily clinical practice.

Future research should aim to correlate grading of injury severity with return to play prognosis after acute ligamentous ankle injuries. Application of the current available grading systems on 3 T MRI is insufficiently reliable for this purpose. Dichotomised scoring for complete discontinuity of ankle ligaments and additional (angulated) MR-planes could potentially improve interrater reliability; however, additional research is required to substantiate these claims.
Implications for clinical practice

Dichotomised grading (absence or presence of complete discontinuity) of the ATFL and AITFL resulted in substantial to almost perfect interrater reliability. Therefore, when interpreting MRI results of an acute ligamentous ankle injury, reported presence of complete discontinuity of the ATFL and AITFL can be considered reliable. Dichotomised grading (absence or presence of complete discontinuity) of the CFL, resulted in fair interrater reliability. Thus, reported discontinuity of the CFL should be interpreted with caution. In clinical practice, this means that MRI can guide treatment of ligamentous ankle injuries, based on the presence of complete discontinuity of the ATFL and AITFL only. This implies that MR imaging can reliably differentiate acute ankle injuries with and without syndesmotic involvement. Since the AITFL is the first syndesmotic ligament to be injured, a prerequisite for syndesmotic instability is complete discontinuity of the AITFL.11 MR imaging could therefore aid the identification of patients, with an increased probability of syndesmotic instability (e.g., torn MRA-T). The use of MR imaging in the diagnosis of acute ligamentous ankle injuries can therefore be considered most useful in those patients with increased probability of syndesmosis instability.

CONCLUSION

In athletes, grading of the three major ligamentous complexes and the individual ankle ligaments according the Schnecken grading system using 3 T MRI resulted in limited reliability. When dichotomised for the presence of complete discontinuity, the interrater reliability of the Schnecken grading system improved to substantial and almost perfect for the ATFL and AITFL, respectively. Classification of syndesmosis injury according the Sikka classification resulted in moderate inter-rater reliability.

ORCID iD

Thomas P A Baltes http://orcid.org/0000-0002-2396-8984

REFERENCES

1. Feng DT-P, Hong Y, Chan L-K, et al. A systematic review on ankle injury and ankle sprain in sports. Sports Med 2007;37:73–94.
2. Doherty C, Delahunt E, Caulfield B, et al. The incidence and prevalence of ankle sprain injury: a systematic review and meta-analysis of prospective epidemiological studies. 2012;41:193-204.12-40
3. Oae K, Takao M, Naito K, et al. Injury of the tibiofibular syndesmosis: value of MR imaging for diagnosis. Radiology 2003;227:155–61.
4. Takao M, Ochi M, Oae K, et al. Diagnosis of a tear of the tibiofibular syndesmosis. The role of arthroscopy of the ankle. J Bone Joint Surg Br 2003;85:324–9.
5. Oae K, Takao M, Uchino Y, et al. Evaluation of anterior talofibular ligament injury with stress radiography, ultrasonography and MRI imaging. Skeletal Radiol 2010;39:841–7.
6. Gaebler C, Kukla C, Breitenseher MJ, et al. Diagnosis of lateral ankle ligament injuries: comparison between talar tilt, MRI and operative findings in 112 athletes. Acta Orthop Scand 1997;68:286–90.
7. Howard DR, Rubin DA, Hillen Tj, et al. Magnetic resonance imaging as a predictor of return to play following syndesmosis (high) ankle sprains in professional football players. Sports Health 2012;4:53–63.
8. Nortunen S, Lepomäki S, Savolainen O, et al. Stability assessment of the ankle mortise in supination-external rotation-type ankle fractures: lack of additional diagnostic value of MRI. J Bone Joint Surg Am 2014;96:1855–62.
9. Roemer FW, Jomaa N, Niu J, et al. Ligamentous injuries and the risk of associated tissue damage in acute ankle Sprains in athletes. Am J Sports Med 2014;42:1549–57.
10. Sikka RS, Fetzer GB, Suparman E, et al. Correlating MRI findings with disability in syndesmotic sprains of NFL players. Foot Ankle Int. 2012;33:371–8.
11. Schneck CD, Mesgarzadeh M, Bonakdarpour A, Mr MRI of the most commonly injured ankle ligaments. Part II. ligament injuries. Radiology 1992;184:507–12.
12. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther 2005;85:257–68.
13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
14. Lee MH, Cha JG, Lee YK, et al. The bright rim sign on MRI for anterior talofibular ligament injury with arthroscopic correlation. JJR Am J Roentgenol 2012;198:885–90.
15. Kim YS, Kim YB, Kim TG, et al. Reliability and validity of magnetic resonance imaging for the evaluation of the anterior Talofibular ligament in patients undergoing ankle arthroscopy. Arthroscopy 2015;31:1540–7.
16. Morvan A, Klouche S, Thes A, et al. Reliability and validity of preoperative MRI for surgical decision making in chronic lateral ankle instability. Eur J Orth Surg Traumatol 2018;28:713–9.
17. Hermans JI, Beumer A, Hop WIC, et al. Tibiofibular syndromes in acute ankle fractures: additional value of an oblique Mr image plane. Skeletal Radiol 2011;40:193–202.
18. Murati HH, Bicjmojla A, Celebi L, et al. Magnetic resonance arthographic evaluation of syndesmotic diastasis in ankle fractures. Arch Orthopa Trauma Surg 2005;125:222–7.
19. McCollum GA, van den Beeker MPJ, Kerkhoffs GMMJ, et al. Syndesmosis and deltoid ligament injuries in the athlete. Knee Surg Sports Traumatol Arthrosc 2013;21:1328–37.
20. Vuurbegra G, Hoorntje A, van Klouche S, et al. Diagnosis, treatment and prevention of ankle sprains: update of an evidence-based clinical guideline. Br J Sports Med 2018;52:956.
21. Großerlinden LG, Hartel M, Yamamura J, et al. Isolated syndesmotic injuries in acute ankle sprains: diagnostic significance of clinical examination and MRI. Knee Surg Sports Traumatol Arthrosc 2016;24:1180–6.
22. Calder D, Bamford R, Petrie A, et al. Stable versus unstable grade II high ankle Sprains: a prospective study predicting the need for surgical stabilization and time to return to sports. Arthroscopy 2016;32:634–42.