University of Groningen

Validation of the diagnostic criteria of the consensus definition of fracture-related infection

Onsea, Jolien; Van Lieshout, Esther M.M.; Zalavras, Charalampos; Sliepen, Jonathan; Depypere, Melissa; Noppe, Nathalie; Ferguson, Jamie; Verhofstad, Michael H.J.; Govaert, Geertje A.M.; IJpma, Frank F.A.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Link to publication in University of Groningen/UMCG research database

Validation of the diagnostic criteria of the consensus definition of fracture-related infection

Jolien Onsea [a,b], Esther M.M. Van Lieshout [c], Charalampos Zalavras [d], Jonathan Sliepen [e],
Melissa Depypere [f], Nathalie Noppe [g], Jamie Ferguson [h], Michael H.J. Verhofstad [i],
Geertje A.M. Govaert [i], Frank F.A. IJpma [e], Martin A. McNally [h], Willem-Jan Metsemakers [a,b,*]

[a] University Hospitals Leuven, Department of Trauma Surgery, Leuven, Belgium
[b] KU Leuven - University of Leuven, Department of Development and Regeneration, Leuven, Belgium
[c] Trauma Research Unit Department of Surgery, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands
[d] Department of Orthopaedic Surgery, Keck School of Medicine, University of Southern California, Los Angeles, USA
[e] Department of Trauma Surgery, University Medical Center Groningen, Groningen, The Netherlands
[f] University Hospitals Leuven, Department of laboratory medicine, Leuven, Belgium
[g] University Hospitals Leuven, Department of radiology, Leuven, Belgium
[h] The Bone Infection Unit, Nuffield Orthopaedic Centre, Oxford University Hospitals, Oxford, United Kingdom
[i] Department of Trauma Surgery, University of Utrecht, University Medical Center Utrecht, Utrecht, The Netherlands

ARTICLE INFO

ABSTRACT

Background: The recently developed fracture-related infection (FRI) consensus definition, which is based on specific diagnostic criteria, has not been fully validated in clinical studies. We aimed to determine the diagnostic performance of the criteria of the FRI consensus definition and evaluated the effect of the combination of certain suggestive and confirmatory criteria on the diagnostic performance.

Methods: A multicenter, multi-national, retrospective cohort study was performed. Patients were subdivided into an FRI or a control group, according to the treatment they received and the recommendations from a multidisciplinary team ('intention to treat'). Exclusion criteria were patients with an FRI diagnosed outside the study period, patients younger than 18 years of age, patients with pathological fractures or patients with fractures of the skull, cervical, thoracic and lumbar spine. Minimum follow up for all patients was 18 months.

Results: Overall, 637 patients underwent revision surgery for suspicion of FRI. Of these, 480 patients were diagnosed with FRI, treated accordingly, and included in the FRI group. The other 157 patients were included in the control group. The presence of at least one confirmatory sign was associated with a sensitivity of 97.5%, a specificity of 100% and a high discriminatory value (AUROC 0.99, $p < 0.001$). The presence of a clinical confirmatory criterion or, if not present, at least one positive culture was associated with the highest diagnostic performance (sensitivity: 98.6%, specificity: 100%, AUROC: 0.99 ($p < 0.001$)). In the subgroup of patients without clinical confirmatory signs at presentation, specificities of at least 95% were found for the clinical suggestive signs of fever, wound drainage, local warmth and redness.

Conclusions: The presence of at least one confirmatory criterion identifies the vast majority of patients with an FRI and was associated with an excellent diagnostic discriminatory value. Therefore, our study validates the confirmatory criteria of the FRI consensus definition. Infection is highly likely in case of the presence of a single positive culture with a virulent pathogen. When certain clinical suggestive signs (e.g., wound drainage) are observed (individually or in combination and even without a confirmatory criterion), it is more likely than not, that an infection is present.

Introduction

Fracture-related infection (FRI) remains an important complication after musculoskeletal trauma, with an enormous impact on patients and healthcare systems [1,2]. An early and accurate di-

* Corresponding author at: Department of Trauma Surgery, University Hospitals Leuven, Belgium.
E-mail address: willem-jan.metsemakers@uzleuven.be (W.-J. Metsemakers).

agnosis is the first, vital step towards a successful treatment outcome. Similar to periprosthetic joint infection (PJI), FRI can present with a variety of clinical signs. However, there are also significant differences from PJI, such as the accompanying soft tissue or vascular injury, and the presence of a fracture. Furthermore, the diagnostic algorithm for PJI is different with certain criteria that are not available in FRI cases (e.g., synovial fluid). Although many studies and guidelines have been published on PJI, scientific evidence regarding the diagnosis of FRI has lagged behind [3,4]. Previous studies focusing on FRI either did not define diagnostic criteria or used self-designed definitions or the more generic Centers for Disease Control and Prevention (CDC) criteria for surgical site infection (SSI) [5]. Using an inadequate definition for FRI risks underestimating or overestimating the actual number of infections. Moreover, such poor diagnostics may make the interventions studied appear better than they really are, thereby resulting in misleading conclusions.

For the above-mentioned reasons, a consensus definition for FRI was created based on expert opinion, with the support of the AO (Arbeitsgemeinschaft für Osteosynthesefragen) Foundation and the European Bone and Joint Infection Society (EBJIS) [3]. An updated version of this definition was published more recently [6]. The FRI consensus definition utilizes two levels of certainty regarding the diagnosis of FRI in the form of confirmatory and suggestive criteria. FRI is only definitively confirmed, if one or more confirmatory criteria are present [3,6]. Although recent studies endorse the FRI consensus definition [7,8], its diagnostic criteria have not been fully validated. Because the clinical presentation of FRI is variable, insight in the diagnostic performance of these criteria is relevant.

We performed a multicenter, multi-national, retrospective cohort study to: (a) determine the diagnostic performance of the individual diagnostic criteria of the FRI consensus definition and (b) evaluate the effect of the combination of certain suggestive and/or confirmatory criteria, especially the ones available upon initial patient presentation, on the diagnostic performance of this definition.

Methods

Study design

This was a multicenter, multi-national, retrospective cohort study. Fracture patients who were suspected of having an FRI between January 2015 and November 2019 from the University Hospitals Leuven (Belgium), the University Medical Center Groningen (the Netherlands), the University Medical Center Utrecht (the Netherlands) and the Oxford University Hospitals (United Kingdom) were included.

Patients and study population

All patients who underwent revision surgery for suspicion of FRI within the study period were included in this study. Exclusion criteria were patients with an FRI diagnosed outside the study period, patients younger than 18 years of age, patients with pathological fractures or patients with fractures of the hand, skull, cervical, thoracic and lumbar spine. Pathological fractures due to malignancy were excluded as these may cause similar clinical and laboratory findings as infection, which may confound the study results. Fracture patients considered to have an infection and treated accordingly ('*intention to treat*'), based on best practice recommendations from a multidisciplinary team, were included in the FRI group. This way of patient allocation was chosen due to the lack of a gold standard to define FRI patients. A multidisciplinary team was present in all participating centers and consisted of surgeons, infectious disease specialists, microbiologists, radiologists and clinical pharmacists [4]. Patients who initially had a clinical suspi-

cion of FRI, but eventually were not diagnosed and treated as such (again based on best practice recommendations from the multidisciplinary team), were included in the control group. All patients were followed up for a minimum of 18 months. Recurrence of infection after cessation of (surgical and antimicrobial) treatment was defined in a similar fashion as the primary infection, using intention-to-treat, based on the recommendations of the multidisciplinary team [4].

Variables and outcome measures

Patient medical records were reviewed and patient demographics including age, sex, body mass index (BMI) and American Society of Anesthesiologists (ASA) score recorded. Data related to the fracture were collected including localization, Gustilo-Anderson (GA) type and time from primary fracture fixation to onset of symptoms. All confirmatory and suggestive diagnostic criteria of the FRI consensus definition (Table 1) were scored as present or absent by different reviewers (JO, JS, JF, FIJ, GG, MMcN, WJM). Serum inflammatory markers were considered elevated in case of a white blood cell (WBC) count $>10 \times 10^9$/L, a C-reactive protein level (CRP) >5 mg/L or an erythrocyte sedimentation rate (ESR) >20 mm/h. Single positive culture tests were recorded only when a virulent pathogen was isolated. Virulent pathogens were defined *a priori* as Gram-negative bacilli, *Staphylococcus aureus, Staphylococcus lugdunensis*, enterococci, beta-hemolytic streptococci, milleri group streptococci, *Streptococcus pneumonia* and *Candida* species [9]. The term 'virulent' refers to pathogens of high clinical importance (i.e., pathogens capable of producing disease and very often causing FRIs). The inclusion of pathogens in the virulent category was based on the clinical experience and consensus opinion of our infectious disease physicians, as these pathogens have a high likelihood of causing disease. Sampling methods used in this study followed standardized protocols [6,10-12]. Single positive cultures with non-virulent pathogens were not further evaluated as they were seen as contaminants.

Statistical analysis

Descriptive and univariate analysis was performed using SPSS for Windows, version 25 (SPSS, Chicago, Illinois, USA). Normality of continuous data was tested with the Shapiro-Wilk test, which showed that all continuous data were nonparametric. Continuous variables were compared using the Mann-Whitney U-test. Categorical variables were compared using the Fisher's exact test or Chi-square test, as appropriate. P-values less than 0.05 (2-sided test) were considered statistically significant. Diagnostic properties and discrimination of the consensus criteria were calculated using MedCalc Statistical Software version 18.2.1. (MedCalc Software bvba, Ostend, Belgium; http://www.medcalc.org; 2018). Sensitivity (or true positive rate, which is the ability of a test to correctly identify patients with FRI), specificity (or true negative rate, which is the ability of a test to correctly identify people without FRI), and area under the receiver operating curve (AUROC) are reported with 95% confidence interval (95% CI).

Selection of suggestive diagnostic criteria for secondary analysis

To investigate the diagnostic performance (i.e., sensitivity and specificity) of sets of suggestive criteria in the entire study population, in the subgroup of patients who did not present with clinical confirmatory criteria or in the subgroup of patients who presented with phenotypically indistinguishable pathogens isolated from at least two separate deep tissue specimens (microbiological confirmatory criterion), suggestive diagnostic criteria were selected

Table 1
Diagnostic criteria based on the consensus definition for fracture-related infection.

Confirmatory criteria
Clinical signs
Fistula, sinus or wound breakdown (with communication with the bone or implant)
Purulent drainage from the wound or presence of pus during surgery
Microbiology signs
Phenotypically indistinguishable microorganisms isolated from at least two separate deep tissue or implant specimens
Histopathological signs
Presence of microorganisms in deep tissue specimens, confirmed by histopathology
For chronic or late-onset cases: the presence of at least five polymorphonuclear neutrophils per high-power field

Suggestive criteria
Local clinical signs of inflammation
Local redness
Pain (without weight bearing, increasing over time, new-onset)
Local swelling
Local warmth
Systemic clinical signs
Fever (single oral temperature measurement of $\geq$38.3°C (101°F)
Other clinical signs
New-onset joint effusion
Persistent, increasing or new-onset wound drainage, beyond the first few days postoperatively, without solid alternative explanation
Radiological (conventional radiography, CT, MRI) signs
Nuclear imaging (WBC scan, ^{18}F FDG-PET) signs
Laboratory signs
Elevated serum inflammatory markers (WBC, CRP and/or ESR)
Microbiology signs
Pathogenic organism identified by culture from a single deep tissue or implant specimen

WBC: white blood cell count; CRP: C-reactive protein; ESR: erythrocyte sedimentation rate.

based on the following conditions: [1] the test had to be available for the majority of patients ($>$50%) in both the FRI and control group; [2] the test should be associated with a high specificity (defined as $\geq$80%) and a significant discriminatory value (AUROC $>$50%, $p < 0.05$).

Ethics approval

The study protocol was conducted following good clinical practice guidelines and was approved by the Ethics Committee of the University Hospitals Leuven, Belgium (Ethics Committee Research UZ/KU Leuven; S62394). A data sharing agreement was signed between participating centers.

Results

Patient demographics, fracture characteristics and outcome

During the study period, 637 patients underwent revision surgery for suspicion of FRI. Of these, 480 patients were diagnosed with FRI, treated accordingly, and included in the FRI group. Four of these patients were diagnosed with two FRIs, at different anatomical localizations and at different moments in time. Therefore, the FRI group consisted of 480 patients with 484 FRIs. The other 157 patients were included in the control group (Fig. 1). Patient characteristics are presented in Table 2. Patients in both groups were similar in age, BMI and ASA score distribution. The FRI group consisted of more men ($n = 329$, 68.5%) compared to the control group ($n = 72$, 45.9%) ($p < 0.001$).

The distribution of fracture locations differed between both groups ($p < 0.001$). In the FRI group, most fractures involved the tibia and/or fibula, followed by the femur and humerus. In the control group, most fractures involved the femur, followed by the tibia and/or fibula, and the humerus. The percentage of open fractures and their severity (GA type) was similar in both groups (29% vs 22%; $p = 0.100$), as was the percentage of unhealed fractures (82% vs 81%; $p = 0.810$). The median time between primary fracture fixation and onset of symptoms was significantly longer in the control

group (284 days, P_{25} - P_{75} 132-447) compared to the FRI group (42 days, P_{25} - P_{75} 15 - 191) ($p < 0.001$).

As mentioned earlier, all patients were followed up for a minimum of 18 months with evaluation of the outcome. During the follow-up period, none of the patients in the control group developed an infection (0/157). The overall recurrence rate in the FRI group was 11.6% (56/480) ($p < 0.001$).

Diagnostic performance of individual criteria

The prevalence and diagnostic performance for each individual diagnostic criterion in the entire study population is shown in Table 3.

Confirmatory criteria (Table 3)

Confirmatory signs were only present in the FRI group (Fig. 1), which corresponds to a specificity of 100% for each separate confirmatory criterion. A fistula, sinus, or wound breakdown was present in 241 (49.8%) FRI patients. Purulent drainage from the wound or pus during surgery was present in 237 (49.0%) patients. In 426 (88.0%) FRI patients, infection was confirmed by phenotypically indistinguishable microorganisms isolated from at least two separate deep tissue cultures. Negative cultures were found in 41 (8.5%) FRI patients and in all (100%) control patients. In the FRI group, 97 patients (20.0%) were treated with antibiotics within 14 days prior to tissue sampling, compared to none of the patients in the control group ($p < 0.001$) (Table 2). The histopathological presence of visible pathogens using different staining techniques for bacteria and fungi ($n = 90$) was found in 13 (14.4%) FRI patients. In the FRI group for whom a polymorphonuclear neutrophil (PMN) count (histopathology) was performed ($n = 90$), at least 5 PMNs per high-power field (PMNs/HPF) were present in 56 (62.2%) patients. Because histopathology was only performed in a small number of patients in both the FRI ($n = 90$, 18.6%) and control ($n = 24$, 15.3%) group, this diagnostic modality was excluded from the secondary analyses. Any confirmatory sign (excluding histopathology) was present in 97.5% of the FRI cases and was associated with a
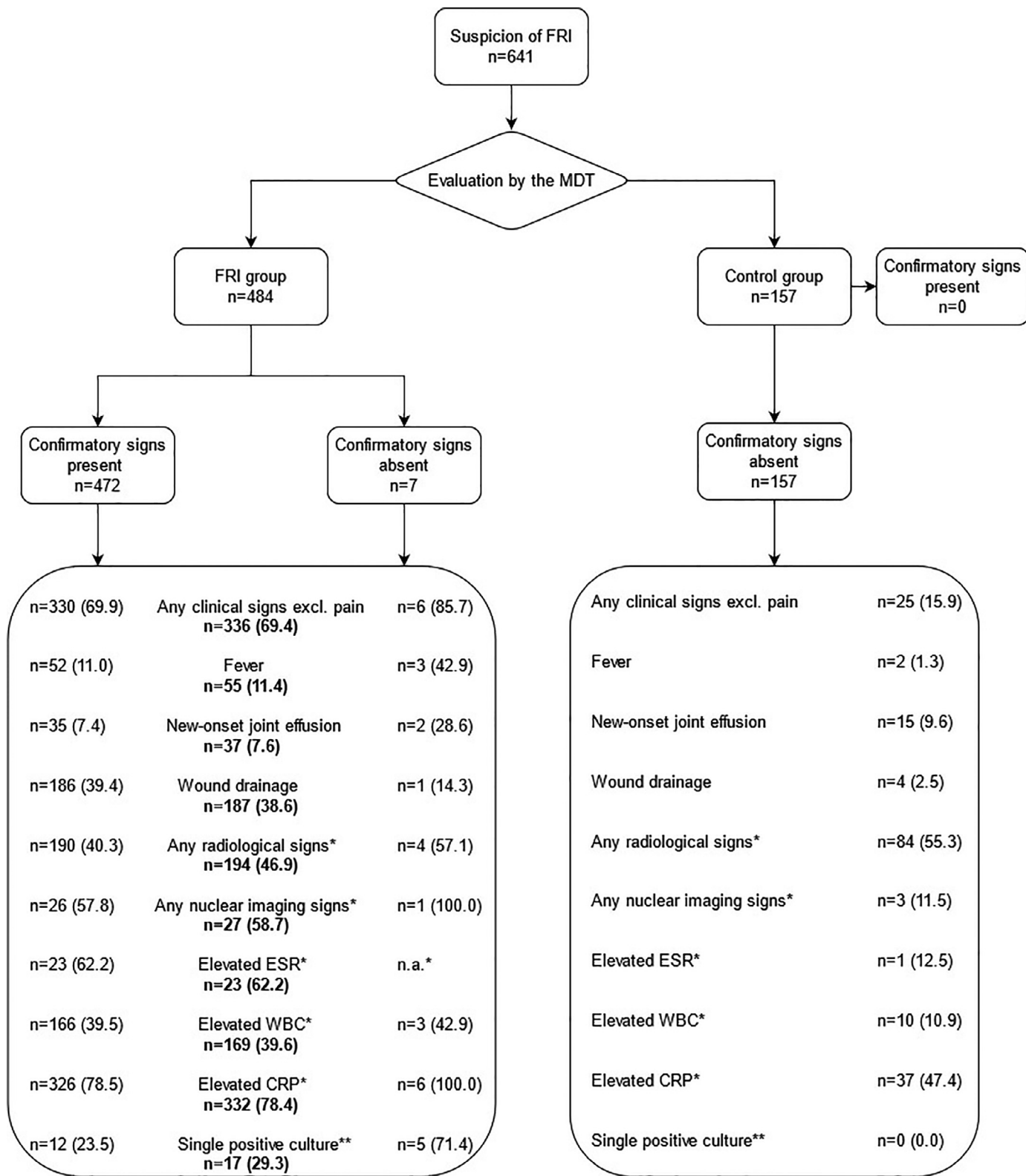
Fig 1. Study flowchart describing the prevalence of confirmatory and suggestive signs in the control group and within the FRI group. Percentages are shown between brackets. Frequencies and percentages in bold are calculated using the FRI group as the denominator, regardless of the presence or absence of confirmatory signs. MDT: multidisciplinary team; FRI: fracture-related infection; ESR: erythrocyte sedimentation rate; WBC: white blood cell count; CRP: C-reactive protein.
*These tests were not performed in all patients, the reported percentages are relative to the total number of patients for whom the test was available.
**Single positive culture caused by virulent pathogens. Virulent pathogens were defined *a priori* as Gram-negative bacilli, *Staphylococcus aureus*, *Staphylococcus lugdunensis*, enterococci, beta-hemolytic streptococci, milleri group streptococci, *Streptococcus pneumonia* and *Candida* species.

Table 2
Patient and fracture characteristics for patients with (FRI group) versus without FRI (control group).

Parameter		Overall(N = 637) N*		FRI group(N = 480) N*		Control group (N = 157) N*	p-value
Age (years)	637	52 (39-65)	480	52 (39-64)	157	53 (40-66)	0.570
Male sex	637	401 (63.0)	480	329 (68.5)	157	72 (45.9)	**<0.001**
BMI (kg/m^2)	632	26.3 (23.2-29.9)	475	26.4 (23.4-30.0)	157	25.6 (22.9-29.6)	0.249
ASA score 1	637	170 (26.7)	480	123 (25.6)	157	47 (29.9)	0.261
2		327 (51.3)		250 (52.1)		77 (49.0)	
3		131 (20.6)		98 (20.4)		33 (21.0)	
4		9 (1.4)		9 (1.9)		0 (0.0)	
Fracture localization[1] Humerus	641	75 (11.7)	484	38 (7.9)	157	37 (23.6)	**<0.001**
Clavicle		23 (3.6)		18 (3.7)		5 (3.2)	
Radius or ulna		49 (7.6)		36 (7.4)		13 (8.3)	
Femur		141 (22.0)		87 (18.0)		54 (34.4)	
Tibia and/or fibula		297 (46.3)		252 (52.1)		45 (28.7)	
Foot		21 (3.3)		19 (3.9)		2 (1.3)	
Pelvic ring or acetabulum		25 (3.9)		24 (5.0)		1 (0.6)	
Patella		5 (0.8)		5 (0.4)		0 (0.0)	
Rib or sternum		4 (0.6)		4 (0.8)		0 (0.0)	
Scapula		1 (0.2)		1 (0.2)		0 (0.0)	
Open fracture[1]	641	177 (27.6)	484	142 (29.3)	157	35 (22.3)	0.100
GA type I	177	41 (23.2)	142	29 (20.4)	35	12 (34.3)	0.104
II		58 (32.8)		51 (35.9)		7 (20.0)	
III		78 (44.1)		62 (43.7)		16 (45.7)	
Healing status[1] Healed	641	113 (17.6)	484	84 (17.4)	157	29 (18.5)	0.810
Unhealed		528 (82.4)		400 (82.6)		128 (81.5)	
Time from primary fixation until onset of symptoms (days)[1]	641	73 (18-316)	484	42 (15-191)	157	284 (132-447)	**<0.001**
Antibiotic therapy less than two weeks prior to revision surgery (tissue sampling)	641	97 (15.1)	484	97 (20.0)	157	0 (0.0)	**<0.001**

Data are shown as median (P$_{25}$ - P$_{75}$) or as N (%). N*, number of patients or FRIs for whom data were available. FRI: fracture-related infection; BMI: body mass index; ASA: American Society for Anaesthesiologists; GA: Gustilo-Anderson; FDA: Food and Drug Administration.
[1] The number of fractures adds up to a total of 641 because four patients suffered from two FRIs at different anatomical localizations.
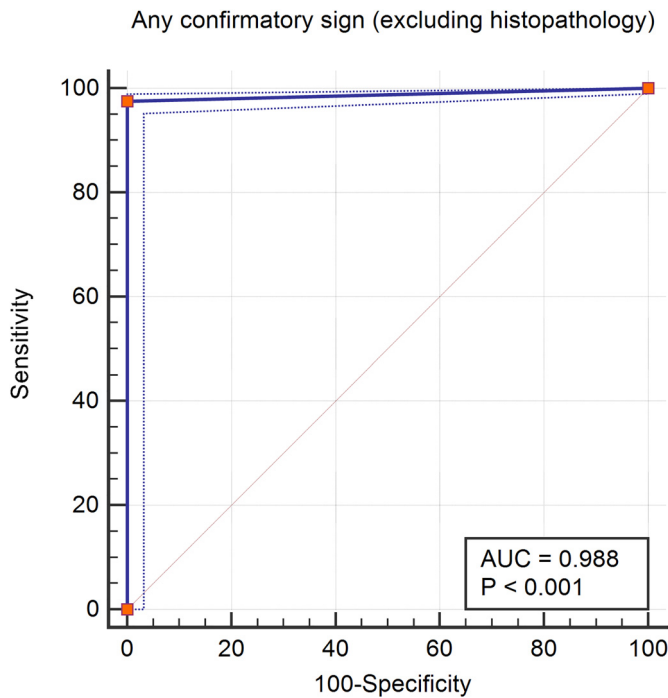

Fig 2. Receiver operating curve for the presence of any confirmatory sign excluding histopathology. The AUC (AUROC) for this parameter is 0.99 (*p* < 0.001). AUC: area under the curve.

specificity of 100% and a high discriminatory value (AUROC 0.99, *p* < 0.001) (Fig. 2).

Twelve FRI patients did not have any clinical or microbiological confirmatory criteria. In five of these 12 patients, histopathological analysis showed the presence of at least 5 PMNs/HPF. The diagnostic profile of the remaining seven patients who did not have any confirmatory criterion is summarized in Table 4. Five of these

patients had a single positive culture with a virulent pathogen (cultures were taken under antibiotic treatment in three of them). These five patients also all had clinical suggestive signs. Furthermore, one of these patients had a positive FDG-PET scan. The remaining two patients had negative cultures, which were taken under antibiotic therapy. These two patients both had clinical suggestive signs (i.e., fever, redness, swelling, local warmth) and an elevated CRP level. For the above-mentioned reasons, the multidisciplinary team considered these seven patients as infected and treated them as such.

Clinical signs

Regarding local clinical suggestive signs, pain was more prevalent in the control (*n* = 107, 68.2%) than in the FRI group (*n* = 233, 48.1%) (*p* < 0.001). The sensitivity and specificity of pain were 48.1% and 31.8%, respectively. The presence of any local clinical sign of inflammation (i.e., local redness, swelling, or warmth), excluding pain, was associated with a sensitivity of 69.4%, a specificity of 84.1% and an AUROC of 0.77 (*p* < 0.001). Fever was more often present in the FRI group (11.4% vs 1.3%, *p* < 0.001) and had a sensitivity of 11.4%, a specificity of 98.7% and an AUROC of 0.55 (*p* < 0.001). Persisting, increasing or new-onset wound drainage had a sensitivity of 38.6%, a specificity of 97.5% and an AUROC of 0.68 (*p* < 0.001). New-onset joint effusion had a sensitivity of 7.6% and a specificity of 90.4%.

Radiological and nuclear imaging signs

Radiological signs [3,4] – as evaluated by conventional radiography (x-ray) in 390 (80.6%) FRI and 138 (87.9%) control patients – were more prevalent in the control (*n* = 80, 58.0%) than in the FRI group (*n* = 171, 43.8%) (*p* = 0.005). The presence of any radiological sign (on conventional radiography, CT or MRI) had no significant discriminatory value to diagnose FRI (AUROC 0.54, *p* = 0.076). Regarding nuclear imaging, the presence of any nuclear imaging sign was associated with a sensitivity of 58.7%, a specificity of 88.5% and an AUROC of 0.74 (*p* < 0.001). Furthermore, the ^{18}F-

Table 3
Prevalence, diagnostic performance and area under the receiver operating curve of individual confirmatory and suggestive signs.

Sign present	Prevalence					Diagnostic performance		Discrimination	
	FRI (N = 484) N*		Control (N = 157) N*		p-value	N*	Sensitivity (95% CI)	Specificity (95% CI)	AUROC (95% CI)	p-value
CONFIRMATORY CRITERIA										
Clinical signs										
Fistula, sinus, or wound breakdown	484	241 (49.8)	157	0 (0.0)	**<0.001**	641	49.8 (45.2-54.3)	100.0 (97.7-100.0)	0.75 (0.73-0.77)	**<0.001**
Purulent drainage or pus	484	237 (49.0)	157	0 (0.0)	**<0.001**	641	49.0 (44.4-53.5)	100.0 (97.7-100.0)	0.75 (0.72-0.77)	**<0.001**
Microbiological signs										
Phenotypically indistinguishable microorganisms isolated from at least two separate deep cultures	484	426 (88.0)	157	0 (0.0)	**<0.001**	641	88.0 (84.8-90.8)	100.0 (97.7-100.0)	0.94 (0.93-0.96)	**<0.001**
Histopathological signs										
Histological presence of microorganisms	90	13 (14.4)	24	0 (0.0)	0.067	114	14.4 (7.9-23.4)	100.0 (85.8-100.0)	0.57 (0.54-0.61)	**<0.001**
Histological presence of ≥ 5 PMNs/HPF	90	56 (62.2)	24	0 (0.0)	**<0.001**	114	62.2 (51.4-72.2)	100.0 (85.8-100.0)	0.81 (0.76-0.86)	**<0.001**
Any confirmatory sign excl. histopathology[1]	484	472 (97.5)	157	0 (0.0)	**<0.001**	641	97.5 (95.7-98.7)	100.0 (97.7-100.0)	0.99 (0.98-1.00)	**<0.001**
Any confirmatory sign incl. histopathology[1]	90	88 (97.8)	24	0 (0.0)	**<0.001**	114	97.8 (92.2-99.7)	100.0 (85.8-100.0)	0.99 (0.97-1.00)	**<0.001**
SUGGESTIVE CRITERIA										
Local clinical signs of inflammation										
Redness/rubor	484	274 (56.6)	157	8 (5.1)	**<0.001**	641	56.6 (52.1-61.1)	94.9 (90.2-97.8)	0.76 (0.73-0.79)	**<0.001**
Local warmth/calor	484	109 (22.5)	157	4 (2.5)	**<0.001**	641	22.5 (18.9-26.5)	97.5 (93.6-99.3)	0.60 (0.58-0.62)	**<0.001**
Swelling/tumor	484	227 (46.9)	157	18 (11.5)	**<0.001**	641	46.9 (42.4-51.5)	88.5 (82.5-93.1)	0.68 (0.64-0.71)	**<0.001**
Pain/dolor	484	233 (48.1)	157	107 (68.2)	**<0.001**	641	48.1 (46.6-52.7)	31.8 (24.6-39.7)	0.60 (0.56-0.64)	**<0.001**
Any local clinical sign of inflammation[1]	484	388 (80.2)	157	115 (73.2)	**<0.001**	641	80.2 (76.3-83.6)	26.8(20.0-34.4)	0.54 (0.50-0.57)	0.082
Any local clinical sign of inflammation (excl. pain)[1]	484	336 (69.4)	157	25 (15.9)	**<0.001**	641	69.4 (65.1-73.5)	84.1 (77.4-89.4)	0.77 (0.73-0.80)	**<0.001**
Systemic clinical signs										
Fever ≥ 38.3°C	484	55 (11.4)	157	2 (1.3)	**<0.001**	641	11.4 (8.7-14.5)	98.7 (95.5-99.8)	0.55 (0.53-0.57)	**<0.001**
Other clinical signs										
New-onset joint effusion	484	37 (7.6)	157	15 (9.6)	0.501	641	7.6 (5.4-10.4%)	90.4 (84.7-94.6)	0.51 (0.48-0.54)	0.471
Wound drainage	484	187 (38.6)	157	4 (2.5)	**<0.001**	641	38.6 (34.3-43.1)	97.5 (93.6-99.3)	0.68 (0.66-0.71)	**<0.001**
Radiological signs										
Any radiological sign[1]	414	194 (46.9)	152	84 (55.3)	0.088	566	46.9 (42.0-51.8)	44.7 (36.7-53.0)	0.54 (0.50-0.59)	0.076
Conventional radiography	390	171 (43.8)	138	80 (58.0)	**0.005**	528	43.8 (38.9-48.9)	42.0 (33.7-50.7)	0.57 (0.52-0.62)	**0.004**
CT-scan	156	75 (48.1)	91	49 (53.8)	0.429	247	48.1 (40.0-56.2)	46.2 (35.6-56.9)	0.53 (0.46-0.59)	0.383
MRI-scan	7	4 (57.1)	1	0 (0.0)	1.000	8	57.1 (18.4-90.1)	100.0 (2.5-100.0)	N.A.	N.A.
Nuclear imaging signs										
Any nuclear imaging sign[1]	46	27 (58.7)	26	3 (11.5)	**<0.001**	72	58.7 (43.2-73.0)	88.5 (69.8-97.6)	0.74 (0.64-0.83)	**<0.001**
WBC scan	24	12 (50.0)	20	3 (15.0)	**0.025**	44	50.0 (29.1-70.9)	85.0 (62.1-96.8)	0.68 (0.55-0.81)	**0.008**
^{18}F-FDG-PET scan	23	15 (65.2)	8	0 (0.0)	**0.002**	31	65.2 (42.7-83.6)	100.0 (63.1-100.0)	0.83 (0.73-0.93)	**<0.001**
Laboratory signs										
ESR > 20 mm/h	37	23 (62.2)	8	1 (12.5)	**0.017**	45	62.2 (44.8-77.5)	87.5 (47.3-99.7)	0.75 (0.60-0.89)	**0.001**
WBC count > 10 × 10⁹/L	427	169 (39.6)	92	10 (10.9)	**<0.001**	519	39.6 (34.9-44.4)	89.1 (80.9-94.7)	0.64 (0.60-0.68)	**<0.001**
CRP level > 5 mg/L	425	332 (78.4)	78	37 (47.4)	**<0.001**	503	78.4 (74.1-82.2)	52.6 (40.9-64.0)	0.66 (0.60-0.71)	**<0.001**
Microbiological signs										
Single positive culture[2]	58	17 (29.3)	157	0 (0.0)	**<0.001**	215	29.3 (18.1-42.7)	100.0 (97.7-100.0)	0.65 (0.59-0.71)	**<0.001**

Prevalence data are shown as N (%) for patients with versus without FRI. Significance of difference is tested using the Chi-square or Fisher's exact test, as applicable. Diagnostic performance parameters sensitivity and specificity are shown as percentages with 95% CI. The area under the receiver operating curve (AUROC) is shown with 95% CI and associated p-value. N*, number of patients for whom data are available. FRI: fracture-related infection; PMN: polymorphonuclear neutrophil; HPF: high-power field; CT: computed tomography; MRI: magnetic resonance imaging; N.A.: not applicable; WBC: white blood cell; FDG-PET: fluorodeoxyglucose positron emission tomography; ESR: erythrocyte sedimentation rate; CRP: C-reactive protein.

[1] The criterion was scored as 'present' if any one of the mentioned signs was present, while it was scored as 'absent' if none of the mentioned signs were present.

[2] The percentage and diagnostic performance of a single positive culture was calculated relative to the patients who had no confirmatory microbiological signs (i.e., either single positive cultures or negative culture results).

Table 4
Characteristics of seven FRI patients who did not present with any confirmatory signs.

Patient demographics (n = 7)	Prevalence
Age	63 (52.0 – 74.0)
Male sex	6 (85.7)
BMI (kg/m²)	29.7 (27.1 – 33.1)
ASA score	
1	3 (42.9)
2	3 (42.9)
3	1 (14.2)
4	0 (0.0)
Fracture localization	
Humerus	1 (14.2)
Femur	2 (28.7)
Tibia	3 (42.9)
Tibia and fibula	1 (14.2)
Open fracture	1 (14.2)
GA type I	0 (0.0)
II	0 (0.0)
III	1 (100.0)
Healing status	
Healed	1 (14.2)
Unhealed	6 (85.8)
Time from primary fixation until onset of symptoms	398.4 (29.0 – 882.0)
Suggestive signs	
Local clinical signs (n = 7)	
Any local clinical sign	7 (100.0)
Any local clinical sign (excl. pain)	6 (85.7)
Redness	4 (57.1)
Pain	5 (71.4)
Swelling	5 (71.4)
Local warmth	2 (28.6)
Systemic clinical signs (n = 7)	
Fever ≥ 38.3°C	3 (42.9)
Other clinical signs (n = 7)	
New-onset joint effusion	2 (28.6)
Wound drainage	1 (14.2)
Radiological signs (n = 7)	
Any radiological sign	4 (57.1)
Standard radiography(n = 7)	3 (42.9)
CT scan (n = 3)	3 (100.0)
Nuclear imaging signs	
^{18}F-FDG-PET (n = 1)	1 (100)
Laboratory signs	
WBC count > 10 × 10⁹/L (n = 7)	3 (42.9)
CRP level > 5 mg/L (n = 6)	6 (100.0)
Microbiological signs (n = 7)	
Single positive culture[2,3]	5 (71.4)
Negative cultures[4]	2 (28.6)
Combination of clinical and microbiological signs (n = 7)	
Any clinical sign (redness/swelling/warmth or fever or wound drainage) and if not present, a single positive culture[1,2]	7 (100.0)

Prevalence is shown as N (%) or as median (P$_{25}$ – P$_{75}$) for patients with FRI who did not have any confirmatory signs (excluding histopathology). In five out of these patients (41.7%), histopathological analysis showed the presence of ≥ 5 PMNs/HPF. FRI: fracture-related infection; BMI: body mass index; ASA: American Society for Anaesthesiologists; GA: Gustilo-Anderson; CT: computed tomography; MRI: magnetic resonance imaging; N.A.: not applicable; WBC: white blood cell; FDG-PET: fluorodeoxyglucose positron emission tomography; ESR: erythrocyte sedimentation rate; CRP: C-reactive protein.

[1] The criterion was scored as 'present' if any one of the mentioned signs was present, while it was scored as 'absent' if none of the mentioned signs were present.

[2] Single positive culture caused by virulent pathogens. Virulent pathogens were defined *a priori* as Gram-negative bacilli, *Staphylococcus aureus, Staphylococcus lugdunensis*, enterococci, beta-hemolytic streptococci, milleri group streptococci, *Streptococcus pneumonia* and *Candida* species. The isolated pathogens in this group consisted of *S. aureus, S. lugdunensis, Burkholderia cepacia* and *Citrobacter koseri*.

[3] In three patients, cultures were taken during antibiotic treatment or antibiotic treatment was given within two weeks prior to sampling.

[4] In both patients, cultures were taken during antibiotic treatment or antibiotic treatment was given within two weeks prior to sampling.

FDG-PET scan was associated with a specificity of 100% and an AUROC of 0.83 ($p < 0.001$). However, nuclear imaging (WBC scan or ^{18}F-FDG-PET scan) was performed in a low sample size of 46 FRI and 26 control patients.

Laboratory signs

WBC count and CRP level were elevated more often in the FRI group compared to the control group ($p < 0.001$). WBC count was associated with a sensitivity of 39.6%, a specificity of 89.1% and an AUROC of 0.64 ($p < 0.001$). CRP was associated with a sensitivity of 78.3%, a specificity of 52.6% and an AUROC of 0.65 ($p < 0.001$).

The sensitivity and specificity of ESR was 62.2% and 87.5%, respectively, with an AUROC of 0.75 ($p < 0.001$). However, ESR was only determined in 37 (7.6%) FRI patients and eight (5.1%) control patients.

Microbiology signs

The diagnostic performance of a single positive culture with a virulent pathogen was determined in the FRI subgroup that did not fulfill the confirmatory criterion of phenotypically indistinguishable microorganisms isolated from at least two separate deep tissue cultures ($n = 58$). Overall, in 17 of these patients a single

positive culture with a virulent pathogen was retrieved. Four were under antibiotic therapy at time of surgery. In the group with a single positive culture, the recurrence rate was 17.6% ($n = 3$). The sensitivity of a single positive culture with a virulent pathogen in the selected population was 29.3% and the AUROC was 0.65 ($p < 0.001$). In the control group, all cultures were negative, which corresponds to a specificity of 100%. As mentioned above, negative cultures were found in 41 FRI patients. Twenty-one were under antibiotic therapy at time of surgery. The recurrence rate in the culture-negative group was 17.1% ($n = 7$).

Diagnostic performance of a selection of confirmatory and suggestive criteria

To simulate our daily clinical practice, where a patient often presents for the first time with a suspicion of FRI, a secondary analysis was performed using combinations of signs. This analysis was performed to assess the performance of the presence of (a) clinical confirmatory criteria, (b) clinical confirmatory signs or, if not present, of select suggestive signs in the entire study population, (c) select suggestive signs in the subgroup of patients who did not present with any clinical confirmatory signs, and (d) select suggestive signs in the subgroup of patients who presented with phenotypically indistinguishable pathogens isolated from at least two separate deep tissue specimens. Herewith we not only evaluate if certain combinations of signs improve the diagnostic performance of the definition but also try to simulate the first patient contact at the outpatient clinic or emergency department.

Diagnostic performance of clinical confirmatory criteria (Table 5)

The presence of clinical confirmatory signs that are evident at presentation, i.e., purulent drainage or the presence of a fistula/sinus/wound breakdown, was determined in the FRI and control groups as a first step. Any one of these clinical confirmatory signs were present in 359 (74.2%) FRI patients and in none of the control patients, corresponding to a sensitivity of 74.2%, a specificity of 100%, and an AUROC of 0.87 ($p < 0.001$) (Table 5).

Diagnostic performance of selected suggestive clinical signs in the whole study population (Table 5)

We evaluated the presence of selected suggestive signs (see Methods sections) (Table 5) in the entire patient population. The following signs were included in the analysis: redness/swelling/local warmth, fever, wound drainage, WBC count elevation, and a single positive culture caused by a virulent pathogen. All individual diagnostic alternatives to the clinical confirmatory signs had an AUROC exceeding 0.80 ($p < 0.001$), regardless of the suggestive clinical criterion that was evaluated. The presence (a) of a clinical confirmatory criterion or (b) if not present, of any of the selected clinical suggestive signs corresponded to a sensitivity of 91.3%, a specificity of 80.9% and an AUROC of 0.86 ($p < 0.001$). This proves the importance of any clinical (inflammatory) sign (excluding pain) in the diagnostic pathway for FRI. When this analysis was repeated with an elevated WBC count as an additional suggestive criterion (i.e., the presence (a) of a clinical confirmatory criterion or (b) if not present, of any of the selected clinical suggestive signs with an elevated WBC count), sensitivity increased to 95.5%, but specificity decreased to 58.5%. The AUROC for this set of criteria was 0.77 ($p < 0.001$). The presence of (a) a clinical confirmatory criterion or (b) if not present, at least one positive culture was associated with a sensitivity of 98.6%, a specificity of 100% and an AUROC of 0.99 ($p < 0.001$).

Diagnostic performance of selected suggestive signs in the subgroup of patients who did not present with any clinical confirmatory signs (Table 6)

Table 5 includes the whole patient population. In order to gain additional insight into the added value of clinical suggestive signs in the absence of clinical confirmatory signs, a subgroup analysis was done. The 359 patients in the FRI group who presented with any of the clinical confirmatory signs were excluded. This subgroup analysis, as detailed in Table 6, therefore only included patients who did not present with clinical confirmatory signs ($n = 125$, 26% for the FRI group, $n = 157$, 100% for the control group). The presence of any clinical suggestive sign was associated with a sensitivity of 66.4%, a specificity of 80.9% and an AUROC of 0.74 ($p < 0.001$). When WBC count elevation was added to this analysis, sensitivity increased to 80.0% but specificity decreased to 65.0% with an AUROC of 0.73 ($p < 0.001$). Finally, the presence of wound drainage as an individual factor seems to be important as well with a sensitivity of 28.8%, a specificity of 97.5% and an AUROC of 0.63 ($p < 0.001$).

Diagnostic performance of selected suggestive signs in the subgroup of patients who presented with phenotypically indistinguishable microorganisms isolated from at least two separate deep tissue specimens (Table 7)

To further support our findings, a subanalysis (Table 7) was performed in which we used the presence of the microbiological confirmatory criterion, which was recently validated [10], as the reference standard. A similar diagnostic performance for all selected suggestive signs was found for this subpopulation of FRI cases ($n = 426$), compared to the entire study population of FRI cases using intention to treat as a reference standard ($n = 484$).

Discussion

This multicenter, multi-national study validated the diagnostic criteria of the FRI consensus definition in a population of patients with suspected FRI who underwent revision surgery. Our results show that the presence of any confirmatory criterion had a high diagnostic discriminatory value and a sensitivity and specificity of 97.5% and 100%, respectively. Below we will discuss the results based on the two main study aims.

1) Determine the diagnostic performance of the individual confirmatory and suggestive criteria of the FRI consensus definition

- **Confirmatory criteria**

Confirmatory criteria, which are considered as pathognomonic for FRI [3], were only present in the FRI group. We demonstrated that the presence of any clinical or microbiological confirmatory criterion has a sensitivity of 97.5% and 100%, respectively. Therefore, our study validates the confirmatory criteria of the FRI consensus definition.

The microbiological confirmatory criterion (phenotypically indistinguishable pathogens isolated from at least two deep tissue specimens) had a high sensitivity of 88.0% and a specificity of 100%. This diagnostic performance is higher than reported in a recent study that aimed to validate the microbiological criteria for the diagnosis of FRI [10]. In that study, a sensitivity and specificity of 68% and 87%, respectively, was reported when at least two out of five cultures were positive with the same pathogen [10]. The authors used only the clinical confirmatory criteria and histopathology to allocate patients either to an infected or a control group. Microbiological criteria were omitted to avoid incorporation bias, which could result in an underestimation of their diagnostic performance [10]. In nonunions thought to be aseptic before surgery,

Table 5

Subanalyses using sets of confirmatory or suggestive signs in the entire study population.

	FRI (N = 484) N*	Control (N = 157) N*	p-value	N*	Sensitivity (95% CI)	Specificity (95% CI)	AUROC (95% CI)	p-value
Evaluation of clinical confirmatory signs								
Purulent drainage/pus or fistula/sinus/wound breakdown[1]	484 359 (74.2)	157 0 (0.0)	**<0.001**	641	74.2 (70.0-78.0)	100.0 (97.7-100.0)	0.87 (0.85-0.89)	**<0.001**
Evaluation of clinical confirmatory signs with suggestive clinical and/or laboratory signs								
Purulent drainage/pus or fistula/sinus/wound breakdown, and if not present, redness/swelling/warmth[1]	484 425 (87.8)	157 25 (15.9)	**<0.001**	641	87.8 (84.6-90.6)	84.1 (77.4-89.4)	0.86 (0.83-0.89)	**<0.001**
Purulent drainage/pus or fistula/sinus/wound breakdown, and if not present, fever[1]	484 371 (76.7)	157 2 (1.3)	**<0.001**	641	76.7 (72.6-80.4)	98.7 (95.5-99.8)	0.88 (0.86-0.90)	**<0.001**
Purulent drainage/pus or fistula/sinus/wound breakdown, and if not present, wound drainage[1]	484 395 (81.6)	157 4 (2.5)	**<0.001**	641	81.6 (77.9-85.0)	97.5 (93.6-99.3)	0.90 (0.87-0.92)	**<0.001**
Purulent drainage/pus or fistula/sinus/wound breakdown, and if not present, any clinical sign (redness/swelling/warmth or fever or wound drainage)[1]	484 442 (91.3)	157 30 (19.1)	**<0.001**	641	91.3 (88.5-93.7)	80.9 (73.9-86.7)	0.86 (0.83-0.89)	**<0.001**
Purulent drainage/pus or fistula/sinus/wound breakdown, and if not present, any clinical sign (redness/swelling/warmth or fever or wound drainage) or WBC count elevation[1,2]	469 448 (95.5)	106 44 (41.5)	**<0.001**	575	95.5 (93.2-97.2)	58.5 (48.5-68.0)	0.77 (0.72-0.82)	**<0.001**
Evaluation of clinical confirmatory signs with microbiological signs (confirmatory and suggestive)								
Purulent drainage/pus or fistula/sinus/wound breakdown and if not present, at least one positive culture[1,3]	484 477 (98.6)	157 0 (0.0)	**<0.001**	641	98.6 (97.0-99.4)	100.0 (97.7-100.0)	0.99 (0.99-1.00)	**<0.001**

Prevalence data are shown as N (%) for patients with versus without FRI. Significance of difference is tested using the Chi-square or Fisher's exact test, as applicable. Diagnostic performance parameters sensitivity and specificity are shown as percentages with 95% CI. The area under the receiver operating curve (AUROC) is shown with 95% CI and associated p-value. FRI: fracture-related infection; WBC: white blood cell.

[1] The criterion was scored as 'present' if any of the mentioned signs was present, while it was scored as 'absent' if none of the mentioned signs were present.

[2] Because WBC count was not performed in all patients, the percentage and diagnostic performance was calculated relative to the patients for whom WBC count was available.

[3] A single positive culture was only included in this study when a virulent pathogen was isolated. Virulent pathogens were defined a priori as Gram-negative bacilli, *Staphylococcus aureus, Staphylococcus lugdunensis*, enterococci, beta-hemolytic streptococci, milleri group streptococci, *Streptococcus pneumonia* and *Candida* species.

positive culture rates of up to 40% have been described [13,14]. However, in our study cohort, none of the patients in the control group had positive cultures, corresponding to a specificity of 100%. This can be explained by the fact that patient inclusion in the FRI or control group was done after surgery, taking into account the culture results. Furthermore, of the 400 fractures in the FRI group that were unhealed, 35 (8.8%) were culture negative. Seven of these patients presented without confirmatory signs and are described in Table 4. Sixteen patients (4.0%) had a single positive culture caused by a virulent pathogen, while the vast majority of patients with unhealed fractures (349 patients, 87.3%) had at least two deep tissue cultures with phenotypically indistinguishable pathogens. This, compared to no double or single positive culture results in the control group, and, considering the fact that all patients were followed up for at least 18 months, indicates that our results are reliable. Although there is no 'gold standard' to diagnose FRI, we believe that our approach using the multidisciplinary team recommendation was optimal because this recommendation was based on all available data for each patient. Finally, in our pa-

tient cohort the specificity of both histopathological criteria was 100%. Although histopathology was only performed in a relatively small subset of patients (and only in chronic cases) and therefore not further evaluated, these results are consistent with recent literature which highlights the importance of histopathology in the diagnostic pathway of orthopedic device-related infections (i.e., PJI, FRI) [15–19].

- **Suggestive criteria**

Clinical signs

To our knowledge, this is the first study that investigates the diagnostic performance of clinical signs for the diagnosis of FRI. Noteworthy is that pain was not specific for FRI and therefore had a low diagnostic performance. Indeed, pain can be caused by multiple conditions and in the trauma patient pain may be related to the fracture or soft tissue injury. On the other hand, local redness, warmth and swelling, which typically occur simultaneously as local signs of inflammation, were associated with high specificities in

Table 6
Subanalyses using sets of suggestive signs in patients who presented without clinical confirmatory signs.

	FRI (N = 125) N*		Control (N = 157) N*		p-value	N*	Sensitivity (95% CI)	Specificity (95% CI)	AUROC (95% CI)	p-value
Evaluation of suggestive clinical signs and/or laboratory signs										
Redness	125	49 (39.2)	157	8 (5.1)	**<0.001**	282	39.2 (30.6-48.3)	94.9 (90.2-97.8)	0.67 (0.62-0.72)	**<0.001**
Swelling	125	49 (39.2)	157	18 (11.5)	**<0.001**	282	39.2 (30.6-48.3)	88.5 (82.5-93.1)	0.64 (0.59-0.69)	**<0.001**
Local warmth	125	20 (16.0)	157	4 (2.5)	**<0.001**	282	16.0 (10.1-23.6)	97.5 (93.6-99.3)	0.57 (0.53-0.60)	**<0.001**
Redness/swelling/warmth[1]	125	66 (52.8)	157	25 (15.9)	**<0.001**	282	52.8 (43.7-61.8)	84.1 (77.4-89.4)	0.68 (0.63-0.74)	**<0.001**
Fever	125	12 (9.6)	157	2 (1.3)	**0.002**	282	9.6 (5.1-16.2)	98.7 (95.5-99.8)	0.54 (0.51-0.57)	**0.003**
Wound drainage	125	36 (28.8)	157	4 (2.5)	**<0.001**	282	28.8 (21.1-37.6)	97.5 (93.6-99.3)	0.63 (0.59-0.67)	**<0.001**
Any clinical sign (redness/swelling/warmth or fever or wound drainage)[1]	125	83 (66.4)	157	30 (19.1)	**<0.001**	282	66.4 (57.4-74.6)	80.9 (73.9-86.7)	0.74 (0.69-0.79)	**<0.001**
WBC count elevation[2]	103	43 (41.7)	92	10 (10.9)	**<0.001**	195	41.7 (32.1-51.0)	89.1 (80.9-94.7)	0.65 (0.60-0.71)	**<0.001**
Any clinical sign (redness/swelling/warmth or fever or wound drainage) or WBC count elevation[1,2]	110	88 (80.0)	103	36 (35.0)	**<0.001**	213	80.0 (71.3-87.0)	65.0 (55.0-74.2)	0.73 (0.66-0.78)	**<0.001**

Prevalence data are shown as N (%) for patients with versus without FRI. Significance of difference is tested using the Chi-square or Fisher's exact test, as applicable. Diagnostic performance parameters sensitivity and specificity are shown as percentages with 95% CI. The area under the receiver operating curve (AUROC) is shown with 95% CI and associated p-value. N*, number of patients for whom data are available. FRI: fracture-related infection; WBC: white blood cell.

[1] The criterion was scored as 'present' if any one of the mentioned signs was present, while it was scored as 'absent' if none of the mentioned signs were present.

[2] Because WBC count was not performed in all patients, the percentage and diagnostic performance was calculated relative to the patients for whom WBC count was available.

Table 7
Subanalyses using sets of suggestive signs in patients who presented with phenotypically indistinguishable microorganisms isolated from at least two separate deep tissue specimens.

	FRI (N = 426) N*		Control (N = 157) N*		p-value	N*	Sensitivity (95% CI)	Specificity (95% CI)	AUROC (95% CI)	p-value
Evaluation of suggestive clinical signs and/or laboratory signs										
Redness	426	241 (56.6)	157	8 (5.1)	**<0.001**	583	56.6 (51.7-61.3)	94.9 (90.2-97.8)	0.76 (0.73-0.79)	**<0.001**
Swelling	426	194 (45.5)	157	18 (11.5)	**<0.001**	583	45.4 (40.7-50.4)	88.5 (82.5-93.1)	0.67 (0.64-0.71)	**<0.001**
Local warmth	426	97 (22.8)	157	4 (2.5)	**<0.001**	583	22.8 (18.9-27.1)	97.5 (93.6-99.3)	0.60 (0.58-0.63)	**<0.001**
Redness/swelling/warmth[1]	426	294 (69.0)	157	25 (15.9)	**<0.001**	583	69.1 (64.4-73.4)	84.1 (77.4-89.4)	0.77 (0.73-0.80)	**<0.001**
Fever	426	47 (11.0)	157	2 (1.3)	**<0.001**	583	11.0 (8.2-14.4)	98.7 (95.5-99.8)	0.55 (0.53-0.57)	**<0.001**
Wound drainage	426	169 (39.7)	157	4 (2.5)	**<0.001**	583	39.7 (35.0-44.5)	97.5 (93.6-99.3)	0.69 (0.66-0.71)	**<0.001**
Any clinical sign (redness/swelling/warmth or fever or wound drainage)[1]	426	350 (82.2)	157	30 (19.1)	**<0.001**	583	82.2 (78.2-85.7)	80.9 (73.9-86.7)	0.82 (0.78-0.85)	**<0.001**
WBC count elevation[2]	375	157 (41.9)	92	10 (10.9)	**<0.001**	467	41.9 (36.8-47.0)	89.1 (80.9-94.7)	0.66 (0.61-0.70)	**<0.001**
Any clinical sign (redness/swelling/warmth or fever or wound drainage) or WBC count elevation[1,2]	401	366 (91.3)	103	36 (35.0)	**<0.001**	504	91.3 (88.1-93.8)	65.0 (55.0-74.2)	0.78 (0.73-0.84)	**<0.001**

Prevalence data are shown as N (%) for patients with versus without FRI. Significance of difference is tested using the Chi-square or Fisher's exact test, as applicable. Diagnostic performance parameters sensitivity and specificity are shown as percentages with 95% CI. The area under the receiver operating curve (AUROC) is shown with 95% CI and associated p-value. N*, number of patients for whom data are available. FRI: fracture-related infection; WBC: white blood cell.

[1] The criterion was scored as 'present' if any one of the mentioned signs was present, while it was scored as 'absent' if none of the mentioned signs were present.

[2] Because WBC count was not performed in all patients, the percentage and diagnostic performance was calculated relative to the patients for whom WBC count was available.

this study. The presence of at least one of the clinical signs of redness, warmth or swelling was associated with a sensitivity of 69.4% and a specificity of 84.1%, indicating that if any of these signs are present, there should be a high index of suspicion for an FRI. Moreover, attributing these signs to a 'superficial' infection (i.e., cellulitis) should be done with caution.

Interestingly, although only a small number of patients presented with fever or wound drainage, resulting in low sensitivities, the individual specificities of these signs exceeded 95%. The high specificity (low false positive rate) indicates that these signs may be very helpful if present during the initial patient presentation (e.g., at the outpatient clinic). If one of these signs is present (in a patient were an FRI is already suspected), the presence of an FRI should be strongly considered. However, absence of these signs does not rule out infection and further investigations and close follow-up may still be required.

Radiological and nuclear imaging signs

Standard radiological techniques such as conventional radiography and CT can detect secondary signs of infection, such as nonunion, bone lysis and implant failure. However, because these signs can also occur in aseptic cases, conventional radiography and CT have a low diagnostic performance for FRI [20,21]. This is supported by our results, as signs on conventional radiography were more common in the control group than in the FRI group. Therefore, these radiological methods (i.e., x-ray, CT) seem more suited for assessing fracture healing and for surgical planning.

MRI generally has a better resolution to detect presence of inflammation in the soft tissues. However, MRI does not differentiate between infection and aseptic inflammation and the presence of fracture fixation devices may cause artefacts [20,22]. In the current study, MRI was only performed in eight cases, which prevents drawing reliable conclusions on its diagnostic value for FRI.

Nuclear imaging was available at all participating centers and all centers used the same imaging protocols [6]. WBC scan had a sensitivity of 50% and a specificity of 85%, which are lower than the pooled sensitivity and specificity rates described in a recent systematic review (86% and 96%, respectively) [21]. In the same systematic review ^{18}F-FDG-PET was associated with a pooled sensitivity and specificity of 93% and 79% [21]. In our study, the sensitivity and specificity of ^{18}F-FDG-PET were 65.2% and 100%, respectively. However, only a small number of our patients underwent nuclear imaging (FDG-PET: $n = 31$, WBC scan: $n = 44$), which resulted in wide confidence intervals, making it difficult to draw conclusions. Prospective studies are therefore required to gain insight in the diagnostic value of imaging modalities for FRI and their potential added value when combined with certain clinical signs [23].

Laboratory signs

Suggestive laboratory signs include the elevation of serum inflammatory markers, such as ESR, WBC count, and CRP. In our study population, an elevated WBC count was associated with a sensitivity of 38.4% and a specificity of 89.1%. These results are in line with previous publications reporting sensitivities and specificities ranging from 22.9% to 72.6% and from 73.5 to 85.7%, respectively [24]. In a recent study by Sigmund et al., a lower sensitivity of 17% and a similar specificity of 95% was reported for WBC count [25]. Therefore, WBC count elevation strongly suggests an FRI, but when the WBC count is not elevated, FRI cannot be ruled out.

In our study, an elevated CRP had a sensitivity of 78.3% and a specificity of 52.6%. This corresponds to values described in the literature (reporting sensitivities between 60% and 100% and specificities between 34.3% and 85.7%) [24]. Sigmund et al. reported a slightly lower sensitivity (67%) and slightly higher specificity (61%) for CRP elevation compared to our study [25]. The relatively low specificity highlights the high false positive rate of CRP, since elevated CRP levels in FRI cases can also be caused by inflammation due to soft tissue injury, the fracture itself, recent surgical procedures as well as infections in other locations, rheumatologic disease, acute coronary syndrome and allergies [6]. In previous studies only patients with chronic/late-onset FRI were included, which makes it difficult to compare their results with ours [24,25]. Furthermore, our results should be interpreted with caution as serum inflammatory markers were not determined in all patients, especially in the control group. This was particular the case for ESR. Our study indicates that serum inflammatory markers should remain suggestive signs for the diagnosis of FRI [6].

Microbiology

The diagnostic performance of a single positive culture with a virulent pathogen was evaluated in patients who did not meet the microbiological confirmatory criteria. A sensitivity of 29.3% and a specificity of 100% was found for this criterion. Therefore, if a single culture is positive with a virulent pathogen, this should raise a very high suspicion that infection is present. The clinical importance of a single positive culture with a virulent pathogen has already been highlighted in the diagnostic criteria for PJI [16] and our study demonstrates, for the first time to our knowledge, the diagnostic importance of this criterion in FRI.

2) Evaluate the effect of the combination of certain suggestive and/or confirmatory criteria on the diagnostic performance of the consensus definition

A secondary analysis was performed using combinations of suggestive and/or confirmatory criteria. It was hypothesized that combining confirmatory signs with other, suggestive signs could further increase sensitivity, without reducing specificity.

The highest diagnostic performance was present when clinical confirmatory criteria were evaluated in combination with microbiological criteria, (i.e., confirmatory and suggestive microbiological criteria), in which a sensitivity of 98.6% and a specificity of 100% was found.

Inclusion of one positive culture, with a virulent pathogen, improved the overall performance of the definition (AUROC of 0.99), making this criterion close to confirmatory. However, our sample size for single positive cultures was small ($n = 17$) so this finding must be interpreted with caution. Furthermore, negative cultures were found in 41 (8.5%) patients in the FRI group. The number of culture-negative infections in our cohort is comparable to the orthopedic device-related (i.e., PJI, FRI) literature, where the rate ranges between 6% and 15% [26–30]. Bacteria are not uniformly distributed in FRI tissue, so it is possible to harvest several specimens with very few or even no bacteria. Therefore, the recommendation to harvest at least five separate samples should be emphasized [6,9,10,12].

To simulate our daily clinical practice, where a patient often presents for the first time with a suspicion of FRI, we specifically evaluated sets of criteria that can readily be assessed upon patient presentation. Finding any of the clinical confirmatory criteria or, in the absence of such criteria, finding any of the suggestive clinical criteria of redness, warmth, swelling, wound drainage or fever resulted in a high diagnostic performance (Table 5). In the subgroup of patients without clinical confirmatory signs at presentation, the diagnostic value of these select clinical suggestive signs was again evaluated. The highest discriminatory value was found in this subgroup of patients when the presence of any clinical sign (redness/swelling/local warmth or fever or wound drainage) was applied as a diagnostic criterion (Table 6). Our study highlights the importance of these suggestive signs and additional research with larger patient cohorts is needed to further clarify the role of these signs. It is possible that a set of suggestive criteria proven to have extremely high specificity may serve as a new confirmatory criterion in the future.

Limitations

This study has several limitations. First, there was a lack of a gold standard to classify patients as FRI or control patients. In this study, we used 'intention to treat' recommended by an experienced multidisciplinary team as a reference standard. Such an approach may have led to over- or under-diagnosing of FRI. However, 97.5% of patients fulfilled confirmatory FRI criteria, therefore the potential for overdiagnosis is extremely small. Furthermore, all control patients were followed up for a minimum of 18 months and none developed an FRI of the affected limb during the follow-up period, therefore the potential for underdiagnosis is extremely small as well. Furthermore, to further support our findings, we performed a subanalysis using the presence of the microbiological confirmatory criterion as the reference standard. The presence of phenotypically indistinguishable microorganisms isolated from at least two separate deep tissue specimens has been validated in an earlier study [10]. The diagnostic performance of the selected suggestive signs was similar compared to the diagnostic performance of these signs calculated when using intention-to-treat as a reference standard (Table 7). Also, the risk of bias due to a possible difference in allocating patients to the FRI or control group based on the absence

(prior to 2018) or presence (from 2018 onwards) of the consensus definition is minimized due to the fact that the majority of cases in our study were diagnosed and treated prior to the development of the FRI consensus definition.

Second, it is a retrospective study that is subject to information bias due to missing data or misclassification of data in the patients' medical files. However, to avoid errors and bias during data collection, medical data were scored by multiple reviewers.

Third, this multicenter, multi-national study is subject to local preferences regarding the diagnostic modalities used and collected data dating back to January 2015. Because the consensus definition was published in 2018 [3] and histopathological analysis was only included as a confirmatory criterion in the updated version (which was published in 2019)(6), histopathology was not considered confirmatory for infection before that time. As a result, during the study period histopathological analysis was only performed in a small sample (90 FRI patients, 24 control patients), and therefore excluded from secondary analyses. Furthermore, other tests, such as certain inflammatory markers (e.g., ESR), radiological imaging techniques (e.g., MRI) and nuclear imaging were not performed in all patients, which resulted in limited sample sizes and wide confidence intervals for the diagnostic performance of some modalities. However, we had an adequate sample size to confidently assess the diagnostic role of most criteria and especially of the clinical criteria that would be the first to be assessed upon patient presentation.

Finally, we did not take the timing of the infection into consideration to evaluate whether the diagnostic performance of criteria varied between acute/early-onset and chronic/late-onset cases. Such an evaluation was not the primary aim of this validation study and our sample size would not have been large enough to stratify our FRI and control groups to subgroups based on time of onset. However, to our knowledge, this study is the largest study assessing the overall performance of the diagnostic consensus criteria for FRI, at all time-points and can serve as a baseline for future investigations to assess the impact of the timing of onset of symptoms.

Conclusions

The presence of any confirmatory criterion identified the vast majority of patients with an FRI and was associated with an excellent diagnostic discriminatory value. Therefore, our study validates the confirmatory criteria of the FRI consensus definition. Infection is highly likely in case of a single positive culture with a virulent pathogen. Specificities of at least 95% were found for the clinical suggestive signs of fever, wound drainage, local warmth and redness. In case these signs (individually or in combination) are observed, it is highly likely that – even in the absence of clinical confirmatory signs – an FRI is present and the treating physician should be careful to avoid misdiagnosing the infection as a superficial one.

Funding and support

Conflict of interest

All authors declare no conflict of interest with respect to the preparation and writing of this article.

Acknowledgements

References

[1] Walter N, Rupp M, Hierl K, Pfeifer C, Kerschbaum M, Hinterberger T, et al. Long-term patient-related quality of life after fracture-related infections of the long bones. Bone Joint Res 2021;10(5):321–7.

[2] Iliaens J, Onsea J, Hoekstra H, Nijs S, Peetermans WE, Metsemakers WJ. Fracture-related infection in long bone fractures: A comprehensive analysis of the economic impact and influence on quality of life. Injury 2021.

[3] Metsemakers WJ, Morgenstern M, McNally MA, Moriarty TF, McFadyen I, Scarborough M, et al. Fracture-related infection: A consensus on definition from an international expert group. Injury 2018;49(3):505–10.

[4] Metsemakers WJ, Morgenstern M, Senneville E, Borens O, Govaert GAM, Onsea J, et al. General treatment principles for fracture-related infection: recommendations from an international expert group. Arch Orthop Trauma Surg 2020;140(8):1013–27.

[5] Metsemakers WJ, Kortram K, Morgenstern M, Moriarty TF, Meex I, Kuehl R, et al. Definition of infection after fracture fixation: A systematic review of randomized controlled trials to evaluate current practice. Injury 2018;49(3):497–504.

[6] Govaert GAM, Kuehl R, Atkins BL, Trampuz A, Morgenstern M, Obremskey WT, et al. Diagnosing fracture-related infection: current concepts and recommendations. J Orthop Trauma 2020;34(1):8–17.

[7] Sliepen J, Onsea J, Zalavras CG, Depypere M, Govaert GAM, Morgenstern M, et al. What is the diagnostic value of the Centers for Disease Control and Prevention criteria for surgical site infection in fracture-related infection? Injury 2021.

[8] Obremskey WT, Metsemakers WJ, Schlatterer DR, Tetsworth K, Egol K, Kates S, et al. Musculoskeletal infection in orthopaedic trauma: assessment of the 2018 international consensus meeting on musculoskeletal infection. J Bone Joint Surg Am Vol 2020;102(10):e44.

[9] Dudareva M, Barrett L, Figtree M, Scarborough M, Watanabe M, Newnham R, et al. Sonication versus tissue sampling for diagnosis of prosthetic joint and other orthopedic device-related infections. J Clin Microbiol 2018;56(12).

[10] Dudareva M, Barrett LK, Morgenstern M, Atkins BL, Brent AJ, McNally MA. Providing an evidence base for tissue sampling and culture interpretation in suspected fracture-related infection. J Bone Joint Surg Am Vol 2021;103(11):977–83.

[11] Hellebrekers P, Rentenaar RJ, McNally MA, Hietbrink F, Houwert RM, Leenen LPH, et al. Getting it right first time: The importance of a structured tissue sampling protocol for diagnosing fracture-related infections. Injury 2019;50(10):1649–55.

[12] McNally M, Govaert G, Dudareva M, Morgenstern M, Metsemakers WJ. Definition and diagnosis of fracture-related infection. EFORT Open Rev 2020;5(10):614–19.

[13] Hackl S, Keppler L, von Rüden C, Friederichs J, Perl M, Hierholzer C. The role of low-grade infection in the pathogenesis of apparently aseptic tibial shaft nonunion. Injury 2021;52(11):3498–504.

[14] Dapunt U, Spranger O, Gantz S, Burckhardt I, Zimmermann S, Schmidmaier G, et al. Are atrophic long-bone nonunions associated with low-grade infections? Ther Clin Risk Manag 2015;11:1843–52.

[15] Bori G, Soriano A, García S, Mallofré C, Riba J, Mensa J. Usefulness of histological analysis for predicting the presence of microorganisms at the time of reimplantation after hip resection arthroplasty for the treatment of infection. J Bone Joint Surg Am Vol 2007;89(6):1232–7.

[16] McNally M, Sousa R, Wouthuyzen-Bakker M, Chen AF, Soriano A, Vogely HC, et al. The EBJIS definition of periprosthetic joint infection. Bone Joint J 2021;103-b(1):18–25.

[17] Morgenstern M, Athanasou NA, Ferguson JY, Metsemakers WJ, Atkins BL, McNally MA. The value of quantitative histology in the diagnosis of fracture-related infection. Bone Joint J 2018;100-b(7):966–72.

[18] Trebse R, Roskar S. Evaluation and interpretation of prosthetic joint infection diagnostic investigations. Int Orthop 2021;45(4):847–55.

[19] Tsaras G, Maduka-Ezeh A, Inwards CY, Mabry T, Erwin PJ, Murad MH, et al. Utility of intraoperative frozen section histopathology in the diagnosis of periprosthetic joint infection: a systematic review and meta-analysis. J Bone Joint Surg Am Vol 2012;94(18):1700–11.

[20] Bosch P, Glaudemans AWJM, de Vries J-PPM, Middelberg TR, Govaert GAM, Ijpma FFA. Nuclear imaging for diagnosing fracture-related infection. Clin Transl Imaging 2020;8(4):289–98.

[21] Zhang Q, Dong J, Shen Y, Yun C, Zhou D, Liu F. Comparative diagnostic accuracy of respective nuclear imaging for suspected fracture-related infection: a systematic review and Bayesian network meta-analysis. Arch Orthop Trauma Surg 2021;141(7):1115–30.

[22] Govaert GA, IJpma FF, McNally M, McNally E, Reininga IH. Glaudemans AW. Accuracy of diagnostic imaging modalities for peripheral post-traumatic osteomyelitis - a systematic review of the recent literature. Eur J Nucl Med Mol Imaging 2017;44(8):1393–407.

[23] Govaert G, Hobbelink M, Reininga I, Bosch P, Kwee TC, de Jong PA, et al. The accuracy of diagnostic Imaging techniques in patients with a suspected Fracture-related Infection (IFI) trial: study protocol for a prospective multicenter cohort study. BMJ Open 2019;9(9):e027772.

[24] van den Kieboom J, Bosch P, Plate JDJ, IJ FFA, Kuehl R, McNally MA, et al. Diagnostic accuracy of serum inflammatory markers in late fracture-related infection: a systematic review and meta-analysis. Bone Joint J 2018;100-B(12):1542–50.

[25] Sigmund IK, Dudareva M, Watts D, Morgenstern M, Athanasou NA, McNally MA. Limited diagnostic value of serum inflammatory biomarkers in the diagnosis of fracture-related infections. Bone Joint J 2020;102-B(7):904–11.

[26] Berbari EF, Marculescu C, Sia I, Lahr BD, Hanssen AD, Steckelberg JM, et al. Culture-negative prosthetic joint infection. Clin Infect Dis 2007;45(9):1113–19.

[27] Kuehl R, Tschudin-Sutter S, Morgenstern M, Dangel M, Egli A, Nowakowski A, et al. Time-dependent differences in management and microbiology of orthopaedic internal fixation-associated infections: an observational prospective study with 229 patients. Clin Microbiol Infect 2019;25(1):76–81.

[28] Palan J, Nolan C, Sarantos K, Westerman R, King R, Foguet P. Culture-negative periprosthetic joint infections. EFORT Open Rev 2019;4(10):585–94.

[29] Rupp M, Baertl S, Walter N, Hitzenbichler F, Ehrenschwender M, Alt V. Is there a difference in microbiological epidemiology and effective empiric antimicrobial therapy comparing fracture-related infection and periprosthetic joint infection? A retrospective comparative study. Antibiotics (Basel) 2021;10(8).

[30] Gitajn IL, Heng M, Weaver MJ, Ehrlichman LK, Harris MB. Culture-negative infection after operative fixation of fractures. J Orthop Trauma 2016;30(10):538–44.