



## Research article

## Semi-quantitative CT assessment of fracture healing: How many and which CT reformats should be analyzed?

Vasiliki Perlepe\*, Nicolas Michoux, Guy Heynen<sup>1</sup>, Bruno Vande Berg

Cliniques Universitaires Saint-Luc, IREC (Institut de Recherche Clinique et Experimentale), Avenue Hippocrate 10, 1200, Brussels, Belgium

## ARTICLE INFO

**Keywords:**  
Bone  
Fracture  
Non-union  
CT

## ABSTRACT

**Purpose:** To assess how many and which CT reformats of long bone non-unions should be analyzed to best approximate the analysis of a larger number of CT reformats obtained in the three orthogonal planes.

**Method:** We used 29 CT examinations of tibial or femoral non-unions to obtain 87 stacks of 7 CT reformats each in the coronal (n = 29), sagittal (n = 29) or transverse (n = 29) planes. Two independent radiologists scored two fracture sites on each CT reformat by using a Tomographic Union Score (TUS) (1: no callus, 2: non-bridging callus; 3: bridging immature callus; 4: bridging remodeled callus). The reference standard was the mean of the three TUS calculated from the cortical scores obtained on all the sagittal, coronal and transverse CT reformats. We determined the agreement (intra-class correlation coefficient (ICC) between the reference standard and 33 models combining one to seven CT reformats from one to three planes. The three best models were compared following a resampling procedure by a Wilcoxon's signed rank test.

**Results:** Three models combining two (mid-coronal and mid-sagittal), three (mid-coronal, mid-sagittal and mid-transverse) or four (two paramedian coronal and sagittal) CT reformats had the highest ICC (ICC ≥ 0.89) for both observers. After resampling, the model combining the two paramedian sagittal and coronal CT reformats statistically outperformed the two other models.

**Conclusion:** Semi-quantitative analysis of the two paramedian sagittal and coronal CT reformats is an acceptable alternative to the analysis of more numerous reformats.

## 1. Introduction

Evaluation of fracture healing is mainly based on clinical and radiographic findings [1–3]. Analysis of radiographs provides an overall estimation of fracture healing albeit with limited inter-observer and intra-observer reproducibility [4–6] most likely due to limitations associated with the projectional nature of radiographic image formation and overlapping metal hardware [7]. Computed tomography (CT) enables to assess the presence and extent of bone callus at fracture edges thanks to its multiplanar capacity and high spatial resolution at the cost of an additional examination and increased ionizing radiation burden [8–12]. The methodology to assess the large amount of data afforded by CT has not been clearly defined yet [13]. We hypothesized that the analysis of a subset of CT reformats could be sufficient to provide a representative overview of the entire fracture. The current study aimed at determining the best combinations of CT reformat that

would provide a reliable overview of non-united fractures in a series of patients with non-united surgically treated femoral or tibial fractures.

## 2. Material and methods

## 2.1. Study population

Authorization of the institutional ethic committee for this retrospective study was obtained and authorization from patients was waived. A search over a period of four years was performed in our picture archiving and communication systems (PACS)(Carestream Client version 11.3; Carestream Health, Rochester, NY, USA) by a fellow in musculoskeletal radiology (VP) with 2 years of experience to select patients aged between 18 and 65 years with a surgically treated tibial or femoral shaft fracture in whom CT had been performed and in whom non-union was later demonstrated at surgery. Exclusion criteria

**Abbreviations:** RUS, Radiographic union score; TUS, Tomographic Union Score; CT, computed tomography; DTS, digital tomosynthesis

\* Corresponding author at: Department of Radiology, Cliniques Universitaires St luc – UC Louvain, Avenue Hippocrate 10, 1200, Brussels, Belgium.

**E-mail addresses:** [vasiliki.perlepe@uclouvain.be](mailto:vasiliki.perlepe@uclouvain.be) (V. Perlepe), [Nicolas.michoux@uclouvain.be](mailto:Nicolas.michoux@uclouvain.be) (N. Michoux), [guy.heyne@bonetherapeutics.com](mailto:guy.heyne@bonetherapeutics.com) (G. Heynen), [bruno.vandenberg@uclouvain.be](mailto:bruno.vandenberg@uclouvain.be) (B. Vande Berg).

<sup>1</sup> Bone Therapeutics SA, Avenue Auguste Piccard, 37 6041 Goselies, Belgium.

<https://doi.org/10.1016/j.ejrad.2019.07.026>

Received 14 April 2019; Received in revised form 19 June 2019; Accepted 18 July 2019

0720-048X/ © 2019 Elsevier B.V. All rights reserved.

were (a) healed fractures, (b) fractures of less than 6 months, (c) non-united fractures that had been treated with allo- or auto-grafts and (d) fractures with imaging or biological signs of infections. Our study population consisted of 17 male and 12 female patients (mean  $\pm$  standard deviation (SD) : 44.44 years  $\pm$  14.4; range 19–64 years). There were 29 cases of non-unions in 12 nailed and four plated tibial fractures and in 12 nailed and one plated femoral fractures. The mean time delay ( $\pm$ SD) between fracture occurrence and CT examination was 595 days ( $\pm$  482) (range: 180–1983 days). The mean ( $\pm$  SD) time delay between the CT examination and surgery was 168 days ( $\pm$  241) (range: 1–1063 days).

## 2.2. Imaging protocols

All unenhanced CT examinations were performed on a 40-detector row CT scanner (Somatom Definition AS; Siemens Healthcare, Forchheim, Germany), patients lying supine with the feet first. Data acquisition was obtained over the entire fractured bone using the following tube parameters: tube voltage, 120 kVp; reference tube current-time product, 275 mAs. Effective current time-product, 178–305 mAs by applying automated tube current modulation (Care Dose 4D; Siemens Healthcare); detector configuration, 40 X 0.6 mm; pitch, 0.8; gantry rotation time, 1 s. CT images were reconstructed with a section width and increment of 0.75 mm using standard bone (B70 h) and soft tissue (B41 s) deconvolution kernels and feed-back retroprojection calculation. Field-of-view varied from 12 x 12 cm to 20 x 20 cm. No iterative reconstruction algorithms were used. All reconstructions were stored in our PACS.

Between January and March 2015, the same fellow obtained sagittal, coronal and transverse reformats in a standardized manner as follows. For each case, the axial images obtained with the bone deconvolution kernel were uploaded in the multiplanar reconstruction software available in our PACS. A stack of nine 1.5 mm-thick reformats in the sagittal, coronal and transverse planes were obtained according to anatomic landmarks. For the coronal and sagittal planes, the first (#1) and last (#9) reformats were placed in contiguity with the target bone which was only visible on reformats #2-8 (Figs. 1–3). For the transverse plane, the first (#1) and last (#9) reformats were positioned above and below the fracture level and fracture sites were visible only on reformats #2-8. The slice intergap (mean  $\pm$  standard deviation) subsequent to the positioning of the stack of reformats in order to

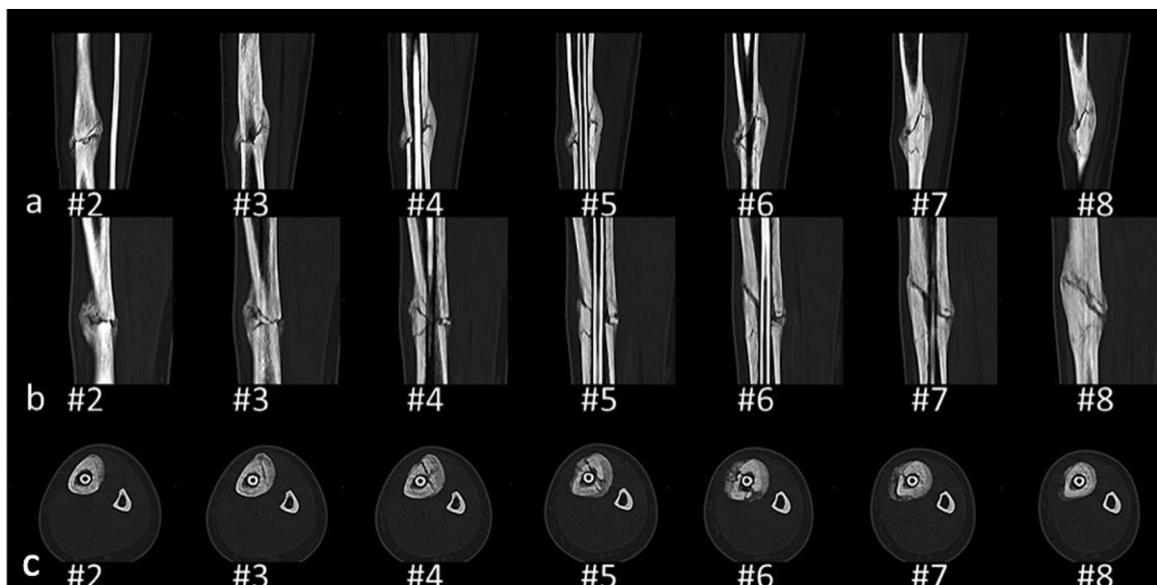
include the entire fracture was 2.2 mm  $\pm$  0.4 for the long axis reformats and 8.7mm  $\pm$  4.9 for the transverse reformats. A data manager anonymized all the stacks, labelled them in a random order by using numerals and stored them in a dedicated research PACS.

## 2.3. Image analysis

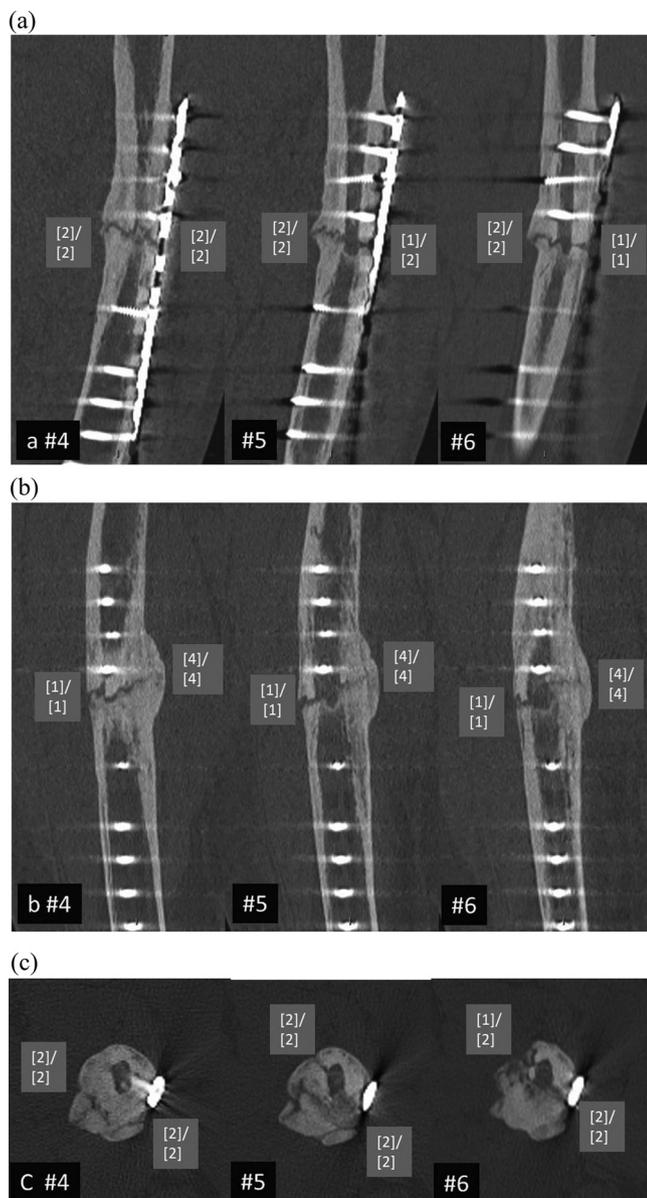
Between May and October 2016, the fellow in musculoskeletal radiology and a musculoskeletal radiologist with 24 years of experience independently scored the CT reformats on a PACS workstation with adapted window width and level values. Coronal, sagittal and transverse CT reformats were analyzed separately during three different reading periods separated by a one-month interval. On each of the seven CT reformats #2 to #8, readers scored two fracture sites by using a Tomographic Union Score (TUS) based on the callus appearance on the CT reformat. The TUS was derived from the previously described Radiological Union Score (RUS) by Litrenta et al who scored fracture healing on radiographs [14]. The score at the fracture edges was defined as follows: score 1: no callus; score 2: discontinuous callus, score 3: continuous bridging immature callus; score 4: continuous remodeled callus. On each CT reformat, two fracture sites had to be scored (one score for the anterior and one for the posterior cortex on the sagittal reformats, one score for the medial and one for the lateral cortex on the coronal reformat, two scores for the transverse reformat). When several fracture sites were detected, the two least healed sites were scored. Scoring was considered to be non-feasible when metal artifacts obscured the fracture, when one of the fracture edges was not visible due to bone deformity or when the fracture plane was not visible mainly when the orientation of the reformats was parallel to that of the fracture.

For each case, a  $TUS^{Sa}$  was derived from the scoring of the seven sagittal reformats, a  $TUS^{Co}$  was derived from the seven coronal reformats and a  $TUS^{Tr}$  was derived from the seven transverse reformats. TUS could range from 14 to 56 because 14 scores ranging from 1 to 4 were derived from the seven reformats and added. In our study, the standard of reference for each fracture was the mean of  $TUS^{Sa}$ ,  $TUS^{Co}$  and  $TUS^{Tr}$ .

In January 2017, the same two readers independently reviewed separately the median sagittal, coronal and transverse CT reformats without seeing the other reformats of the same stack or of the other planes to assess intra-observer and inter-observer agreement in scoring.



**Fig. 1.** (a) Coronal, (b) sagittal and (c) transverse CT reformats of a nailed non-united tibial fracture in a 55-year-old man. Lateral (#2 and #8), paralateral (#3 and #7), paramedian (#4 and #6) and median (#5) CT reformats were obtained in the three orthogonal planes.

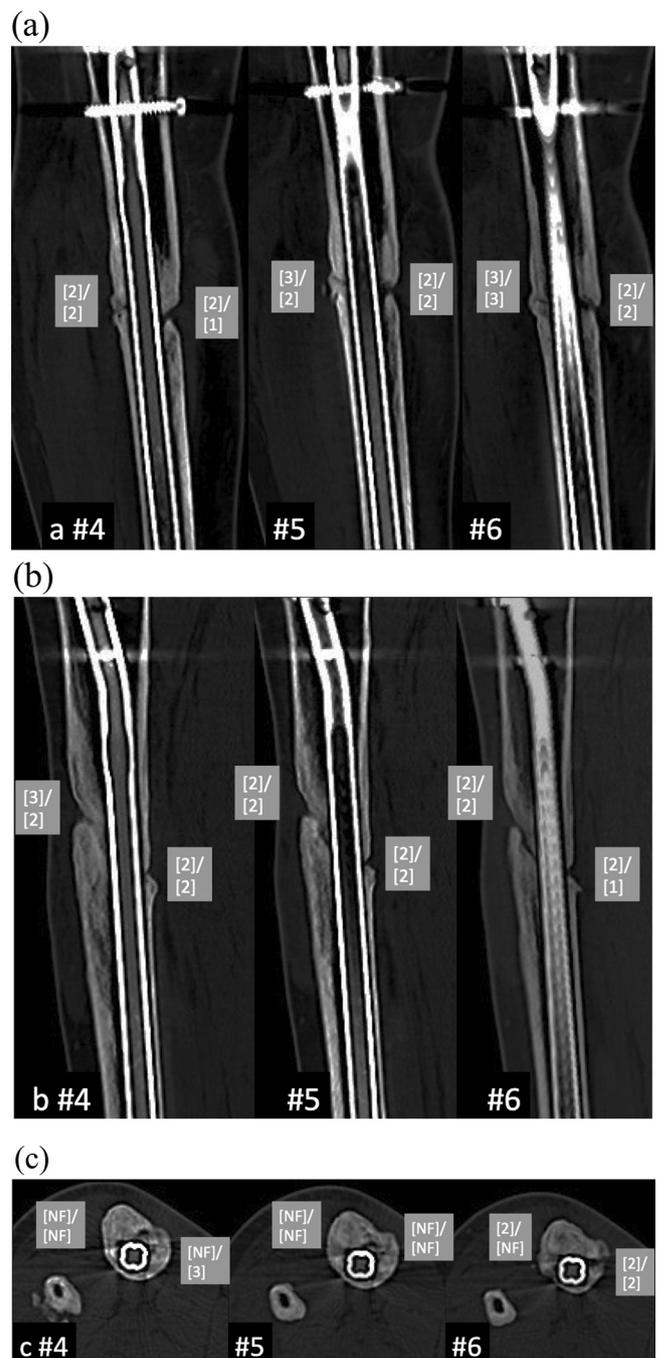


**Fig. 2.** (a) Median (#5) and paramedian (#4 and #6) coronal, (b) sagittal and (c) transverse CT reformats of a plated non-united femoral fracture in a 27-year-old man. Numbers in brackets indicate scores given by reader 1 / reader 2, respectively (1: no callus, 2: discontinuous callus, 3: continuous immature, non-remodeled callus, 4: continuous remodeled callus). The ratio between the sum of the scores and the maximum score obtained on the three median reformats (12/24 for reader 1 and 13/24 for reader 2) is very close to that obtained on the two paramedian reformats (24/48 for reader 1 and 25/48 for reader 2).

#### 2.4. Statistical analysis

The rate of feasibility of the cortical scoring was determined for the two readers on the three planes and differences in frequency of non-feasibility were compared between the three planes. Intra- and inter-observer reproducibility of the cortical scoring of the median CT reformat of the three planes was assessed by determining the intra-class correlation coefficient (ICC) based on a two-way model after exclusion of the non-assessable areas. Inter-observer agreement for TUS<sup>Sa</sup>, TUS<sup>Co</sup> and TUS<sup>Tr</sup> was assessed by determining the intra-class correlation coefficient (ICC) based on a two way model after exclusion of the non-assessable areas for all slices obtained in each plane.

Thirty-three models including different combinations of one to seven reformats and one to three planes were constructed (Tables 1 and



**Fig. 3.** (a) Median (#5) and paramedian (#4 and #6) coronal, (b) sagittal and (c) transverse CT reformats of a nailed non-united tibial fracture in a 28-year-old man. Numbers in brackets indicate scores given by reader 1 / reader 2, respectively. The ratio between the sum of the scores and the maximum score obtained on the three median reformats (9/16 for reader 1 and 8/16 for reader 2) is very close to that obtained on the two paramedian reformats (22/40 for reader 1 and 20/40 for reader 2).

2). The TUS calculated for each of these models were compared with the TUS resulting from all reformats in the three planes by calculating ICC and median difference in total score. Missing cortical scores were calculated as the mean of the available scores from the same plane.

Finally, statistically significant differences in ICC between the models combining one (#5), two (#4 and #6) or three (#4, #5 and #6) CT reformats obtained in the coronal and sagittal planes with those combining either the transverse and sagittal or the transverse and coronal planes were investigated following a resampling procedure

**Table 1**  
Correlation coefficients (ICC) and median differences in score value between the reference standard and 33 models combining different planes and image numbers for reader 1. Numbers in [] are 95%CI.

Slice number	Three planes	Two planes			One plane		
	Tr, Co, Sa	Co Sa	Tr Co	Tr Sa	Tr	Co	Sa
#5	0.90 [0.80; 0.95] -1.5 [-4.0; +1.5]	0.91 [0.81; 0.95] +0.5 [-1.5; +2.5]	0.81[0.64; 0.91] -1.0[-4.5; +2.5]	0.80 [0.62; 0.90] -2.0 [-6.0; +2.5]	0.57 [0.26; 0.77] -5.0 [-11.0; +0.5]	0.73 [0.50; 0.86] +0.5 [-3.5; +4.5]	0.83 [0.67; 0.92] +0.0 [-4.0; +4.0]
#4 # 6	0.97 [0.93; 0.98] -1.0 [-2.5; +0.5]	0.97[0.93; 0.98] -0.5 [-2.5; +1.0]	0.94 [0.88; 0.97] -1.5[-3.5; +1.0]	0.90 [0.81; 0.95] -1.5[-4.5; +0.5]	0.83 [0.66; 0.91] -3.0[-6.0; +0.0]	0.85 [0.71; 0.93] +0.0 [-2.5; +3.0]	0.83 [0.67; 0.92] -1.0 [-4.5; +3.5]
#4 #5 #6	0.96 [0.91; 0.98] -1.5 [-3.0; +0.5]	0.96 [0.92; 0.98] -0.0 [-2.0; +1.5]	0.93 [0.85; 0.96] -1.0 [-3.5; +2.0]	0.89 [0.79; 0.95] -2.0 [-4.5; +1.0]	0.79[0.61; 0.90] +3.0 [-6.0; +0.0]	0.83 [0.66; 0.91] +0.5 [-3.0; +3.0]	0.85 [0.70; 0.93] -1.0 [-4.5; +3.0]
#3 #7	NA	NA	NA	NA	0.87 [0.74;0.94] +1.5 [-2.5; +4.5]	0.87 [0.75;0.94] +0.5 [-3.0; +4.0]	0.85 [0.71;0.93] +0.5 [-3.0; +4.0]
#3 #5 #7	NA	NA	NA	NA	0.93 [0.87;0.97] -0.5 [-3.0; +1.5]	0.86 [0.72;0.93] +0.5 [-3.0; +4.0]	0.87 [0.75;0.94] +0.5 [-3.0; +3.5]
#2 #4 #6 #8	NA	NA	NA	NA	0.81 [0.63;0.90] +2.5 [-2.0; +7.0]	0.87 [0.73;0.93] -0.5 [-2.5; +2.5]	0.90 [0.80;0.95] -1.0 [-4.5; +2.0]
#2-#8	NA	NA	NA	NA	0.90 [0.80;0.95] +1.0 [-2.0; +4.5]	0.87 [0.74;0.94] -0.0 [-2.5; +2.5]	0.90 [0.79;0.95] -0.5 [-3.5; +2.0]

Tr: transverse; Co: Coronal; Sa:Sagittal; #5: median CT reformat; #4 and# 6: paramedian CT reformats, #3 and #7: paralateral CT reformats; #2 and #8 : lateral CT reformats; NA: not assessed.

based on 200 samples of 14 patients randomly drawn from the whole series. A Wilcoxon signed rank test was performed to compare differences between the models. A Bonferroni-type correction for performing nine comparisons between models was applied and a p-value < 0.0056 was regarded as statistically significant.

### 3. Results

#### 3.1. Feasibility and reproducibility of the scoring

The rate of feasibility of the scoring for reader 1/reader 2 was 85% / 85% for the transverse reformats, 97%/100% for the coronal reformats and 98.5%/99% for the sagittal reformats. Missing scores among the expected 1218 scores were due to presence of metal artifacts (n = 13 and 5 for reader 1 and 2), absence of one of the fracture edge on the reformat (n = 8 and 5 for reader 1 and 2) and non-visibility of the fracture (n = 49 and 52 for reader 1 and 2). Missing scores were statistically significantly more frequent in the transverse than in the coronal and sagittal planes (p < 10<sup>-4</sup>).

The intra-observer reproducibility for the scoring performed on the median reformat was moderate on the transverse plane (ICC: +0.54 [+0.24; +0.75]for R1 and +0.60 [+0.30; +0.79] for R2) and good on

the coronal (ICC: +0.72 [+0.48; +0.86]for R1 and +0.62 [+0.28; +0.81]for R2) and sagittal plane (ICC: +0.76 [+0.60; +0.89] for R1 and +0.61 [+0.26; +0.81] for R2). The inter-observer reproducibility for the scoring performed on the medial reformat was fair on the transverse plane (ICC: +0.33 [-0.02; +0.61]for R1 and +0.26 [-0.07; +0.55] for R2), moderate on the coronal plane (ICC: +0.46 [+0.12; +0.71]for R1 and +0.27 [-0.06; +0.56] for R2)and good on the sagittal plane (ICC: +0.65 [+0.33; +0.82]for R1 and +0.66 [+0.41; +0.83] for R2).

#### 3.2. Reference TUS values and reproducibility

The TUS values of reference for each plane ranged from 23 to 26 with very good inter-observer reproducibility in the transverse plane (ICC : 0.83 [+0.66; +0.91]) and good inter-observer reproducibility in the sagittal (ICC: +0.72 [+0.38; +0.87] and coronal (+0.72 [+0.48; +0.86]) planes.

#### 3.3. Comparison between tested models and reference TUS values

##### 3.3.1. Models combining three planes

All tested models based on three planes had an ICC value equivalent

**Table 2**

: Correlation coefficients (ICC) and median differences in score value between the reference standard and 33 models combining different planes and image numbers for reader 2. Numbers in [] are 95%CI. Numbers in [] are 95%CI.

Combinations	Three planes	Two planes			One plane		
	Tr, Co, Sa	Co Sa	Tr Co	Tr Sa	Tr	Co	Sa
#5	0.93 [0.86; 0.97] -1.5 [-3.5; +1.0]	0.93 [0.85; 0.96] +2.0 [-0.5; +4.5]	0.82[0.61; 0.92] -5.0 [-9.0; -2.0]	0.86[0.73; 0.93] -2.0 [-4.5; +2.0]	0.64[0.19; 0.84] -9.5 [-16.5; -4.5]	0.84 [0.68; 0.92] -1.0 [-0.5; +2.5]	0.73 [0.48; 0.86] +5.0 [-0.0; +11]
#4 # 6	0.97 [0.94; 0.99] +0.5 [-1.0; +1.5]	0.95 [0.89; 0.98] +2.5 [-0.0; +4.0]	0.93[0.85; 0.97] -2.0 [-4.0; +0.5]	0.91 [0.82; 0.96] +0.5[-2.0; +3.5]	0.75 [0.51; 0.88] -4.0[-8.5; +0.5]	0.88 [0.75; 0.94] +0.5 [-4.0; +2.5]	0.79[0.59; 0.90] +3.5 [+0.0; +7.5]
#4 #5 #6	0.98 [0.95; 0.99] -0.5 [-1.5; +1.0]	0.96 [0.91; 0.98] +2.0 [+0.0; +3.5]	0.93 [0.81; 0.97] -3.0 [-5.0; -0.5]	0.91 [0.82; 0.96] +0.0 [-3.0; +3.0]	0.76 [0.44; 0.89] -6.0 [-10.0; -2.0]	0.89 [0.78; 0.95] +0.0 [-3.5; +2.0]	0.80[0.59; 0.90] +4.0 [+0.5; +8.5]
#3 #7	NA	NA	NA	NA	0.80 [0.62;0.90] -0.5 [-5.5; +2.0]	0.80 [0.61;0.90] +1.0 [-4.0; +4.0]	0.76 [0.52;0.88] +4.0 [-0.0; +8.0]
#3 #5 #7	NA	NA	NA	NA	0.85 [0.65;0.93] -4.0 [-7.0;-0.5]	0.86 [0.72;0.93] -0.0 [-3.5; +3.0]	0.80 [0.58;0.91] +4.0 [+0.5; +8.5]
#2 #4 #6 #8	NA	NA	NA	NA	0.80 [0.61;0.90] -1.0 [-4.5; +2.5]	0.89 [0.77;0.95] -2.5 [-5.5; +0.5]	0.89 [0.78;0.95] +3.0[+0.0; +6.0]
#2-#8	NA	NA	NA	NA	0.84 [0.69;0.92] -2.5 [-6.0; +0.5]	0.90 [0.81;0.95] -1.0 [-4.5; +1.0]	0.87 [0.72;0.94] +3.5[+0.5; +6.0]

Tr: transverse; Co: Coronal; Sa:Sagittal; #5: median CT reformat; #4 and# 6: paramedian CT reformats, #3 and #7: paralateral CT reformats; #2 and #8 : lateral CT reformats; NA: not assessed.

or superior to 0.90 and difference in median values between -1.5 and +0.5 in comparison with the standard of reference for both readers without any statistically significant difference in value (Table 1,2).

### 3.3.2. Models combining two planes

Five of the nine tested models based on two planes had a ICC value equivalent or superior to 0.90 and difference in median values between -1.5 and + 2.5 in comparison with the standard of reference for both readers without any statistically significant difference in value (Table 1,2).

### 3.3.3. Models with one plane

Only four (reader 1) and one (reader 2) of the 21 models based on one plane had a ICC value equivalent or superior to 0.90 and difference in median values between -1.0 and +1.0 in comparison with the standard of reference without any statistically significant difference in value (Table 1,2).

### 3.4. Comparison between best models

For both readers, the median differences in ICC were statistically significantly closest to the reference standard with the models combining coronal and sagittal CT reformats than those combining coronal and transverse or sagittal and transverse CT reformats (Table 3). The median differences in ICC were the lowest for the models that combined the two paramedian reformats for the two readers (Table 3).

## 4. Discussion

Our study performed in a series of tibial and femoral shaft non-unions demonstrated that the semi-quantitative analysis of the median CT reformat obtained in the three orthogonal planes or of the 2 paramedian sagittal and coronal CT reformats was a valuable alternative to the analysis of a large number of CT reformats obtained in the three orthogonal planes.

First, we demonstrated that the scores derived from the analysis of the median sagittal, coronal and transverse CT reformats or from the two paramedian sagittal and coronal CT reformats had a very high correlation coefficient (> 0.90) with the scores obtained at analysis of all sagittal, coronal and transverse CT reformats. Most likely, the scoring of all reformats from the three planes leads to data over-sampling as the same cortical areas are analyzed on the three stacks i.e. the fracture area located in the most lateral segment of the diaphysis can be assessed on the most lateral sagittal CT reformat and on the medial coronal CT reformat.

Second, we demonstrated that the scores derived only from a limited number of sagittal and coronal CT reformats appeared to have a

higher correlation coefficient with the reference values than when the analyzed reformats were derived from the transverse CT reformats and one of the two long-axis CT reformats. Actually on the transverse reformats, analysis of the transverse component of the fractures was more difficult due to partial volume artifacts. In addition, analysis of the fracture sites were more often non feasible on short- than on long-axis reformats due to more metal-related artifacts on transverse images.

Finally, the analysis of all reformats from only one plane provided results that poorly correlated with the standard of reference. This observation could derive from the complex and variable spatial orientation of cortical interruption in fractures that, depending on their relation with the orientation of the plane may or may not be conspicuous on one of the plane.

Several difficulties interfered with the feasibility of the analysis including metal-related artifacts, non-visibility of one of the fracture edges due to bone deformity on long-axis images and limited conspicuity of the fracture line due to partial volume artifacts mainly on transverse images. Although metal-related artifacts could have been reduced by using specific post-processing software, their use could have induced some loss in spatial resolution and they were not used in this study [15].

The results of the current study could pave the way to the use of digital tomosynthesis (DTS) for semi-quantitative fracture healing evaluation. Tomographic images obtained at DTS enable to detect nondisplaced fractures and to evaluate the healing process [16–19]. The fact that DTS does not provide transverse tomographic images due to the translational nature of X-ray source displacement and to its limited multiplanar capabilities could therefore have limited implication in the evaluation of fracture healing with DTS since short -axis images that are lacking at DTS seem to be of limited interest.

Several limitations of the current study should be noted. First, a limited number of patients were studied and we focused on surgically proven non-united tibial and femoral shaft fractures. Second, we used a semi-quantitative scoring system derived from a radiographic scoring system that has been validated for the analysis of fracture healing on antero-posterior and lateral radiographs [14]. The current study did not assess the value of this tomographic scoring system and did not compare it with results obtained on radiographs of the same lesion. Third, the reproducibility of the score calculated on a single reformat was fair to moderate whereas it was good to very good when numerous reformats were assessed indicating that the analysis of one reformat may be influenced by that of the adjacent reformats. Finally, readers were not blinded to the planes of the reformats during their analysis.

In conclusion, the current study performed in a series of tibial and femoral shaft non-unions demonstrated that the analysis of the median coronal, sagittal and transverse CT reformats or of the two paramedian sagittal and coronal CT reformats was a valuable alternative to the

**Table 3**

Comparison for the two readers of the performance of the models combining one (#5), two (#4 and #6) or three (#4,#5 and #6) CT reformats obtained in the coronal and sagittal planes with those combining either the transverse and sagittal or the transverse and coronal planes.

Plane combinations	Reader 1		Reader 2	
	P value	Median difference in ICC [95%CI]	P value	Median difference in ICC [95%CI]
<b>CoSA<sup>+</sup> vs TrSa<sup>+</sup></b>				
#5 <sup>**</sup>	< 10 <sup>-4</sup>	+0.104 [+0.101; +0.107]	< 10 <sup>-4</sup>	+0.039 [+0.037; +0.042]
#4 and #6 <sup>**</sup>	< 10 <sup>-4</sup>	+0.031 [+0.030; +0.033]	< 10 <sup>-4</sup>	+0.039 [+0.038; +0.040]
#4, #5 and #6	< 10 <sup>-4</sup>	+0.072 [+0.070; +0.073]	< 10 <sup>-4</sup>	+0.051 [+0.050; +0.052]
<b>CoSA<sup>+</sup> vs TrCo<sup>+</sup></b>				
#5 <sup>**</sup>	< 10 <sup>-4</sup>	+0.095 [+0.092; +0.098]	< 10 <sup>-4</sup>	+0.078 [+0.076; +0.080]
#4 and #6 <sup>**</sup>	< 10 <sup>-4</sup>	+0.007 [+0.006; +0.008]	< 10 <sup>-4</sup>	+0.020 [+0.019; +0.020]
#4, #5 and #6	< 10 <sup>-4</sup>	+0.038 [+0.038; +0.039]	< 10 <sup>-4</sup>	+0.031 [+0.030; +0.032]

The median difference in ICC was statistically significantly closest to the reference standard with the models combining the coronal and sagittal imaging planes and difference was the lowest for the model combining the two paramedian reformats.

\* CoSa combines coronal and sagittal planes; TrSa combines transverse and sagittal planes; TrCo combines transverse and coronal planes.

\*\* #5 is median CT reformat; #4 and #6 are the two paramedian CT reformats.

analysis of a large number of CT reformats obtained in the three orthogonal planes. The observation that short-axis reformats were not mandatory could pave the way to the use of DTS - that only provides long-axis reformats- to semi-quantitatively evaluate the healing process of long bone fractures.

#### Funding source

This work was funded by an educational grant from Bone Therapeutics SA.

#### Declaration of Competing Interest

Guy Heynen is an employee of Bone Therapeutics SA.

#### References

- [1] P. Harwood, J. Newman, A. Michael, An update on fracture healing and non-union, *Orthop. Trauma* 24 (2010) 9–20123.
- [2] M. Bhandari, G.H. Guyatt, M.F. Swiontkowski, P. Tornetta 3rd, S. Sprague, E.H. Schemitsch, A lack of consensus in the assessment of fracture healing among orthopaedic surgeons, *J. Orthop. Trauma* 16 (2002) 562–566.
- [3] L.A. Corrales, S. Morshed, M. Bhandari, T. Miclau 3rd, Variability in the assessment of fracture-healing in orthopaedic trauma studies, *J. Bone Joint Surg. Am.* 90 (2008) 1862–1868.
- [4] D.B. Whelan, M. Bhandari, M.D. McKee, et al., Interobserver and intraobserver variation in the assessment of the healing of tibial fractures after intramedullary fixation, *J. Bone Joint Surg. Br.* 84 (2002) 15–18.
- [5] B.J. Davis, P.J. Roberts, C.I. Moorcroft, M.F. Brown, P.B. Thomas, R.H. Wade, Reliability of radiographs in defining union of internally fixed fractures, *Injury* 35 (2004) 557–561.
- [6] B.G. Dijkman, S. Sprague, E.H. Schemitsch, M. Bhandari, When is a fracture healed? Radiographic and clinical criteria revisited, *J. Orthop. Trauma* 24 (Suppl 1) (2010) 76–80.
- [7] V. Perlepe, P. Omoumi, A. Larbi, et al., Can we assess healing of surgically treated long bone fractures on radiograph? *Diagn. Interv. Imaging* 99 (2018) 381–386, <https://doi.org/10.1016/j.diii.2018.02.004>.
- [8] J.S. Fisher, J.J. Kazam, D. Fufa, R.J. Bartolotta, Radiologic evaluation of fracture healing, *Skeletal Radiol.* 48 (2019) 349–361, <https://doi.org/10.1007/s00256-018-3051-0>.
- [9] J.E. Kuhlman, E.K. Fishman, D. Magid, W.W. Scott, A.F. Brooker, S.S. Siegelman, Fracture nonunion: CT assessment with multiplanar reconstruction, *Radiology*. 167 (483) (1988) 48.
- [10] M. Grigoryan, J.A. Lynch, A.L. Fierlinger, et al., Quantitative and qualitative assessment of closed fracture healing using computed tomography and conventional radiography, *Acad. Radiol.* 10 (2003) 1267–1273.
- [11] T. Bhattacharyya, K.A. Bouchard, A. Phadke, J.B. Meigs, A. Kassarian, H. Salamipour, The accuracy of computed tomography for the diagnosis of tibial nonunion, *J. Bone Joint Surg. Am.* 88 (2006) 692–697 49.
- [12] C.R. Krestan, H. Noske, V. Vasilevska, et al., MDCT versus digital radiography in the evaluation of bone healing in orthopedic patients, *AJR Am. J. Roentgenol.* 186 (2006) 1754–1760.
- [13] Y.V. Kleinlugtenbelt, V.A. Scholtes, J. Toor, et al., Does Computed Tomography change our observation and management of fracture non-unions? *Arch. Bone Surg.* 4 (4) (2016) 337–342.
- [14] J. Litrenta, P. Tornetta, S. Mehta, et al., Determination of radiographic healing. An assessment of Consistency Using RUST and modified RUST in Metadiaphyseal Fractures, *J. Orthop. Trauma*. 29 (2015) 516–520.
- [15] R.H.H. Wellenberg, E.T. Hakvoort, C.H. Slump, M.F. Boomsma, M. Maas, G.J. Streekstra, Metal artifact reduction techniques in musculoskeletal CT-imaging, *Eur. J. Radiol.* 107 (2018) 60–69, <https://doi.org/10.1016/j.ejrad.2018.08.010>.
- [16] A. Blum, A. Noël, D. Regent, et al., Tomosynthesis in musculoskeletal pathology, *Diagn. Interv. Imaging* 99 (7-8) (2018) 423–441, <https://doi.org/10.1016/j.diii.2018.05.001>.
- [17] M.A. Ottenin, A. Jacquot, O. Grospretre, et al., Evaluation of the diagnostic performance of tomosynthesis in fractures of the wrist, *AJR Am. J. Roentgenol.* 198 (1) (2012) 180–186, <https://doi.org/10.2214/AJR.11.6374>.
- [18] A.S. Ha, A.Y. Lee, D.S. Hippe, S.H. Chou, F.S. Chew, Digital tomosynthesis to evaluate fracture healing: prospective comparison with radiography and CT, *AJR Am. J. Roentgenol.* 205 (1) (2015) 136–141, <https://doi.org/10.2214/AJR.14.13833>.
- [19] A. De Silvestro, K. Martini, A.S. Becker, et al., Postoperative imaging of orthopaedic hardware in the hand and wrist: is there an added value for tomosynthesis? *Clin. Radiol.* 73 (2) (2018), <https://doi.org/10.1016/j.crad.2017.08.001> 214.e1-214.e9.