Virtual Mechanical Testing Based on Low-Dose Computed Tomography Scans for Tibial Fracture

A Pilot Study of Prediction of Time to Union and Comparison with Subjective Outcomes Scoring

Hannah L. Dailey, PhD, Peter Schwarzenberg, MS, Charles J. Daly, BSc, Sinéad A.M. Boran, MB(BCh), BAO(BCh), FRCSI, Michael M. Maher, MD, FRCSI, and James A. Harty, MB(BCh), BAO, MSc, FRCSI

Investigation performed at Lehigh University, Bethlehem, Pennsylvania, and Cork University Hospital, University College Cork, Cork, Ireland

Background: Quantitative outcomes assessment remains a persistent challenge in orthopaedic trauma. Although patient-reported outcome measures (PROMs) and radiographic assessments such as Radiographic Union Scale for Tibial Fractures (RUST) scores are frequently used, very little evidence has been presented to support their validity for measuring structural bone formation or biomechanical integrity.

Methods: In this pilot study, a sequential cohort of patients with a tibial shaft fracture were prospectively recruited for observation following standard reamed intramedullary nailing in a level-I trauma center. Follow-up at 6, 12, 18, and 24 weeks included radiographs and completion of PROMs (EuroQol 5-Dimension [EQ-5D] and pain scores). Low-dose computed tomography (CT) scans were also performed at 12 weeks. Scans were reconstructed in 3 dimensions (3D) and subjected to virtual mechanical testing via the finite element method to assess torsional rigidity in the fractured limb relative to that in the intact bone.

Results: Patients reported progressive longitudinal improvement in mobility, self-care, activity, and health over time, but the PROMs were not correlated with structural bone healing. RUST scoring showed moderate intrarater agreement (intraclass coefficient [ICC] = 0.727), but the scores at 12 weeks were not correlated with the time to union ($R^2 = 0.104$, $p = 0.193$) and were only moderately correlated with callus structural integrity ($R^2 = 0.347$, $p = 0.010$). In contrast, patient-specific virtual torsional rigidity (VTR) was significantly correlated with the time to union ($R^2 = 0.383$, $p = 0.005$) and clearly differentiated 1 case of delayed union (VTR = 10%, union at 36 weeks) from the cases in the normally healing cohort (VTR > 60%; median union time, 19 weeks) on the basis of CT data alone.

Conclusions: PROMs provide insight into the natural history of the patient experience after tibial fracture but have limited utility as a measure of structural bone healing. RUST scoring, although reproducible, may not reliably predict time to union. In contrast, virtual mechanical testing with low-dose CT scans provides a quantitative and objective structural callus assessment that reliably predicts time to union and may enable early diagnosis of compromised healing.

Level of Evidence: Therapeutic Level IV. Please see Instructions for Authors for a complete description of levels of evidence.

Clinical research in orthopaedic trauma requires assessments that can capture the gradual progress of bone healing in the context of the patient's return to normal activity. Patient-reported outcome measures (PROMs) such as pain scores and multidimensional health status instruments are becoming increasingly central to robust trial design[1-3]. Unfortunately, interpreting PROM data can be challenging in trauma research. For example, in a trial of angular-stable nailing of distal tibial fractures, pain scores were highly variable and showed a strong trend toward pain decreasing over time

Disclosure: Funding to support study management was provided by a grant from OrthoXel, DAC. On the **Disclosure of Potential Conflicts of Interest** forms, *which are provided with the online version of the article*, one or more of the authors checked "yes" to indicate that the author had a relevant financial relationship in the biomedical arena outside the submitted work and "yes" to indicate that the author had a patent and/or copyright, planned, pending, or issued, broadly relevant to this work (http://links.lww.com/JBJS/F229).

without a detectable difference in pain between groups based on fixator type[4]. Similarly, pilot data for the Warwick Hip Trauma Evaluation (WHiTE) trial indicated that nearly 1,000 patients would be needed to detect a clinically relevant difference in EuroQol 5-Dimension (EQ-5D) scores between implant groups[5].

In contrast to PROMs, radiographic assessments offer the promise of objectivity, but historically they have been limited by concerns about reliability[6]. In response to this need, the Radiographic Union Scale for Tibial Fractures (RUST) was developed as a structured semiquantitative method for assessing callus. RUST scoring has demonstrable reliability[7,8], has been adapted for use for distal femoral fractures[9], has been used to diagnose fracture nonunion (a RUST score of ≥10 is given for union)[10-12], and has been adopted in the design of large randomized controlled trials[13]. Despite this wide use, we are not aware of any data on whether RUST scores are a reliable measure of structural bone healing in clinical (not preclinical) applications.

The objective of this study was to assess bone healing after tibial fracture by using a comprehensive suite of radiographic measures and PROMs and to critically evaluate these instruments with reference to a new objective measure of biomechanical integrity: the virtual torsional rigidity (VTR) of the fractured limb relative to that of the intact bone derived using patient-specific image-based finite element models based on low-dose computed tomography (CT) scans.

Materials and Methods

Study Design

A sequential cohort of adults (≥18 years of age) with a tibial shaft fracture were prospectively recruited from April 1, 2016, to December 31, 2016, at Cork University Hospital, a level-I trauma center serving southwest Ireland. All fractures were treated by reamed intramedullary nailing. Patients were permitted to bear weight as tolerated and were followed at 6, 12, 18, and 24 weeks or until clinical union, which was defined as radiographic evidence of union with pain-free ambulation and no requirement for continuing follow-up related to the tibial fracture. All cases were also reviewed at least 1 year after surgery, and clinical assessments were blinded with regard to findings of the CT-based analysis.

Outcome measures included EQ-5D and Numeric Rating Scale (NRS) pain scores, RUST scores, and quantitative CT-derived morphometric and structural measures of callus. To minimize bias, RUST scoring was completed by the senior clinician using blinded radiographs that were randomly

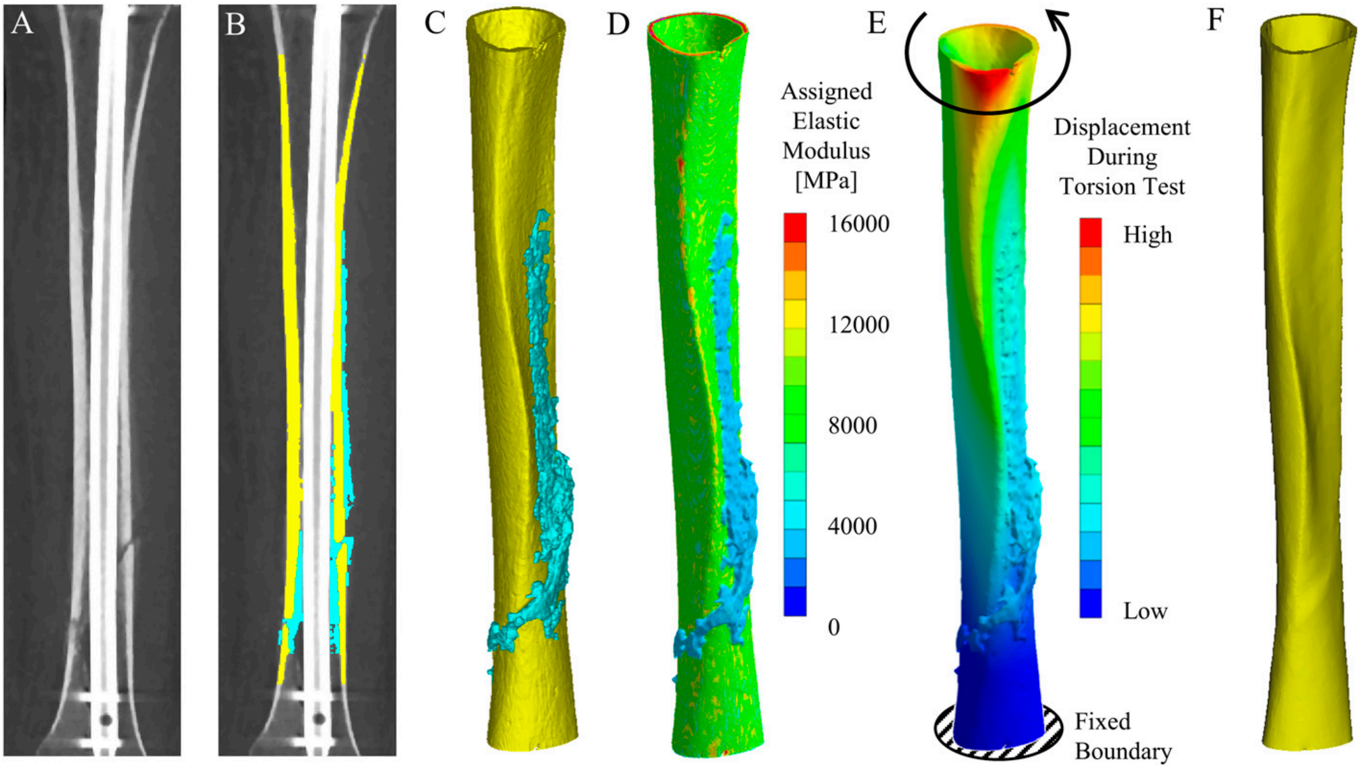

Fig. 1

Figs. 1-A through 1-F Workflow for virtual mechanical testing. **Fig. 1-A** Sagittal slice view of a low-dose CT scan without masks applied. **Fig. 1-B** Density-based threshold masks applied for cortical bone (yellow, 1,400 to 2,700 HU) and callus (cyan, 400 to 1,400 HU). **Fig. 1-C** 3D reconstruction after surface optimization. **Fig. 1-D** 3D tetrahedral volumetric model discretization with elementwise mechanical properties derived from voxel radiodensity. **Fig. 1-E** Virtual torsion test with the distal end clamped and the proximal end rotated about the anatomic axis of the tibia. **Fig. 1-F** Reconstructed intact tibia for comparison with the VTR of the fractured bone.

shuffled such that images from the different follow-up intervals were not presented in chronological order or grouped by patient. RUST scores were assigned to the blinded shuffled images by the same rater after a period of 4 months to assess intrarater agreement. None of the CT-derived measures were made available for comparison with the RUST scores or clinical findings until after all evaluations were completed.

This pilot study design was reviewed and approved by the local institutional review board, and all patients provided written informed consent.

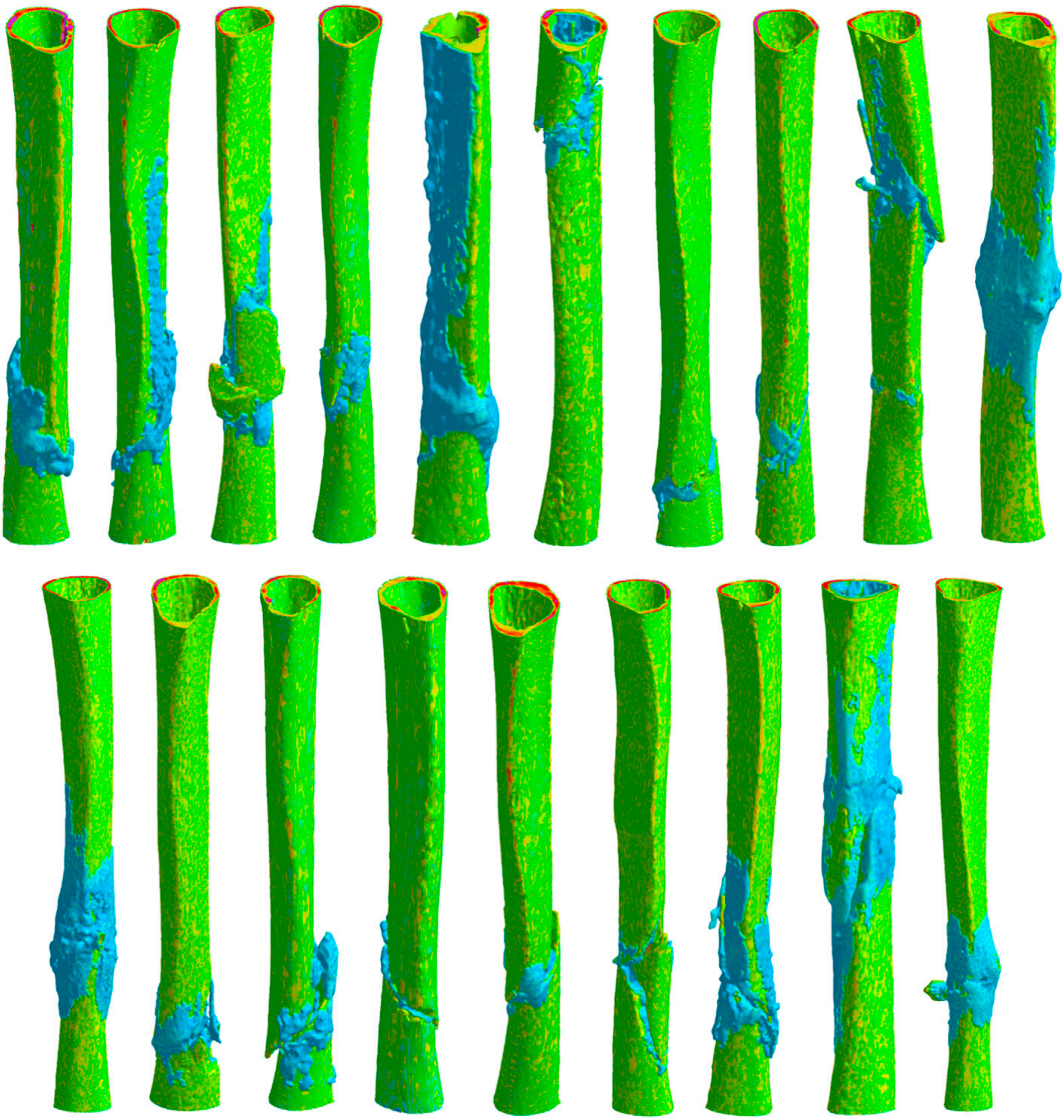

Fig. 2

Reconstructed models of all 19 cases. The colors indicate radiodensity, with lower density (blue) in the callus zone and higher density (yellow and red) in the cortex.

CT Scan Protocol

Patients underwent low-dose CT scanning at 12 weeks after surgery. CT scans were performed on a GE (General Electric Healthcare) Discovery CT750 HD scanner with x-ray tube voltage of 80 kV, current-time product of 10 mA, gantry rotation speed of 0.4 second, and pitch of 0.51. A pure iterative reconstruction algorithm, model-based iterative reconstruction (MBIR), was used for image reconstruction with a resolution improvement filter kernel, RP05[14,15]. The resulting scan resolution had a slice thickness of 0.625 mm and median pixel spacing of 0.373 mm.

CT Scan Processing and Reconstruction of Injured Limb

Image sets were processed using the Mimics Innovation Suite (version 20; Materialise). A segmentation workflow was developed to create 3-dimensional (3D) surface models by applying density-based threshold rules to identify different tissues. Callus included tissue within threshold values of 400 to 1,400 Hounsfield units (HUs), and cortical bone included tissue within 1,400 to 2,700 HU. Automated surface optimization sequences were applied to refine the segmented models, and then the surfaces were enclosed and volumetrically discretized to create a unified tetrahedral mesh for finite element analysis. The mesh was refined and elementwise mechanical properties were interpolated from voxel radiodensities using a previously published elastic modulus scaling law based on local HUs of the original CT scan: $E = 0.00704 \times HU$ (GPa)[16]. Each scan was also digitally reconstructed to recreate an intact bone model representing the preinjury anatomy. The complete workflow for generating the injured and intact models is summarized in Figure 1 and discussed in detail elsewhere[17].

A nondestructive virtual torsion test was chosen as the summary measure of mechanical integrity in each model. Torsional stiffness relative to intact paired controls is often used as a summary indicator of healing in preclinical models[18-22] because it is direction-independent[23]. Torsion testing yields virtual torsional rigidity: $VTR = ML/\phi$ (in N-m^2/°), where M is the twisting moment, L is the length of the test segment, and ϕ is the angle of twist. All structural simulations were carried out in ANSYS 17.2 (ANSYS).

Statistical Analysis

Descriptive statistics were generated using Microsoft Excel (2016) and MATLAB (R2016a; The MathWorks). All other analyses were carried out with SPSS Statistics 24 (IBM). Repeated-measures testing for differences over time was carried out using Friedman nonparametric tests as is appropriate for ordinal and non-normally distributed interval data. One-way repeated-measures analysis of variance (ANOVA) was also used for testing differences in normally distributed interval variables. Relationships between outcome measures were assessed using Pearson correlations. Normality was assessed with Shapiro-Wilk and Kolmogorov-Smirnov testing.

Results

Patient and Injury Characteristics

Twenty-four patients met the inclusion criteria and 19 completed the follow-up protocol, including undergoing

TABLE I Patient and Injury Characteristics

CT Case Number	Sex	Age *(yr)*	Injury Type	OTA/AO Classification	Gustilo-Anderson Classification	Tscherne Classification
CT01	Female	51	Closed	42A2		0
CT02	Male	52	Closed	42A1		0
CT03	Male	32	Open	42B3	I	
CT04	Male	32	Closed	42A3		1
CT05	Male	55	Closed	42A1		0
CT06	Male	58	Closed	42A3		3
CT07	Male	33	Closed	42A2		0
CT08	Female	39	Open	42B2	I	
CT09	Male	65	Closed	42C2		1
CT10	Male	33	Closed	42A3		0
CT11	Male	20	Closed	42A3		0
CT12	Male	45	Open	42A2	II	
CT13	Male	50	Closed	42A1		0
CT14	Male	24	Closed	42A1		0
CT15	Male	39	Closed	42A2		0
CT16	Male	29	Closed	42B2		1
CT17	Male	57	Closed	42B3		0
CT18	Male	53	Closed	42A1		1
CT19	Male	18	Closed	42A3		0

low-dose CT scanning (Fig. 2). Demographic and injury characteristics for these patients are shown in Table I. The median age was 39 years. Two patients were female, and 3 had open injuries. The predominant fracture morphology was OTA/AO type A[24] (15 cases). Injury mechanisms included sports (5 patients), a simple fall (4), a blow (3), a fall from a height (2), crush (2), pedestrian struck by a motor vehicle (2), and other (1). The median time to clinical union was 18.7 weeks. There was 1 delayed union, indicated by progressive autodynamization and eventual union at 36 weeks. There were no nonunions or reoperations to induce union.

PROMs

Complete descriptive statistics for all EQ-5D component scores and the NRS pain scores are shown in Table II, and selected statistically or clinically meaningful temporal trends

TABLE II Detailed Temporal Data for All PROMs

Score/Follow-up*	Median	Interquartile Range	Mean ± Standard Deviation	Friedman Test P Value*
EQ-5D mobility component†				**0.007**
6 wk[a]	3.0	2.0-4.3	3.1 ± 1.1	
12 wk	3.0	2.0-3.0	2.8 ± 1.0	
18 wk	2.0	2.0-3.0	2.5 ± 1.2	
24 wk[a]	2.0	1.0-2.5	2.2 ± 1.2	
EQ-5D self-care component†				**0.002**
6 wk[a]	2.0	2.0-2.3	2.1 ± 0.8	
12 wk	1.0	1.0-2.0	1.7 ± 0.9	
18 wk	1.0	1.0-1.0	1.3 ± 0.8	
24 wk[a]	1.0	1.0-1.0	1.3 ± 0.6	
EQ-5D usual activities component†				
6 wk[a,b,c]	4.0	3.0-5.0	3.8 ± 1.1	**<0.001**
12 wk[a]	3.0	2.0-4.0	2.9 ± 1.2	
18 wk[b]	2.3	2.0-3.0	2.7 ± 1.2	
24 wk[c]	2.0	2.0-3.0	2.3 ± 0.9	
EQ-5D pain/discomfort component†				0.666
6 wk	2.0	2.0-3.0	2.4 ± 0.7	
12 wk	3.0	2.0-3.0	2.4 ± 0.7	
18 wk	2.0	2.0-3.0	2.2 ± 0.7	
24 wk	2.0	2.0-3.0	2.3 ± 0.7	
EQ-5D anxiety/depression component†				0.252
6 wk	1.0	1.0-2.0	1.5 ± 0.8	
12 wk	1.0	1.0-1.0	1.5 ± 1.1	
18 wk	1.0	1.0-1.8	1.4 ± 0.7	
24 wk	1.0	1.0-1.3	1.3 ± 0.7	
EQ-5D health component‡				**0.040**
6 wk[a]	80	65-80	74 ± 12	
12 wk	80	75-90	79 ± 14	
18 wk	88	80-90	81 ± 15	
24 wk[a]	90	89-92	86 ± 15	
NRS pain§				0.324
6 wk	2.0	1.0-3.5	2.4 ± 2.1	
12 wk	2.0	1.0-4.0	2.7 ± 2.3	
18 wk	2.0	0.3-3.8	2.5 ± 2.4	
24 wk	2.5	0.8-4.0	2.4 ± 1.8	

*Significant differences between time points (least significant difference [LSD] post hoc) at the $p < 0.05$ level are indicated if the Friedman test showed significance overall (bolded p values). Time points that share a letter within a category were significantly different from one another. †1 = no problems, 2 = slight problems, 3 = moderate problems, 4 = severe problems, and 5 = unable. ‡0 = worst health you can imagine, and 100 = best health you can imagine. §0 = no pain, 1 to 3 = mild pain, 4 to 6 = moderate pain, and 7 to 10 = severe pain.

are illustrated in Figure 3. Patients tended to report significantly improved mobility, capacity for self-care, engagement in usual activities, and general health over time. Post-hoc testing showed that this trend was generally statistically significant only between the first and last time points. One exception was the usual activities score, with the greater difficulty at the 6-week follow-up being significantly different from the scores reported at all later time points. Pain scores were very low for most patients and were steady over time, both on the EQ-5D pain component subscale and on the NRS pain scale.

RUST Scores

RUST scores at the 4 follow-up intervals are also shown in Figure 3, together with time to union. The intraclass correlation coefficient, ICC (3,1), for test-retest reliability of RUST scoring with consistency effects (2-way random single measures procedure in SPSS) was 0.727 (95% confidence interval [CI] = 0.597 to 0.820). The RUST scores significantly increased over time (p = 0.001 overall) and, after 6 weeks, became non-normally distributed with a notable ceiling effect, with at least 75% of patients achieving a RUST score of ≥10 from 18 weeks onward.

CT-Derived Morphometry and Virtual Mechanical Testing

Complete morphometric data for each patient-specific model are provided in Table III, with selected values illustrated in Figure 4. For each patient-specific model, the VTR of the fractured limb was normalized by the VTR of the reconstructed intact tibia of the same patient to produce a dimensionless indicator of healing relative to the preinjury state (normalized VTR). Across all patients, the median normalized VTR was 99% (interquartile range [IQR] = 86% to 113%) at 12 weeks.

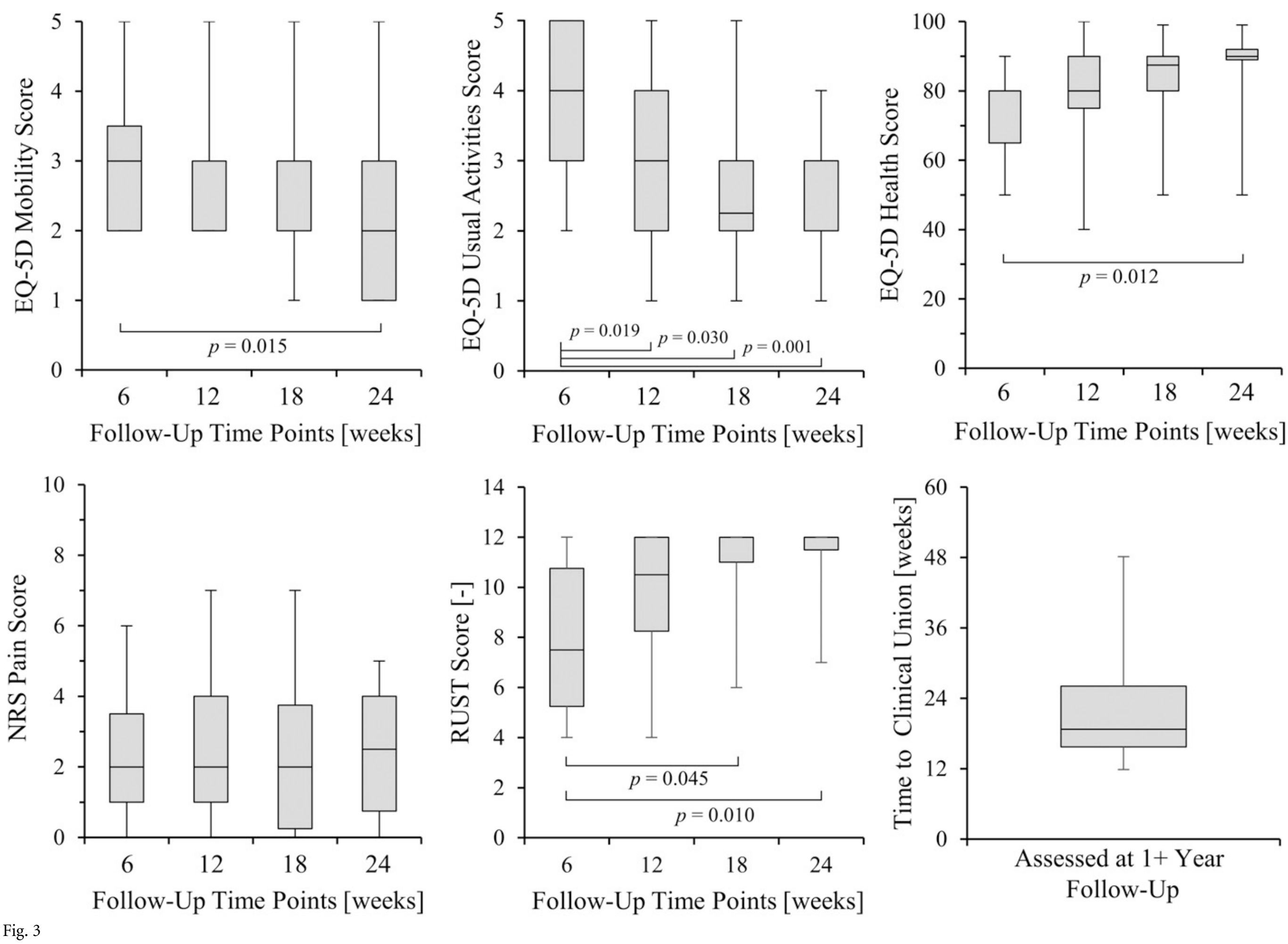

Fig. 3

Selected clinical findings at the 4 follow-up time points: 6, 12, 18, and 24 weeks. The EQ-5D mobility and usual activities components were recorded on an interval scale of 1 to 5 with 1 representing "no problems." The EQ-5D health component score was recorded on an interval visual analog scale (VAS) of 0 to 100 with 100 being "best health you can imagine" state. The NRS pain score was also recorded on an interval scale of 0 to 10 with 10 being "worst pain imaginable." The RUST scores and time to union were assessed by the senior clinician. Tops and bottoms of boxes = 25th and 75th percentiles, horizontal lines in boxes = medians, and whiskers = minimum and maximum.

TABLE III Complete Morphometric and Structural Data for All Patients

CT Case	Cortical Segment Length *(mm)*	Median Cortical Bone Density *(HU)*	Median Callus Density *(HU)*	Callus Volume *(cm³)*	VTR *(N-m²/°)*		Normalized VTR (Fractured/Intact)	12-Wk RUST Score from Radiograph
					Fractured	Intact		
CT01	202	2,121	806	7.4	1.48	1.28	1.16	12
CT02	269	2,059	694	14.2	3.66	3.21	1.14	8
CT03	263	2,034	743	9.0	1.63	2.68	0.61	8
CT04	275	2,062	700	5.3	2.41	2.43	0.99	11
CT05	189	1,737	839	15.3	1.94	1.48	1.31	12
CT06	230	1,926	738	8.2	1.37	1.68	0.82	11
CT07	291	1,940	782	9.1	4.06	3.70	1.10	12
CT08	265	2,053	772	8.1	1.73	1.73	1.00	5
CT09	261	1,984	767	32.2	2.82	3.31	0.85	11
CT10	201	1,916	925	14.8	1.99	1.73	1.15	12
CT11	251	2,017	784	23.6	2.61	2.09	1.25	12
CT12	256	1,954	706	11.8	2.71	2.86	0.95	9
CT13	265	1,961	681	8.8	0.19	1.85	0.10	4
CT14	231	1,868	781	2.3	2.09	1.88	1.11	9
CT15	264	2,069	793	4.0	2.34	2.59	0.90	5
CT16	268	2,103	763	8.7	2.60	3.20	0.81	10
CT17	259	1,929	692	13.0	1.68	1.91	0.88	12
CT18	213	1,732	948	29.1	1.61	1.84	0.88	*
CT19	276	1,922	833	13.7	2.83	2.56	1.11	10
Median	261	1,961	772	9.1	2.09	2.09	0.99	10.5

*This patient had a CT scan and PROMs at 12 weeks but had no 12-week radiographs; for reference, this individual's RUST score at 18 weeks was 12.

Correlations Between Outcomes Measures

Relationships among PROMs, RUST scores, and CT-derived properties at 12 weeks were assessed with Pearson correlations. The RUST scores were not significantly correlated with the CT-derived VTR (in N-m²/°) of either the fractured or the intact tibia (both $p > 0.217$), but the RUST scores had a moderate and significant positive correlation with the normalized (fractured relative to intact) VTR ($p = 0.010$; Fig. 5). Callus volume had no observable correlation with the normalized VTR ($R^2 = 0.012$, $p = 0.651$). Callus density showed a weak and non-significant

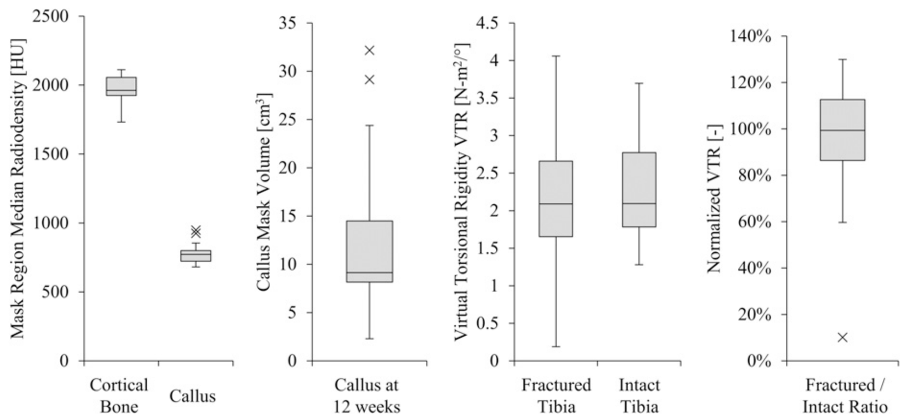

Fig. 4

CT-derived patient-specific morphometric and structural parameters, including (left to right) cortical bone and callus density, callus volume, VTR of the fractured and virtually reconstructed intact limbs, and normalized VTR of the fractured tibia relative to that of the reconstructed intact bone. At least 75% of the fractured limbs (Q1 and above) demonstrated at least 85% of the VTR of the intact bone by 12 weeks after surgery. Tops and bottoms of boxes = 25th and 75th percentiles, horizontal lines in boxes = medians, whiskers = minimum and maximum, and X = outliers.

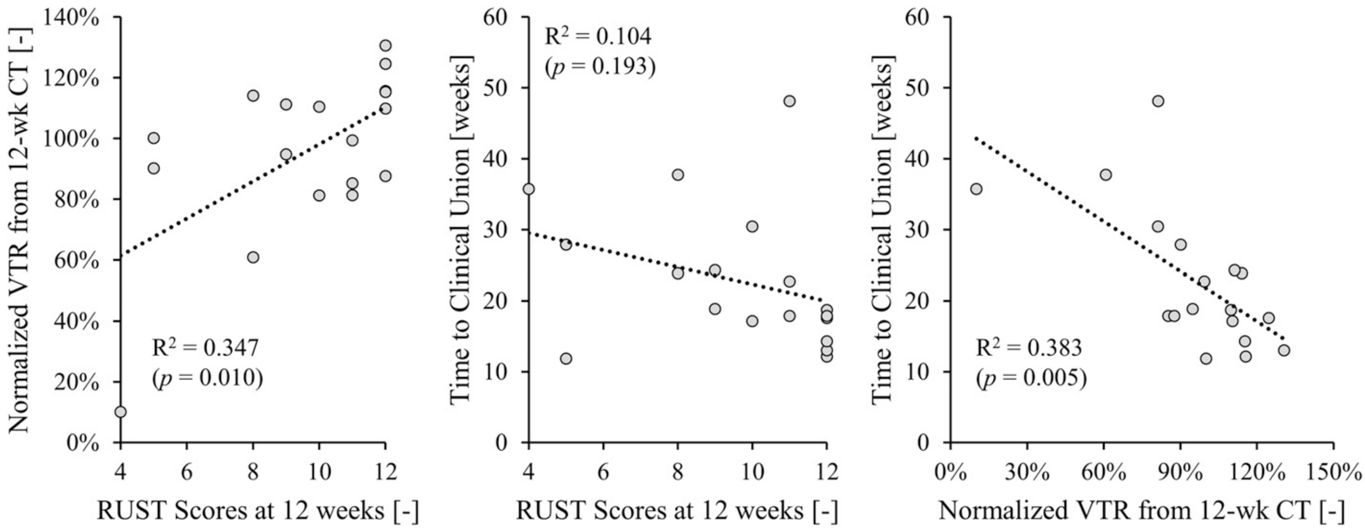

Fig. 5

Left to right: The RUST score was moderately and significantly correlated with the CT-derived normalized VTR at 12 weeks, the RUST score was not correlated with the time to clinical union, and the CT-derived normalized VTR was moderately and significantly correlated with the time to clinical union.

correlation with the normalized VTR ($R^2 = 0.167$, $p = 0.082$). None of the EQ-5D component scores or pain measures were correlated with the normalized VTR (all $R^2 < 0.051$, all nonsignificant). The RUST scores at 12 weeks were not significantly correlated with the time to union, but the normalized VTR at 12 weeks was moderately and significantly correlated with the time to clinical union (Fig. 5) and clearly differentiated 1 case of delayed union (VTR = 10%, union at 36 weeks) from the cases in the normally healing cohort (VTR > 60%; median union time, 19 weeks). One case (CT06 in Tables I and III) had a clinical follow-up at 48 weeks that was a statistical outlier and was not consistent with the radiographic and functional evidence of good healing (CT analysis and blinded RUST scores, with no implant failure or clinical concerns). Excluding this case from the correlations shown in Figure 5 produced the following results: VTR versus RUST score ($R^2 = 0.381$, $p = 0.008$), union time versus RUST score ($R^2 = 0.285$, $p = 0.027$), and union time versus VTR ($R^2 = 0.491$, $p = 0.001$).

Discussion

When evaluating this new clinical assessment method, normalized VTR, investigators must consider whether the technique produces data that are objective (free from bias), precise (repeatable and reproducible), systematically accurate (in agreement with true values, when known), and clinically useful[25]. In terms of objectivity, the workflow for processing CT scans and carrying out structural mechanics simulations follows a predefined sequence of actions applied uniformly for all models, with limited opportunity for introduction of researcher bias. The process has been documented in detail elsewhere[26]. The strict workflow control produces data that are highly repeatable, and any 2 individuals following the process would produce nearly identical summary measures.

The question of systematic accuracy with patient-specific virtual structural analysis is more difficult to resolve because the torsional rigidity of partially healed human fractures fixed with intramedullary nailing cannot be directly measured. The closest set of available reference data arises from published values for the torsional rigidity of intact cadaver tibiae, which was reported to average 2.42 ± 0.80 N-m²/° in physical mechanical testing[27]. For comparison, the mean VTR of the intact tibiae in our case series was 2.32 ± 0.68 N-m²/°, indicating that the modeling technique produces results very near the accepted reference value based on the limited data available. Systematic model bias may have been present in our study—for example in the choice of a mathematical scaling law for defining mechanical properties on the basis of radiodensity—but this source of error would be consistent across all models and is likely to be small in light of the close agreement between our results and the published cadaver test data[27].

Finally, we considered whether virtual structural analysis is clinically useful via an a priori power analysis based on all data collected at 12 weeks. For each outcome measure, we calculated the required sample size to detect clinically meaningful 20% differences between group means, assuming 80% power and a significance level of 0.05. On the basis of these data, significant differences between means would be detectable for the following measures and group sample sizes: NRS pain score (n = 273), EQ-5D mobility component score (n = 55), objective callus volume measured on CT (n = 151), RUST score (n = 30), and normalized VTR (n = 30). Notably, unlike in preclinical research, callus volume is not a useful measure of healing progress because the variety of clinical fracture morphologies leads to large variability in the amount of callus. Both the RUST score and the normalized VTR had similar expected statistical power, but the utility of RUST scoring should be considered carefully because of ceiling effects and known challenges with interpretation and repeatability[8]. Furthermore, in this analysis, VTR at 12 weeks successfully predicted time to union whereas the RUST score did not.

Aside from questions about statistical power, this study illustrates some concerning limitations of some assessment instruments that are commonly used to assess fracture-healing and outcomes. First, our data indicated that PROMs such as the EQ-5D may be likely to show a strong regression to the mean effect, in which patients report feeling better over time independent of their structural healing progress. Published normative population data from the EuroQol Group for individuals in the U.K. show that the median self-reported health score across all age groups is 90[28], which is precisely equal to the median health score reported by our participants at 24 weeks after surgery. The fact that none of the PROMs were significantly correlated with the normalized VTR suggests that PROMs provide insight into the functional response to injury but are not a reliable indicator of the physiological progress of fracture union for nailed tibial fractures.

This study also suggests some potential difficulties with the widely used RUST score. The ICC (0.727) for the RUST scores in our study was comparable with the repeatability recently reported for the RUST score by many other investigators[8,9,21,22,29,30], but it was not as good as that in the original reports[7]. Notably, none of these reports show that RUST scoring meets the strictest definition of ICC > 0.9 for an assessment to be deemed "reasonable for clinical measurements,"[31] even in studies using simple transverse osteotomies in animal models[21,22]. Furthermore, our data indicate that RUST scores may not have prognostic validity for predicting longitudinal outcomes, an issue that was raised early in the history of RUST scoring[32] but has been neglected in favor of repeated examination of interrater and intrarater reliability. Interpreted cautiously, our data support the general conclusion that the RUST is useful for differentiating between healed and unhealed fractures but may be lacking in sensitivity and specificity as a surrogate measure of callus structural integrity, particularly in early healing (a nonunited fracture as indicated by a RUST score of <10).

This study has a few noteworthy limitations. First, we did not include any objective activity monitoring, which would have provided a valuable comparison with the EQ-5D mobility and usual activities component scores. We also did not collect data at very early time points (<6 weeks). By 6 weeks, patients were reporting low levels of pain with a limited need for analgesia, so it is possible that more meaningful early pain data were missed. The time to clinical union also probably overestimates the actual union time and is subject to uncertainties if patients reschedule follow-ups for non-clinical reasons. Nevertheless, our median time to union (18.7 weeks) agrees well with that of large-scale epidemiological reports on reamed nailing (18 weeks)[33]. Finally, although the virtual mechanical test that we employed shows promise, this was a pilot study with a small sample size in a relatively low-risk cohort. Stronger inferences about diagnostic or prognostic validity, particularly for early nonunion detection, require further study.

In this inaugural proof-of-concept study, we showed that normalized virtual torsional rigidity (VTR) from CT-derived patient-specific structural models is predictive of the time to union, offers greater statistical power than PROMs, and avoids difficulties associated with RUST scoring such as repeatability and ceiling effects. Patient-specific modeling does incur increased investigational burden compared with using only PROMs or standard radiographic measures, but in exchange it offers precision and accuracy that cannot be achieved by those routine assessments. Accordingly, virtual mechanical testing merits future evaluation as an outcome measure in clinical trials that may require enhanced sensitivity and objectivity to enable comparison between groups. ■

NOTE: The authors gratefully acknowledge the supporting contributions from Karl James, MB, FFR(RCSI); Richard Kavanagh, MB, FFR(RCSI); and Owen J. O'Connor, MB, FFR(RCSI).

Hannah L. Dailey, PhD[1]
Peter Schwarzenberg, MS[1]
Charles J. Daly, BSc[2]
Sinéad A.M. Boran, MB(BCh), BAO(BCh), FRCSI[3]
Michael M. Maher, MD, FRCSI[3]
James A. Harty, MB(BCh), BAO, MSc, FRCSI[3]

[1]Department of Mechanical Engineering & Mechanics, Lehigh University, Bethlehem, Pennsylvania

[2]OrthoXel, DAC, Cork, Ireland

[3]Departments of Radiology (M.M.M.) and Orthopaedic Surgery (S.A.M.B. and J.A.H.), Cork University Hospital, University College Cork, Wilton, Cork, Ireland

E-mail address for H.L. Dailey: hannah.dailey@lehigh.edu

ORCID iD for H.L. Dailey: 0000-0002-4574-800X
ORCID iD for P. Schwarzenberg: 0000-0002-8747-6786
ORCID iD for C.J. Daly: 0000-0002-8778-2513
ORCID iD for S.A.M. Boran: 0000-0001-8619-9791
ORCID iD for M.M. Maher: 0000-0001-7423-7439
ORCID iD for J.A. Harty: 0000-0001-8437-4357

References

1. Gagnier JJ. Patient reported outcomes in orthopaedics. J Orthop Res. 2017 Oct;35(10):2098-108. Epub 2017 Jun 13.
2. Christensen DL, Dickens JF, Freedman B, Mauntel T, Owens BD, Potter BK, Provencher M, Tokish JM, Waterman BR, Antosh I, Bellamy J, Cameron KL, Cook KF, Eckel TT, Eichinger JK, Garcia ESJ, Helgeson M, Joyner PW, Kang DG, Kilcoyne KG, Krueger CA, LeClere LE, Li X, Martin KD, McCallum JR, McDonald LS, Ortiz D III, Pallis MP, Parada SA, Patzkowski JC, Posner MA, Rivera JC, Robins RJ, Rothrock N, Schmitz MR, Slabaugh M, Song DJ, Stinner DJ, Tintle SM, Tucker CJ, Waltz RA, Wilson K; MOTION Group. Patient-reported outcomes in orthopaedics. J Bone Joint Surg Am. 2018 Mar 7;100(5):436-42.
3. Hamilton DF, Giesinger JM, Giesinger K. It is merely subjective opinion that patient-reported outcome measures are not objective tools. Bone Joint Res. 2017 Dec;6(12):665-6.
4. Höntzsch D, Schaser KD, Hofmann GO, Pohlemann T, Hem ES, Rothenbach E, Krettek C, Attal R. Evaluation of the effectiveness of the angular stable locking system in patients with distal tibial fractures treated with intramedullary nailing: a multicenter randomized controlled trial. J Bone Joint Surg Am. 2014 Nov 19;96(22):1889-97.
5. Griffin XL, Parsons N, McArthur J, Achten J, Costa ML. The Warwick Hip Trauma Evaluation One: a randomised pilot trial comparing the X-Bolt Dynamic Hip Plating System with sliding hip screw fixation in complex extracapsular hip fractures: WHiTE (One). Bone Joint J. 2016 May;98-B(5):686-9.

6. Bhandari M, Guyatt GH, Swiontkowski MF, Tornetta P 3rd, Sprague S, Schemitsch EH. A lack of consensus in the assessment of fracture healing among orthopaedic surgeons. J Orthop Trauma. 2002 Sep;16(8):562-6.

7. Whelan DB, Bhandari M, Stephen D, Kreder H, McKee MD, Zdero R, Schemitsch EH. Development of the radiographic union score for tibial fractures for the assessment of tibial fracture healing after intramedullary fixation. J Trauma. 2010 Mar;68(3):629-32.

8. Leow JM, Clement ND, Tawonsawatruk T, Simpson CJ, Simpson AHRW. The radiographic union scale in tibial (RUST) fractures: Reliability of the outcome measure at an independent centre. Bone Joint Res. 2016 Apr;5(4):116-21.

9. Litrenta J, Tornetta P 3rd, Mehta S, Jones C, O'Toole RV, Bhandari M, Kottmeier S, Ostrum R, Egol K, Ricci W, Schemitsch E, Horwitz D. Determination of Radiographic Healing: An Assessment of Consistency Using RUST and Modified RUST in Metadiaphyseal Fractures. J Orthop Trauma. 2015 Nov;29(11):516-20.

10. Bishop JA, Palanca AA, Bellino MJ, Lowenberg DW. Assessment of compromised fracture healing. J Am Acad Orthop Surg. 2012 May;20(5):273-82.

11. O'Halloran K, Coale M, Costales T, Zerhusen T Jr, Castillo RC, Nascone JW, O'Toole RV. Will my tibial fracture heal? predicting nonunion at the time of definitive fixation based on commonly available variables. Clin Orthop Relat Res. 2016 Jun;474(6):1385-95.

12. van Houten AH, Heesterbeek PJC, van Heerwaarden RJ, van Tienen TG, Wymenga AB. Medial open wedge high tibial osteotomy: can delayed or nonunion be predicted? Clin Orthop Relat Res. 2014 Apr;472(4):1217-23. Epub 2013 Nov 19.

13. Castillo R, Raja S. METRC. Improving pain management in high energy orthopaedic truama. https://metrc.org/54-infopages/crsd/292-painstudydescription. Accessed 21 Jun 2017.

14. Li G, Liu X, Dodge CT, Jensen CT, Rong XJ. A noise power spectrum study of a new model-based iterative reconstruction system: Veo 3.0. J Appl Clin Med Phys. 2016 Sep 8;17(5):428-39.

15. Li K, Tang J, Chen GH. Statistical model based iterative reconstruction (MBIR) in clinical CT systems: experimental assessment of noise performance. Med Phys. 2014 Apr;41(4):041906.

16. Snyder SM, Schneider E. Estimation of mechanical properties of cortical bone by computed tomography. J Orthop Res. 1991 May;9(3):422-31.

17. Schwarzenberg P, Maher MM, Harty JA, Dailey HL. Virtual structural analysis of tibial fracture healing from low-dose clinical CT scans. J Biomech. 2019 Jan 23;83:49-56. Epub 2018 Nov 20.

18. Epari DR, Kassi JPP, Schell H, Duda GN. Timely fracture-healing requires optimization of axial fixation stability. J Bone Joint Surg Am. 2007 Jul;89(7):1575-85.

19. Kaspar K, Schell H, Seebeck P, Thompson MS, Schütz M, Haas NP, Duda GN. Angle stable locking reduces interfragmentary movements and promotes healing after unreamed nailing. Study of a displaced osteotomy model in sheep tibiae. J Bone Joint Surg Am. 2005 Sep;87(9):2028-37.

20. Bottlang M, Tsai S, Bliven EK, von Rechenberg B, Klein K, Augat P, Henschel J, Fitzpatrick DC, Madey SM. Dynamic stabilization with active locking plates delivers faster, stronger, and more symmetric fracture-healing. J Bone Joint Surg Am. 2016 Mar 16;98(6):466-74.

21. Cooke ME, Hussein AI, Lybrand KE, Wulff A, Simmons E, Choi JH, Litrenta J, Ricci WM, Nascone JW, O'Toole RV, Morgan EF, Gerstenfeld LC, Tornetta P 3rd. Correlation between RUST assessments of fracture healing to structural and biomechanical properties. J Orthop Res. 2018 Mar;36(3):945-53. Epub 2017 Sep 20.

22. Litrenta J, Tornetta P 3rd, Ricci W, Sanders RW, O'Toole RV, Nascone JW, Faber H, Wilson D. O'Toole RV, Nascone JW, Faber H, Wilson D. In vivo correlation of radiographic scoring (radiographic union scale for tibia fractures) and biomechanical data in a sheep osteotomy model: can we define union radiographically? J Orthop Trauma. 2017 Mar;31(3):127-30.

23. Augat P, Penzkofer R, Nolte A, Maier M, Panzer S, v Oldenburg G, Pueschl K, Simon U, Bühren V. Interfragmentary movement in diaphyseal tibia fractures fixed with locked intramedullary nails. J Orthop Trauma. 2008 Jan;22(1):30-6.

24. Kellam JF, Meinberg EG, Agel J, Karam MD, Roberts CS. Introduction: Fracture and Dislocation Classification Compendium-2018: International Comprehensive Classification of Fractures and Dislocations Committee. J Orthop Trauma. 2018 Jan;32(Suppl 1):S1-10.

25. Langlois J, Hamadouche M. Current recommendations for assessing the reliability of a measurement tool: a survival guide for orthopaedic surgeons. Bone Joint J. 2016 Feb;98-B(2):166-72.

26. Schwarzenberg P. Virtual structural analysis of bone fracture healing from low-dose clinical CT scans [Thesis]. 2018 May.

27. Heiner AD. Structural properties of fourth-generation composite femurs and tibias. J Biomech. 2008 Nov 14;41(15):3282-4. Epub 2008 Oct 1.

28. Szende A, Janssen B, Cabasés J, editors. Self-reported population health: an international perspective based on EQ-5D. 1st ed. Dordrecht: Springer; 2014.

29. Azevedo Filho FA, Cotias RB, Azi ML, Teixeira AA. Reliability of the radiographic union scale in tibial fractures (RUST). Rev Bras Ortop. 2016 Dec 15;52(1):35-9.

30. Perlepe V, Cerato A, Putineanu D, Bugli C, Heynen G, Omoumi P, Berg BV. Value of a radiographic score for the assessment of healing of nailed femoral and tibial shaft fractures: A retrospective preliminary study. Eur J Radiol. 2018 Jan;98(98):36-40. Epub 2017 Nov 7.

31. Trevethan R. Intraclass correlation coefficients: clearing the air, extending some cautions, and making some requests. Health Serv Outcomes Res Methodol. 2017 Jun;17(2):127-43.

32. Kooistra BW, Dijkman BG, Busse JW, Sprague S, Schemitsch EH, Bhandari M. The radiographic union scale in tibial fractures: reliability and validity. J Orthop Trauma. 2010 Mar;24(Suppl 1):S81-6.

33. Dailey HL, Wu KA, Wu PS, McQueen MM, Court-Brown CM. Tibial fracture nonunion and time to healing after reamed intramedullary nailing: risk factors based on a single-center review of 1003 patients. J Orthop Trauma. 2018 Jul;32(7):e263-9.